Constr. Approx. (1993) 9: 1–21

CONSTRUCTIVE APPROXIMATION

Rational Approximation in the Real Hardy Space H_2 and Stieltjes Integrals: A Uniqueness Theorem

L. Baratchart and F. Wielonsky

Abstract. The paper deals with rational approximation over the real Hardy space $H_{2,\mathbf{R}}(V)$, where V is the complement of the closed unit disk. The results concern Stieltjes functions

$$f(z) = \int \frac{d\mu(t)}{z - t},$$

where μ is a positive measure. It is shown that there is a unique critical point and hence both a unique local and a unique global best rational approximation in each degree, provided the support of μ lies within some absolute bounds which are explicitly estimated.

1. Introduction

The goal of this paper is to prove that Stieltjes functions of the form

$$f(z) = \int \frac{d\mu(t)}{z - t}, \tag{1}$$

with μ a finite nonnegative Borel measure on $\mathbf{R}$ whose support lies within $[-\lambda_0, \lambda_0]$, where λ_0 is the unique positive real number smaller than 1 such that

$$(1 - \lambda_0^2)^2 - 2\lambda_0^2 = 0,$$

have, for any positive integer n, a unique local (hence global) best approximation among rational functions of degree at most n in the real Hardy space $H_{2,\mathbf{R}}(V)$, where $V := \{z \in \mathbf{C}; |z| > 1\}$. Note that $\lambda_0 < 1$ so that f indeed belongs to $H_{2,\mathbf{R}}(V)$. Note also that λ_0 is *independent* of n, which is perhaps unexpected. Numerical estimation yields $\lambda_0 = 0.5176\ldots$. When μ is one-sided, that is when the support of μ lies entirely within $\mathbf{R}^+$ or $\mathbf{R}^-$, the value for λ_0 can be improved to

Date received: May 3, 1991. Date revised: April 7, 1992. Communicated by Dietrich Braess.
AMS classification: 30E10.
Key words and phrases: Rational approximation in H_2, Critical points, Uniqueness of best approximation, Stieltjes functions.

$1/\sqrt{2} \simeq 0.7071\ldots$. It is not difficult to design an even measure μ (hence defining some odd f) such that no best approximation of degree 1 has a pole at zero. A simple computation further shows that the support of such a μ may be arbitrary in $(-1, 1)$, provided it is not included in $[-1/\sqrt{2}, 1/\sqrt{2}]$. By symmetry, there must then be two distinct best approximations of degree 1 to f, so that the sharp value for λ_0 cannot exceed $1/\sqrt{2}$ as was pointed out to the authors by D. Braess. When μ is a positive measure on $(0, 1)$, it has also been shown by E. B. Saff [24] that the best upper bound on the support for uniqueness of best approximation in degree 1 is $2\sqrt{2}/3 = 0.9428\ldots$. This number is *a fortiori* an upper bound on the sharp value of λ_0 in the one-sided case.

What precedes has been stated in terms of $H_{2,\mathbf{R}}(V)$, instead of the more familiar real Hardy space of the disk, because this is the context in which Stieltjes functions appear. It can be translated at once into a uniqueness property for rational approximation to functions of the form

$$g(z) = \int \frac{d\mu(t)}{1 - zt}. \tag{2}$$

The above results are obtained by combining the index theorem appearing in [4] and [6] with a direct computation that enables us to estimate second derivatives in the special case of Stieltjes functions. The next two sections contain differential preliminaries and, in particular, the above-mentioned theorem (stated here as Theorem 2) essentially asserting that

$$\sum_{c_i} (-1)^{\varepsilon_i} = 1,$$

where the sum ranges over all the critical points c_i of the H_2-norm and where ε_i is the Morse index of the critical point c_i, provided f is smooth in a neighborhood of the unit circle T. In Section 4 we derive the main result of the paper, namely that all critical points are local minima if f is a Stieltjes function satisfying the conditions above. More precisely, the critical points have Morse index zero. From the index theorem, uniqueness of the best approximation is thus established. Moreover, since there is a unique local minimum, a gradient algorithm as described in [5] is globally convergent so that the best approximation can be numerically estimated. Stieltjes functions are already well-known for their nice behavior in important areas of approximation theory, for instance Padé approximation [1], and our result only adds to this picture.

After they proved the above property, the authors became aware that a completely similar approach was taken by D. Braess and N. Dyn to tackle the problem of uniqueness of generalized monosplines of least L_p norm [11], [13], thereby extending to every p such that $1 \leq p < \infty$ some of their previous results obtained by a fix-point method for $p = 1$ and $p = 2$ [12], [16]. We comment on this in the last section. This use of degree theory was also in the spirit of some earlier work by B. D. Bojanov on algebraic monosplines [10]. This interplay between topology and function theory is, in our opinion, an attractive feature of the subject.

2. A Rational H_2 Approximation Problem

Let T be the unit circle, and consider the *real* Hilbert space $L_{2,\mathbf{R}}(T)$ consisting of square summable functions with real Fourier coefficients. The scalar product can be expressed as a line integral:

$$\langle f, g\rangle = \frac{1}{2i\pi}\int_T f(z)g\left(\frac{1}{z}\right)\frac{dz}{z}. \tag{3}$$

We consider two closed subspaces in $L_{2,\mathbf{R}}(T)$. The first is the real Hardy space $H_{2,\mathbf{R}}(U)$ of the open unit disk U, consisting of analytic functions of the form

$$h(z) = \sum_{k=0}^{\infty} h_k z^k,$$

where $h_k \in \mathbf{R}$ and the squared norm $\|h\|^2 = \sum_k h_k^2$ is finite. The second is the real Hardy space $H_{2,\mathbf{R}}(V)$ of the complement of the disk, consisting of functions f analytic in V and at infinity

$$f(z) = \sum_{k=0}^{\infty} f_k z^{-k},$$

with $f_k \in \mathbf{R}$ and $\|f\|^2 = \sum_k f_k^2 < \infty$. Changing z into $1/z$ clearly defines an isometry between our two Hardy spaces. In the following we assume that a real analytic function is one taking real values for real arguments. Hence, the above requirement that h_k and f_k be real means that $h(z)$ and $f(z)$ are real.

We further single out the closed subspace $H^0_{2,\mathbf{R}}(V)$ of $H_{2,\mathbf{R}}(V)$ consisting of those functions vanishing at infinity, or equivalently such that the coefficient f_0 in the above expansion is zero.

The degree of a rational function p/q, where p and q are coprime polynomials, is defined to be $\max\{\deg(p), \deg(q)\}$. The set of rational functions of degree at most n in $H^0_{2,\mathbf{R}}(V)$ is denoted by $\mathscr{R}^0_n(V)$. Note that a rational function p/q belongs to $\mathscr{R}^0_n(V)$ if and only if p and q are real polynomials, q has all its roots in U, and $\deg(p) < \deg(q)$. Thus, $\deg(p/q) = \deg(q)$ in our case. The approximation problem that we consider in the following can be stated as follows:

Given $f \in H^0_{2,\mathbf{R}}(V)$ and some integer $n > 0$, find some rational function $p/q \in \mathscr{R}^0_n(V)$ that minimizes $\|f - p/q\|$.

This problem is easily seen to be equivalent to the corresponding problems in $H_{2,\mathbf{R}}(U)$ and $H_{2,\mathbf{R}}(V)$ (see, e.g., [7]).

We only consider real analytic functions though this approximation problem can be stated in the full Hardy space and everything in the next two sections could be extended with minor modifications. Note, however, that the best complex approximation to a real function may fail to be real. In fact, the authors' interest in this question originally stemmed from identification problems for linear dynamical control systems, where it is essential to find real rational approximations to functions that are themselves real [21]. It is perhaps worth noting that Stieltjes functions admit a system-theoretic interpretation as transfer functions of so-called relaxation systems [26], for instance RC circuits.

Let us briefly describe a few known results concerning this problem, which are more or less standard for any L^2 rational approximation problem. First, a best approximation does, indeed, exist. While the proof in the complex case seems to go back to [25], the first reference in our situation that we know of is [15]. Second, a best approximation is not always unique. Examples can be given where at least two best approximations exist, by letting f enjoy a symmetry property [23]. In the complex case, using rotation invariance, we could even exhibit some f having infinitely many best approximations [14], but such examples are not known in the real case. Examples without symmetry can also be obtained by transposing to our situation Theorem 1.6 in Chapter 5 of [11]. More precisely, we get that any $(n+1)$-subspace of $H^0_{2,\mathbf{R}}(V)$ comprising no member of $\mathscr{R}^0_n(V)$ except for 0, contains some function with at least two best approximants. An instance of such a subspace is the span of $r_1, r_2, \ldots, r_{n+1}$ where the r_i's are rational functions of degree greater than n whose denominators are pairwise coprime. The proof is another illustration of the use of topology in this context, since it appeals to the Borsuk–Ulam theorem. It can be carried out as in [11], using Propositions 1 and 5 of [2]. Third, the problem is normal [23], meaning that any best approximation of degree at most n to a function which is not itself rational of degree less than n is in fact of degree precisely n. This property, which we call the *order lemma*, is in fact true for local *minima* and *maxima* [7]. As for the rate of approximation, we refer the reader to [19] and the bibliography therein.

Our main concern, in this section, is to compute the optimal numerator p from f and q, so as to end up with a criterion depending on q only. We first define some notation.

Let $\mathbf{P}_k[z]$ denote the space of real polynomials of degree at most k, and let $\mathscr{M}_k$ be the subset of monic polynomials of degree k. Let $\mathscr{M}^1_k \subset \mathscr{M}_k$ consist of those polynomials whose roots lie in U. For $q \in \mathscr{M}^1_n$, we set

$$V_q := \frac{\mathbf{P}_{n-1}[z]}{q}.$$

For fixed q, it is clear that the smallest value for $\|f - p/q\|$ is obtained when p/q is the orthogonal projection of f onto the n-dimensional linear subspace V_q of $H^0_{2,\mathbf{R}}(V)$. We denote this projection by L_q/q, where $L_q \in \mathbf{P}_{n-1}[z]$, and seek the minimum with respect to q of the following function:

$$\psi_n(q) := \left\| f - \frac{L_q}{q} \right\|^2.$$

To compute L_q more precisely, we need two more pieces of notation. For $p \in \mathbf{P}_k[z]$, we define its reciprocal polynomial in $\mathbf{P}_k[z]$ by

$$\tilde{p}(z) := z^k p\left(\frac{1}{z}\right).$$

We offer a word of warning about this notation. If $k' > k$ and some member p of

$\mathbf{P}_k[z]$ is considered as a member of $\mathbf{P}_{k'}[z]$ whose leading coeffcients do vanish, the two definitions of $\tilde{p}$ may be inconsistent. For this reason, we always specify which $\mathbf{P}_k[z]$ is involved in the process.

For any function f, we set

$$f^\sigma(z) := \frac{1}{z} f\left(\frac{1}{z}\right).$$

Clearly, $f \to f^\sigma$ maps $H_{2,\mathbf{R}}^0(V)$ isometrically onto $H_{2,\mathbf{R}}(U)$ and conversely.

If g is holomorphic in U and $q \in \mathscr{M}_n^1$, we have the well-known division formula

$$g = vq + r, \tag{4}$$

where the quotient v is again holomorphic in U and the remainder r is the unique polynomial in $\mathbf{P}_{n-1}[z]$ interpolating g at the zeros of q. The quotient and the remainder can also be represented by the Hermite integral formulae [25]:

$$v(z) = \frac{1}{2i\pi} \int_{T_\rho} \frac{g(\xi)}{q(\xi)} \frac{d\xi}{\xi - z}, \tag{5}$$

$$r(z) = \frac{1}{2i\pi} \int_{T_\rho} \frac{g(\xi)}{q(\xi)} \left[\frac{q(\xi) - q(z)}{\xi - z}\right] d\xi, \tag{6}$$

where T_ρ is a circle centered at 0 of radius $\rho < 1$ encompassing z and all the roots of q. From these formulae, it is obvious that v and r are real if g and q are real. Moreover, since $1/q(z)$ is bounded for $|z| \geq \rho$, it is evident from (4) that v belongs to $H_{2,\mathbf{R}}(U)$ if g does. Our first proposition is just a reformulation of a classical result due to Walsh [25, Chapter 9, Theorem 2] in terms that do not involve the roots of q. A proof is given in [6].

Proposition 1. *For $f \in H_{2,\mathbf{R}}^0(V)$ and $q \in \mathscr{M}_n^1$, let us denote the division of $f^\sigma\tilde{q}$ by q:*

$$f^\sigma \tilde{q} = v_q q + r,$$

where $r \in \mathbf{P}_{n-1}[z]$ and $v_q \in H_{2,\mathbf{R}}(U)$. Then $L_q = \tilde{r}$.

Proposition 1 is often nicely rephrased in terms of interpolation by saying that, for fixed denominator q, the best approximation to f interpolates f at the reciprocal of the roots of q.

For later use, we derive from the previous proposition the following corollary.

Corollary 1. *Using the notation of Proposition* 1, *assume f is orthogonal to V_q. Then the function $z^n f(z)/\tilde{q}(z)$ belongs to $H_{2,\mathbf{R}}^0(V)$.*

Proof. The fact that f is orthogonal to V_q simply means that $L_q = 0$. Therefore, it follows from Proposition 1 that $f^\sigma/q = v_q/\tilde{q}$ belongs to $H_{2,\mathbf{R}}(U)$. Upon changing z into $1/z$ and dividing by z, we get the desired conclusion. ■

We now turn to differentiation. We consider $\mathscr{M}_n$ as being $\mathbf{R}^n$ by identifying the polynomial

$$q(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$$

with the vector $(a_{n-1}, a_{n-2}, \ldots, a_0)$. In this way, $\mathscr{M}_n^1$ becomes an open subset of $\mathbf{R}^n$. Similarly, we identify $\mathbf{P}_k[z]$ and $\mathbf{R}^{k+1}$ and formula (6) shows at once, by differentiating under the integral sign, that the map

$$L: \mathscr{M}_n^1 \to \mathbf{P}_{n-1}[z] \quad \text{given by} \quad q \to L_q$$

is a smooth map. Consequently,

$$\psi_n: \mathscr{M}_n^1 \to \mathbf{R}$$

is also smooth. By definition, a *critical point* of ψ_n will be any $q \in \mathscr{M}_n^1$ such that the derivative $D\psi_n(q)$ vanishes. Critical points fall into two classes: they are called irreducible if L_q is prime to q, and reducible otherwise. As usual, a critical point is said to be nondegenerate if the second derivative at this point is a nondegenerate quadratic form on $\mathbf{R}^n$. In this case the signature of this quadratic form, that is the number of negative eigenvalues of the Hessian matrix, is called the *Morse index* of the critical point.

We sometimes have to deal with critical points of ψ_k for several values of k. Therefore, we often say that q is critical, meaning that it is a critical point of $\psi_{\deg(q)}$. It is clear that the (monic) denominator of any best approximation to f lies among the critical points. The order lemma tells us, furthermore, that this is an irreducible critical point unless $f \in \mathscr{R}_{n-1}^0(V)$. If it is nondegenerate, its Morse index is obviously 0. However, there may be many other critical points: any *local minimum*, for instance, is critical, and again its Morse index is 0 if it is nondegenerate. Critical points are, of course, easier to analyze than best approximations, because they are defined locally. The next section is devoted to the set of critical points.

A final remark in this section is that our definition of critical points looks different from the natural one, namely that (p_0, q_0) is critical if the map $\mathbf{P}_{n-1}[z] \times \mathscr{M}_n^1 \to \mathbf{R}$ given by

$$(p, q) \to \left\| f - \frac{p}{q} \right\|^2$$

has vanishing derivative at (p_0, q_0). It is, however, easy to see that any critical point (p_0, q_0) of the above map is of the form (L_{q_0}, q_0) where q_0 is critical in our sense, and conversely. Hence, the two definitions are equivalent. It is only slightly more demanding to check that our definition also agrees with the general definition, using tangent cones if p_0 and q_0 are not coprime [11], of a critical point with respect to $\mathscr{R}_n^0(V) \subset H_{2,\mathbf{R}}^0(V)$.

3. Critical Points

Our next goal is to give a characterization of critical points in terms of the division introduced before. To this end we recall some results (see Proposition 2 of [7]):

Theorem 1. *Let q belong to $\mathcal{M}_n^1$, and let*

$$f^\sigma \tilde{q} = v_q q + \tilde{L}_q \tag{7}$$

be the division of $f^\sigma \tilde{q}$ by q. The following two assertions are equivalent:

(i) *q is a critical point of ψ_n.*
(ii) *q divides $v_q L_q$.*

In particular, if q is an irreducible critical point, then q divides v_q.

Like Proposition 1, part (ii) of the theorem can be rephrased in terms of interpolation. When applied to the minimum of Ψ_n, it says that a best approximation to f with free denominator q interpolates f at the reciprocal of the roots of q with order at least two.

Concerning reducible critical points, we complement Proposition 1 as follows (see Proposition 3 of [7]):

Proposition 2. *Let $q \in \mathcal{M}_n^1$ be such that*

$$L_q = ds, \qquad q = dp,$$

with s/p irreducible and $d \in \mathcal{M}_k^1$. Then $s = L_p$ and the following two assertions are equivalent:

(i) *q is a critical point of ψ_n.*
(ii) *p is a critical point of ψ_{n-k} and the polynomial d divides v_p/p.*

To obtain the results of Section 4, we need to know that the set of critical points satisfies some arithmetic constraint, namely the index theorem. This requires an extra assumption concerning the smoothness of f, which is to the effect that we shall be able to replace $\mathcal{M}_n^1$ by a *compact* subset of $\mathcal{M}_n$ as follows. Let Δ_n denote the closure, in $\mathcal{M}_n$, of $\mathcal{M}_n^1$. It is clear that Δ_n is the compact set consisting of those monic polynomials whose roots are of modulus at most 1. Proposition 2 of [6] allows us to extend the domain of ψ_n under a certain condition:

Proposition 3. *Assume there exists $\eta > 0$ such that $f(z)$ is analytic for $|z| > 1 - \eta$. Then the map ψ_n extends smoothly to a neighborhood of Δ_n in $\mathcal{M}_n$.*

The argument in [6] is based on the homotopy invariance of the Cauchy integral (6) which allows us to choose ρ slightly greater that 1, thereby defining a smooth extension of r. Note, for later use, that the same argument applied to the integral representation (5) for v_q shows that, for fixed z, $v_q(z)$ is also smoothly extended as a function of q in a neighborhood of Δ_n.

We continue to denote the extended function by ψ_n. Let us call $\partial\Delta_n$ the boundary, in $\mathcal{M}_n$, of Δ_n. This boundary consists of polynomials in Δ_n having at least one root of modulus 1. Such a polynomial can be factored for some positive k as $q = q_1 q_2$ where $q_2 \in \Delta_{n-k}$ and q_1, whose degree is k, has only roots of modulus 1. In this case, since $\tilde{q}_1 = \pm q_1$, it follows from Proposition 1 that

$$L_{q_1 q_2} = q_1 L_{q_2} \qquad \text{so that} \quad \psi_n(q_1 q_2) = \psi_{n-k}(q_2). \tag{8}$$

It is shown in [6] that Δ_n is topologically an n-ball embedded in $\mathscr{M}_n$, so that $\partial\Delta_n$ is actually an $(n-1)$-sphere which is, however, not smooth. It may of course happen now that ψ_n has critical points on $\partial\Delta_n$ and those are automatically reducible by (8). To treat such critical points, we need the following result (see the Corollary to Lemma 3 of [6]):

Proposition 4. *Let $q \in \Delta_n$. Then*

$$\frac{\partial}{\partial a}\psi_{n+1}((z-a)q)_{|a=1} = 2v_q^2(1),$$

$$\frac{\partial}{\partial a}\psi_{n+1}((z+a)q)_{|a=1} = 2v_q^2(-1),$$

and if $z^2 - 2\alpha z + \beta = (z-\xi)(z-\bar{\xi})$, where ξ_1 is any member of T,

$$\frac{\partial}{\partial \beta}\psi_{n+2}((z^2 - 2\alpha z + \beta)q)_{|\xi=\xi_1} = 2|v_q(\xi_1)|^2.$$

To conclude this section we state our global result concerning critical points.

Theorem 2 (The Index Theorem). *Assume f is such that ψ_n has only nondegenerate critical points, none of which lies on the boundary $\partial\Delta_n$. Let $\mathscr{C}_f$ be the set of critical points in Δ_n. Then $\mathscr{C}_f$ is finite, and, if we denote by $M(x)$ the Morse index of $x \in \mathscr{C}_f$, we have*

$$\sum_{x \in \mathscr{C}_f} (-1)^{M(x)} = 1. \tag{9}$$

We refer the reader to [6] for a proof. It can be shown that critical points on $\partial\Delta_n$ are automatically degenerate, so that the hypotheses in the above theorem are somewhat redundant. It is proved in [4] that they are generically satisfied.

4. H^2-Rational Approximation of Stieltjes Integrals

In the following we always assume that the function f that we approximate is a Stieltjes function of type (1). *The symbol λ will always denote a positive real number less than* 1, *such that the support of μ is included in $[-\lambda, \lambda]$.*

We first need to compute the function v_q, for $q \in \mathscr{M}_n^1$, in the Stieltjes case. In fact, it is easier to derive the value of v_q at $\bar{z}$, with $z \in U$: substituting $f^\sigma\tilde{q}$ for g in

(5) yields

$$
\begin{aligned}
v_q(\bar{z}) &= \overline{v_q(z)} \\
&= -\overline{\frac{1}{2i\pi} \int_T \frac{f^\sigma \tilde{q}}{q}(\xi) \frac{d\xi}{\xi - z}} \\
&= \frac{1}{2i\pi} \int_T \frac{fq}{\tilde{q}}(\xi) \frac{d\xi}{1 - \xi\bar{z}} \\
&= \frac{1}{2i\pi} \int_T \int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}}(\xi) \frac{d\mu(t)}{\xi - t} \frac{d\xi}{1 - \xi\bar{z}},
\end{aligned}
$$

the integration with respect to ξ being taken over the unit circle T as f^σ is analytic in some open set containing $\bar{U}$. Using successively Fubini and the residue formula,

$$
\begin{aligned}
v_q(\bar{z}) &= \frac{1}{2i\pi} \int_{-\lambda}^{\lambda} d\mu(t) \int_T \frac{q}{\tilde{q}}(\xi) \frac{d\xi}{(\xi - t)(1 - \xi\bar{z})} \\
&= \int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}}(t) \frac{d\mu(t)}{1 - t\bar{z}}.
\end{aligned}
$$

Thus

$$
v_q(z) = \int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}}(t) \frac{d\mu(t)}{1 - tz}. \tag{10}
$$

So far, this expression of v_q has been established only when $q \in \mathscr{M}_n^1$, but it follows from the continuity of $v_q(z)$ with respect to q pointed out after Proposition 3 that (10) is still valid when $q \in \Delta_n$. A consequence of this formula is the following lemma.

Lemma 1. *Assume $f \notin \mathscr{R}_n^0(V)$, and let $q \in \Delta_n$. Then v_q has at most n roots in $\bar{U}$, counting multiplicities.*

Proof. Let $\eta_1, \ldots, \eta_k$ be distincts roots of v_q, where η_i has multiplicity m_i. Let $\sum m_i = m$, and define $d \in \mathbf{P}_m[z]$ by

$$
d(z) = \prod_{i=1}^{k} (z - \eta_i)^{m_i}.
$$

For each index i,

$$
v_q(\eta_i) = v_q'(\eta_i) = \cdots = v_q^{(m_i - 1)}(\eta_i) = 0. \tag{11}
$$

Differentiating under the integral yields

$$
v_q^{(l)}(z) = \int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}}(t) \frac{l!\, t^l}{(1 - tz)^{l+1}} d\mu(t).
$$

Since the polynomials $t^l(1 - t\eta_i)^{m_i - 1 - l}$ are linearly independent over $\mathbf{C}$ for $0 \le l \le m_i - 1$, they form a basis of $\mathbf{P}_{c,m_i - 1}[t]$, the set of all complex polynomials of degree at most $m_i - 1$. Therefore, combining equations (11) together, we get

$$\int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}}(t) \frac{P_{m_i-1}(t)}{(1-t\eta_i)^{m_i}}\, d\mu(t) = 0, \qquad \forall i \in \{1, \ldots, k\}, \tag{12}$$

where P_{m_i-1} is any polynomial of $\mathbf{P}_{c,m_i-1}[t]$. Next, observe that the family of polynomials

$$\sum_i P_{m_i-1}(t) \frac{\tilde{d}(t)}{(1-t\eta_i)^{m_i}},$$

where P_{m_i-1} ranges over $\mathbf{P}_{c,m_i-1}[t]$ coincides with $\mathbf{P}_{c,m-1}[t]$, since the polynomials $\tilde{d}(t)/(1-t\eta_i)^{m_i}$ are coprime. Combining equations (12) together, this implies

$$\int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}\tilde{d}}(t) P_{m-1}(t)\, d\mu(t) = 0, \qquad \forall P_{m-1} \in \mathbf{P}_{c,m-1}[t]. \tag{13}$$

If $m > n$, then $q(t) \in \mathbf{P}_{m-1}[t]$ and we can substitute q for P_{m-1} in (13). Since $\tilde{q}$ and $\tilde{d}$ are of constant positive sign on $(-1, 1)$, the integrand is nonnegative, and therefore should vanish a.e. with respect to μ. However, the integrand can vanish in at most n points of $[-\lambda, \lambda]$ whereas the support of μ contains more than n points since $f \notin \mathscr{R}_n^0(V)$. ■

Now, we deduce from the above some information about the roots of critical points. It is shown in [17], using methods given in [9], [20], and [22], that a best approximation to a Stieltjes function is itself Stieltjes with poles in the convex hull of the support of μ. More generally, this turns out be true of L_q/q at any critical point, but establishing this entails an assertion on L_q that we do not need nor prove and we content ourselves with the following result:

Proposition 5. *Let q in Δ_n and f be a Stieltjes function with $f \notin \mathscr{R}_n^0(V)$. Assume that q is a critical point of ψ_n. Then q is in $\mathscr{M}_n^1$ and*

(i) *the critical point q is irreducible,*
(ii) *the roots of q are real,*
(iii) *the roots of q are distinct, and*
(iv) *the roots of q lie in the minimal segment containing the support of the measure $d\mu$.*

Proof. First suppose that $q \in \mathscr{M}_n^1$ and that $q = dp$ with $d = \text{g.c.d.}(q, L_q)$. Then p is critical and irreducible by Proposition 2. Following Theorem 1, p divides v_p. Then v_p/p has no other zero in $\bar{U}$ by Lemma 1. Using Proposition 2 again, we get $d = 1$ which means that q is irreducible. Assume now that $q = q_1 q_2 \in \Delta_n$ is any critical point where $q_1 \in \Delta_k$ has only roots of modulus 1, while q_2 has none, i.e., $q_2 \in \mathscr{M}_{n-k}^1$. Then q_2 is critical, because we see from (8) that the derivative of ψ_{n-k} at q_2 factors through the derivative of ψ_n at $q_1 q_2$, which is zero since q is critical. From the first part of the proof, we know that q_2 is irreducible, hence divides v_{q_2} by Theorem 1, so that v_{q_2} has no other zero in $\bar{U}$ by Lemma 1. In particular, v_{q_2}

has no zero on T. Now, let d be some irreducible factor over $\mathbf{R}$ of q_1, so that we have either $d = (z \pm 1)$ or $d = (z - \xi_1)(z - \bar{\xi}_1)$ with $\xi_1 \in T$. Replacing q_2 by dq_2, we get from (8) by the same reasoning as before that dq_2 is critical. However, then, Proposition 4 implies that v_{q_2} has a root on T, a contradiction showing that q cannot lie on $\partial\Delta_n$ and hence belongs to $\mathcal{M}_n^1$. Now, we know that q divides v_q. Suppose that

$$q(z) = \prod_{i=1}^{k} (z - \xi_i)^{n_i},$$

each root ξ_i being distinct. From the previous lemma, we see that (13) holds with m replaced by n and d replaced by q:

$$\int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}^2}(t) P_{n-1}(t)\, d\mu(t) = 0, \qquad \forall P_{n-1} \in \mathbf{P}_{\mathrm{c},n-1}[t]. \tag{14}$$

If the roots of q were not all real, we would write $q(t)$ as $|t - \xi|^2\, d(t)$, ξ a complex root of q, and take d as value of P_{n-1} in (14). As in the lemma, this would contradict the fact that μ is positive with a support comprising more than n points. This establishes (ii). If the roots of q were not all distinct, we would write $q(t)$ as $(t - \xi)^2$ $d(t)$, ξ now designating a multiple real root of q, and use the same positivity argument as before. This proves (iii). Finally, if they were not lying in the minimal segment containing the support of the measure $d\mu$, there would be a nonconstant factor of q of constant sign on the integration path in (14), again yielding a contradiction. This shows that (iv) holds, and the proof is complete. ■

We now focus on the quadratic form given by the Hessian matrix describing the second-order variation of the criterion in a neighborhood of an irreducible critical point $q \in \mathcal{M}_n^1$ following a direction given by the real vector $(\lambda_1, \ldots, \lambda_n)$. Induced by the index theorem, the condition we want to ensure is that this quadratic form is positive at any irreducible critical point. Indeed, each critical point would then have Morse index zero, which is even. In the course of the computation, we make linear changes of variables in the quadratic form. It is understood that this cannot affect the signature of the form.

Before starting the computation of the Hessian matrix, we claim that for each irreducible critical point $q \in \mathcal{M}_n^1$, $\mathbf{P}[z]$ denoting the space of all real polynomials, we have

$$\left\langle f - \frac{L_q}{q}, \frac{P}{q^2} \right\rangle = 0, \qquad \forall P \in \mathbf{P}[z]. \tag{15}$$

This is proved in Section 3.2 of [6], when $P \in \mathbf{P}_{2n-1}[z]$. However, if $\deg(P) \geq 2n$, we may use Euclidean division to write

$$\frac{P}{q^2} = P_1 + \frac{P_2}{q^2},$$

where $P_1 \in \mathbf{P}[z]$ and $P_2 \in \mathbf{P}_{2n-1}[z]$. Since $f - L_q/q \in H_{2,\mathbf{R}}^0(V)$ is orthogonal to $\mathbf{P}[z] \subset H_{2,\mathbf{R}}(U)$, we get (15). Also, for *any* $q \in \mathcal{M}_n^1$, L_q/q is the orthogonal projection of f onto the linear space V_q:

$$\left\langle f - \frac{L_q}{q}, \frac{z^k}{q} \right\rangle = 0, \qquad \forall k \in \{0, \ldots, n-1\}. \tag{16}$$

Taking partial derivatives in (16) with respect to the coefficients a_i of q leads us to

$$-\left\langle \frac{\partial}{\partial a_i}\left(\frac{L_q}{q}\right), \frac{z^k}{q} \right\rangle - \left\langle f - \frac{L_q}{q}, \frac{z^{k+i}}{q^2} \right\rangle = 0, \qquad \forall i \in \{0, \ldots, n-1\}.$$

If, in addition, q is critical, (15) shows that the second term on the left-hand side is zero and we get the orthogonality relations:

$$\left\langle \frac{\partial}{\partial a_i}\left(\frac{L_q}{q}\right), \frac{z^k}{q} \right\rangle = 0, \qquad \forall k \in \{0, \ldots, n-1\}. \tag{17}$$

Now, we start computing. The entries of the Hessian matrix at a critical point $q \in \mathcal{M}_n^1$ are by definition:

$$\begin{aligned}
&\frac{\partial^2}{\partial a_i \, \partial a_j}\left\langle f - \frac{L_q}{q}, f - \frac{L_q}{q} \right\rangle \\
&= -2\frac{\partial}{\partial a_i}\left\langle f - \frac{L_q}{q}, \frac{\partial}{\partial a_j}\frac{L_q}{q} \right\rangle \\
&= -2\left[-\left\langle \frac{\partial}{\partial a_i}\left(\frac{L_q}{q}\right), \frac{\partial}{\partial a_j}\left(\frac{L_q}{q}\right) \right\rangle + \left\langle f - \frac{L_q}{q}, \frac{\partial}{\partial a_i}\frac{\partial}{\partial a_j}\left(\frac{L_q}{q}\right) \right\rangle \right] \\
&= 2\left[\left\langle \frac{\partial}{\partial a_i}\left(\frac{L_q}{q}\right), \frac{\partial}{\partial a_j}\left(\frac{L_q}{q}\right) \right\rangle - \left\langle f - \frac{L_q}{q}, \frac{\partial}{\partial a_i}\left(\frac{\partial L_q/\partial a_j}{q} - \frac{L_q z^j}{q^2}\right) \right\rangle \right] \\
&= 2\left[\left\langle \frac{\partial}{\partial a_i}\left(\frac{L_q}{q}\right), \frac{\partial}{\partial a_j}\left(\frac{L_q}{q}\right) \right\rangle - \left\langle f - \frac{L_q}{q}, 2z^{j+i}\frac{L_q}{q^3} \right\rangle \right],
\end{aligned} \tag{18}$$

where the last equality uses (15).

Putting everything over the common denominator q^2, the functions $(\partial/\partial a_i)(L_q/q)$ can be written as R_i/q^2, with $R_i \in \mathbf{P}_{2n-1}[z]$. Since they are orthogonal to V_q by (17), Corollary 1 implies that $z^n R_i(z)/q^2\tilde{q}$ belongs to $H^0_{2,\mathbf{R}}(V)$. Since the roots of $\tilde{q}$ belong to V, it follows that $\tilde{q}(z)$ divides $z^n R_i(z)$, but $\tilde{q}(z)$ is prime to z, hence divides $R_i(z)$. So, there exist polynomials v_i of $\mathbf{P}_{n-1}[z]$ such that

$$\frac{\partial}{\partial a_i}\left(\frac{L_q}{q}\right) = \frac{\tilde{q}}{q^2} v_i, \qquad \forall i \in \{0, \ldots, n-1\}, \tag{19}$$

or

$$q\frac{\partial L_q}{\partial a_i} - z^i L_q = \tilde{q} v_i, \qquad \forall i \in \{0, \ldots, n-1\}. \tag{20}$$

Let $L_q^* \in \mathbf{P}[z]$ be any polynomial interpolating $1/L_q$ at the roots of q. Such a polynomial exists because q and L_q are coprime. Since

$$L_q L_q^* \equiv 1 \pmod{q},$$

where $\equiv$ denotes congruence, it follows from (20) that

$$z^{j+i} L_q \equiv \tilde{q}^2 L_q^* v_i v_j \pmod{q}. \tag{21}$$

Consider now the bracketed difference in (18). Using (19) and the fact that $\tilde{q}/q$ is a Blaschke product, the first term of the difference is simply

$$\left\langle \frac{v_i}{q}, \frac{v_j}{q} \right\rangle,$$

and, using (21) and (15), the second term is

$$\left\langle f - \frac{L_q}{q}, 2\frac{\tilde{q}^2 L_q^*}{q^3} v_i v_j \right\rangle.$$

The second-order variation δ_q of the criterion is thus given by

$$(\tfrac{1}{2})\delta_q(\lambda_1, \ldots, \lambda_n) = \left\| \frac{\sum \lambda_i v_i}{q} \right\|^2 - \left\langle f - \frac{L_q}{q}, 2\frac{\tilde{q}^2 L_q^*}{q^3} (\textstyle\sum \lambda_i v_i)^2 \right\rangle. \tag{22}$$

Let $v = \sum \lambda_i v_i$ and

$$q(z) = \prod_{i=1}^{n} (z - \xi_i),$$

where the ξ_i's are distinct real numbers by Proposition 5. We then have

$$\frac{v}{q}(z) = \sum_i \frac{v(\xi_i)}{q'(\xi_i)} \frac{1}{z - \xi_i}$$

and the squared norm in (22) becomes, thanks to the residue formula,

$$\left\| \frac{v}{q} \right\|^2 = \left\langle \frac{v}{q}, \frac{v}{q} \right\rangle = \sum_{i,j} \frac{v(\xi_i)}{q'(\xi_i)} \frac{v(\xi_j)}{q'(\xi_j)} \frac{1}{1 - \xi_i \xi_j}.$$

Also, the second term on the right-hand side of (22) can be computed as

$$-\left\langle f - \frac{L_q}{q}, 2\frac{\tilde{q}^2 L_q^*}{q^3} v^2 \right\rangle = -\left\langle v_q^\sigma \frac{\tilde{q}}{q}, 2\frac{\tilde{q}^2 L_q^*}{q^3} v^2 \right\rangle = -\left\langle 2\frac{\tilde{q} L_q^*}{q^2} v^2, v_q^\sigma \right\rangle.$$

As the polynomial q divides the function v_q, we denote by w the function v_q/q. Using (3) and the residue formula again, the previous expression becomes

$$-2\left(\frac{1}{2i\pi}\right) \int_T zw \frac{\tilde{q}}{q}(z) L_q^*(z) v^2(z) \frac{dz}{z} = -2 \sum_i \frac{w(\xi_i)\tilde{q}(\xi_i) L_q^*(\xi_i) v^2(\xi_i)}{q'(\xi_i)}$$

$$= -2 \sum_i \frac{w(\xi_i)\tilde{q}(\xi_i) v^2(\xi_i)}{q'(\xi_i) L_q(\xi_i)}.$$

Finally, the variation δ_q of the criterion can be expressed as

$$(\tfrac{1}{2})\delta_q(\lambda_1, \ldots, \lambda_n) = \sum_{i,j} \frac{v(\xi_i)}{q'(\xi_i)} \frac{v(\xi_j)}{q'(\xi_j)} \frac{1}{1 - \xi_i\xi_j} - 2 \sum_i \frac{w(\xi_i)\tilde{q}(\xi_i)v^2(\xi_i)}{q'(\xi_i)L_q(\xi_i)}.$$

Because of (20) where q and $L(q)$ are coprime, it is permissible to define new variables

$$X_i = \frac{v(\xi_i)}{q'(\xi_i)}$$

since the values $v(\xi_i)$ can be assigned arbitrarily over $\mathbf{R}$ by a suitable choice of the λ_i's. Hence we get

Proposition 6. *Let q be an irreducible critical point of the criterion, then the second-order variation in the neighborhood of q is given by*

$$\tfrac{1}{2}\delta_q = \sum_{i,j} \frac{X_iX_j}{1 - \xi_i\xi_j} - 2 \sum_i \frac{w(\xi_i)\tilde{q}(\xi_i)q'(\xi_i)}{L_q(\xi_i)} X_i^2.$$

Just as we gave an integral expression for v, it is also possible to express the quotient $w = v_q/q$ by

$$w(z) = \int_{-\lambda}^{\lambda} t^n \frac{q(t)}{\tilde{q}(t)^2} \frac{d\mu}{1 - tz}, \tag{23}$$

and the omitted proof follows the same lines. Since the values of $w(\xi_i)$ appear in δ_q, we need a new expression for them. Assume first that $\xi_i \neq 0$ and perform the Euclidean division of t^n by $1 - t\xi_i$:

$$t^n = Q_{n-1}(t)(1 - t\xi_i) + \left(\frac{1}{\xi_i}\right)^n, \qquad Q_{n-1} \in \mathbf{P}_{n-1}[t].$$

Then, from (23),

$$\begin{aligned} w(\xi_i) &= \int_{-\lambda}^{\lambda} Q_{n-1}(t) \frac{q(t)}{\tilde{q}^2(t)}\, d\mu + \left(\frac{1}{\xi_i}\right)^n \int_{-\lambda}^{\lambda} \frac{q(t)}{\tilde{q}^2(t)} \frac{d\mu}{1 - t\xi_i} \\ &= \left(\frac{1}{\xi_i}\right)^n \int_{-\lambda}^{\lambda} \frac{q(t)}{\tilde{q}^2(t)} \frac{d\mu}{1 - t\xi_i} \end{aligned}$$

using relation (14). Now dividing q by $(1 - t\xi_i)$:

$$q(t) = R_{n-1}(t)(1 - t\xi_i) + q\left(\frac{1}{\xi_i}\right), \qquad R_{n-1} \in \mathbf{P}_{n-1}[t],$$

we get again from (14)

$$\int_{-\lambda}^{\lambda} \frac{q^2(t)}{\tilde{q}^2(t)} \frac{d\mu}{1 - t\xi_i} = q\left(\frac{1}{\xi_i}\right) \int_{-\lambda}^{\lambda} \frac{q(t)}{\tilde{q}^2(t)} \frac{d\mu}{1 - t\xi_i}.$$

Therefore,

$$w(\xi_i) = \frac{1}{\tilde{q}(\xi_i)} \int_{-\lambda}^{\lambda} \frac{q^2(t)}{\tilde{q}^2(t)} \frac{d\mu}{1 - t\xi_i}. \tag{24}$$

If $\xi_i = 0$, (24) is still valid since it is merely the result of the substitution $z = 0$ in (23), taking into account (14) and the fact that $\tilde{q}(0) = 1$. In the same way, we get an expression for $(L_q/\tilde{q})(\xi_i)$. In the complement of the closed unit disk $\bar{U}$,

$$fq = v_q^\sigma \tilde{q} + L_q$$

and, except at points $1/\xi_i$, we have

$$\frac{L_q}{\tilde{q}} = f\frac{q}{\tilde{q}} - v_q^\sigma$$

or

$$\frac{L_q}{\tilde{q}}(z) = \int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}}(z)\frac{d\mu}{z-t} - \int_{-\lambda}^{\lambda} \frac{q}{\tilde{q}}(t)\frac{d\mu}{z-t} = \int_{-\lambda}^{\lambda}\left(\frac{q}{\tilde{q}}(z) - \frac{q}{\tilde{q}}(t)\right)\frac{d\mu}{z-t}.$$

Except at points $\{1/\xi_i\}$, this integral is well defined over the entire complex plane because $(z - t)$ is a factor of $(q/\tilde{q})(z) - (q/\tilde{q})(t)$ which has no poles in $[-\lambda, \lambda]$ as a function of t. We deduce that

$$\frac{L_q}{\tilde{q}}(\xi_i) = \int_{-\lambda}^{\lambda} \frac{q_i(t)}{\tilde{q}(t)}\, d\mu,$$

where, from now on, q_i stands for the polynomial

$$q_i(t) := \frac{q(t)}{t - \xi_i}, \qquad i \in \{1, \ldots, n\}.$$

However,

$$\tilde{q}(t) = S_{n-1}(t)(t - \xi_i) + \tilde{q}(\xi_i), \qquad S_{n-1} \in \mathbf{P}_{n-1}[t],$$

which yields by (14)

$$\frac{L_q}{\tilde{q}}(\xi_i) = \tilde{q}(\xi_i) \int_{-\lambda}^{\lambda} \frac{q_i(t)}{\tilde{q}^2(t)}\, d\mu.$$

Moreover,

$$q_i(t) = T_{n-2}(t)(t - \xi_i) + q'(\xi_i), \qquad T_{n-2} \in \mathbf{P}_{n-2}[t]$$

and thus

$$\frac{L_q}{\tilde{q}}(\xi_i) = \frac{\tilde{q}(\xi_i)}{q'(\xi_i)} \int_{-\lambda}^{\lambda} \frac{q_i^2(t)}{\tilde{q}^2(t)}\, d\mu. \tag{25}$$

Putting

$$Q_i := \frac{\int_{-\lambda}^{\lambda}(q^2(t)/\tilde{q}^2(t))(d\mu/(1 - t\xi_i))}{\int_{-\lambda}^{\lambda}(q_i^2(t)/\tilde{q}^2(t))\, d\mu}, \tag{26}$$

we get from (25), (24), and Proposition 6

Proposition 7. *The second-order variation $\frac{1}{2}\delta_q$ is given in the variables X_i's by*

$$\sum_{i,j} \frac{X_i X_j}{1 - \xi_i \xi_j} - 2 \sum_i \left[\frac{q'(\xi_i)}{\tilde{q}(\xi_i)}\right]^2 Q_i X_i^2. \tag{27}$$

Next, we provide a positivity criterion for this quadratic form.

Proposition 8. *A sufficient condition for the quadratic form* (27) *to be positive definite is that*

$$\sum_i \frac{2Q_i}{1 - \xi_i^2} < 1.$$

Proof. The hypothesis implies that there exist positive real numbers $(\gamma_i)_{1 \le i \le n}$ such that $\sum_i \gamma_i = 1$ and

$$\frac{2Q_i}{1 - \xi_i^2} < \gamma_i. \tag{28}$$

Recall that $q_i \in \mathbf{P}_{n-1}[z]$ is given by

$$q_i(z) = \prod_{j \neq i} (z - \xi_j),$$

and let

$$h_i(z) = \left(\frac{2Q_i}{\gamma_i(1 - \xi_i^2)}\right)^{1/2} \frac{q_i(z)}{\tilde{q}_i(z)}.$$

The h_i's are real analytic functions in U and from (28), since $q_i/\tilde{q}_i$ is a Blaschke product,

$$\sup_{z \in U} |h_i(z)| = \left(\frac{2Q_i}{\gamma_i(1 - \xi_i^2)}\right)^{1/2} < 1.$$

It now follows from Pick's theorem (e.g., Theorem 2.2 of [18]) that the real quadratic form

$$P_i(X_1, \ldots, X_n) = \sum_{k,l=1}^{n} \frac{1 - h_i(\xi_k) h_i(\xi_l)}{1 - \xi_k \xi_l} X_k X_l$$

is positive definite. Taking into account the fact that $q_i(\xi_i) = q'(\xi_i)$ and $q_i(\xi_j) = 0$ if $i \neq j$, it is but a simple computation to check that the quadratic form (27) can be written as

$$\sum_{i=1}^{n} \gamma_i P_i(X_1, \ldots, X_n),$$

hence is positive definite. ∎

We now proceed with the final portion of the proof, showing how a condition on the support of the measure μ can ensure that the positivity criterion of Proposition 8 is satisfied. Let us consider the real Hilbert space $L^2(d\mu)$ with scalar product

$$\langle f, g\rangle_{L^2} := \int_{-\lambda}^{\lambda} f(t)g(t)\, d\mu.$$

The functions $(q_i/\tilde{q})$ form an orthogonal family in $L^2(d\mu)$ since, for each pair (i, j) of distinct integers in $\{1, \dots, n\}$, we have by relation (14)

$$\int_{-\lambda}^{\lambda} \frac{q_i}{\tilde{q}}(t) \frac{q_j}{\tilde{q}}(t)\, d\mu = 0.$$

Let

$$e_i := \frac{q_i/\tilde{q}}{\|q_i/\tilde{q}\|_{L^2}}.$$

Then (e_i) is an orthonormal family in $L^2(d\mu)$. Define further

$$T_i := \frac{\int_{-\lambda}^{\lambda}(q^2(t)/\tilde{q}^2(t))\, d\mu}{\int_{-\lambda}^{\lambda}(q_i^2(t)/\tilde{q}^2(t))\, d\mu}. \tag{29}$$

From the definition of Q_i, we get

$$Q_i \le \left(\frac{1}{1 - \lambda|\xi_i|}\right) T_i. \tag{30}$$

On the other hand, using relation (14),

$$T_i = \frac{\int_{-\lambda}^{\lambda}(q_i/\tilde{q})(t)(tq(t)/\tilde{q}(t))\, d\mu}{\int_{-\lambda}^{\lambda}(q_i^2/\tilde{q}^2)(t)\, d\mu}$$

which can be rewritten as

$$T_i = \frac{\langle tq/\tilde{q}, e_i\rangle_{L^2}}{\|q_i/\tilde{q}\|_{L^2}}. \tag{31}$$

The squared $L^2(d\mu)$-norm of the function $tq/\tilde{q}$ is not less than the squared norm of its projection on the space spanned by the e_i's:

$$\int_{-\lambda}^{\lambda} t^2 \frac{q^2}{\tilde{q}^2}(t)\, d\mu \ge \sum_i \left|\left\langle t\frac{q}{\tilde{q}}, e_i\right\rangle_{L^2}\right|^2. \tag{32}$$

Assume now that $f \notin \mathscr{R}_n^0(V)$. Then strict inequality holds in (32) because equality would mean that $tq(t)$ is equal μ-almost everywhere to a linear combination with real coefficients of the $q_i(t)$, and checking the degree shows that this is impossible when μ has at least $(n + 1)$ points in its support. Comparing (32) and (31),

$$\int_{-\lambda}^{\lambda} t^2 \frac{q^2}{\tilde{q}^2}(t)\, d\mu > \sum_i T_i^2 \left\|\frac{q_i}{\tilde{q}}\right\|_{L^2}^2$$

and thus

$$\lambda^2 \int_{-\lambda}^{\lambda} \frac{q^2}{\tilde{q}^2}(t)\,d\mu > \sum_i T_i^2 \int_{-\lambda}^{\lambda} \frac{q_i^2}{\tilde{q}^2}(t)\,d\mu. \tag{33}$$

Inequality (33) shows that

$$\lambda^2 > \sum_i T_i,$$

implying by (30) that

$$\sum_i \frac{2Q_i}{1-\xi_i^2} < 2\frac{\lambda^2}{(1-\lambda^2)^2}. \tag{34}$$

Note that if the measure $d\mu$ was one-sided, that is if the defining integral for f was to be taken between 0 and λ or between $-\lambda$ and 0, then applying (32) to the function $(t-\lambda/2)q/\tilde{q}$ or $(t+\lambda/2)q/\tilde{q}$ rather than to $tq/\tilde{q}$ would improve the term λ^2 in (33) to $\lambda^2/4$. Accordingly, the numerator on the right-hand side of (34) would change from λ^2 to $\lambda^2/4$. Finally, we get our result as follows:

Theorem 3. *Let λ_0 be the unique real positive root smaller than* 1 *of*

$$1 = \frac{2\lambda^2}{(1-\lambda^2)^2}$$

($\lambda_0 \simeq 0.5176$). *For each Stieltjes function of type* (1) *which is not in $\mathscr{R}_n^0(V)$ and whose support lies in $[-\lambda_0, \lambda_0]$, there is a unique critical point (hence local minimum) for the $H^0_{2,\mathbf{R}}(V)$ rational approximation problem in degree n. If the measure $d\mu$ is one-sided, the same result holds true when λ_0 is replaced by the unique real positive root smaller than* 1 *of*

$$1 = \frac{\lambda^2}{2(1-\lambda^2)^2}$$

($\lambda_0 = 1/\sqrt{2} \simeq 0.7071$).

Proof. As $f \notin \mathscr{R}_n^0(V)$, Proposition 5 implies that each critical point in degree n is irreducible and at such a point, under the foregoing assumptions on the support of μ, it follows from (34) that the positivity criterion for the Hessian matrix given in Proposition 8 is met. In particular, each critical point is nondegenerate of Morse index 0, and the index theorem applies to show uniqueness. The case where μ is one-sided is treated similarly using the remark after (34). ∎

This theorem leaves out the case where $f \in \mathscr{R}_n^0(V)$. It would remain true, however, provided the degree of f is n. When the degree of f is less than n, the problem degenerates because any multiple of the denominator of f is a critical point, but it is easy to show that there are no others, so that f is the unique local best approximation to itself in $\mathscr{R}_n^0(V)$. We leave it to the reader to check this from what precedes. In fact, it is possible to prove that the above two assertions concerning the case where $f \in \mathcal{R}_n^0(V)$ hold true even if f is not Stieltjes. The first one is established in [4].

5. Related Work and Open Questions

There is a strong connection between the works [11]–[13] and the problem under study here. Consider the defining formula of a monospline with simple knots

$$M_{a,x}(t) = \int K(x, t)\, d\mu(t) - \sum_{k=0}^{n} a_k K(x_k, t), \tag{35}$$

where μ is a positive measure on some open real interval I and K is some totally positive kernel, while $a = (a_1, \ldots, a_n)$ and $x = (x_1, \ldots, x_n)$ belong to $\mathbf{R}^n$ and the x_k's lie in the support of μ. Let $\|\cdot\|_p$ denote the norm in $L^p(I)$. In [13] it is shown that when μ is the Lebesgue measure, $K(x, t)$ is of the form $g(t - x)$ and $1 \leq p < \infty$, there is a unique pair (a^*, x^*) such that all derivatives of $\|M_{a^*,x^*}\|_p$ with respect to the a_k's and x_k's do vanish, in other words (a^*, x^*) is the unique critical point of the L^p norm. This implies that M_{a^*,x^*} is the unique monospline in the family $M_{a,x}$ which is of minimum norm, both locally and globally with respect to the parameters (a, x). These authors also indicate that the result could be carried over to any μ such that $d\mu$ is Log-concave, and that the L^p norm could also involve a Log-concave weight. It is to be noted that the condition on the support of μ that we require in the present work and the Log-concave condition given in [13] both assert, in different ways, that μ should not concentrate at the endpoints of the interval on which it is defined.

If we let $K(x, t) = 1/(1 - xt)$, and if we allow, in formula (35), for x to be complex, keeping $|t| < 1$ and $|x| < 1$, we get the difference between a function g of type (2) and a real rational function of degree n with real simple poles. As we have seen, the problem of finding a best rational approximation to g in $H_{2,\mathbf{R}}(U)$ amounts to minimizing the above difference, but the minimization, in our case, is taken with respect to $L^2(T)$ and not $L^2(I)$.

Another formal complexification of the monospline problem, which is perhaps closer to general rational approximation in Hardy spaces, is to allow for x and t to belong to the closed unit disk $\bar{U}$ in $\mathbf{C}$, and to let $d\mu$ be some complex function in $L^p(T)$. Due to the Cauchy formula, if we set $K(x, t) = 1/(1 - x\bar{t})$ (note that the latter is positive in the usual sense of complex analysis), (35) becomes $f - p/q$ where f could be any function in the Hardy space $H_p(U)$ and p/q any rational function of degree n in this Hardy space. Therefore, minimizing the norm of this expression in $L^p(T)$ really means performing rational approximation of degree n in Hardy spaces, which is our concern here in the case $p = 2$, and with the additional requirement that both f and p/q should be real.

In view of the above, it is natural to ask whether the uniqueness property in Theorem 3 holds without restriction on the support of μ when $d\mu$ is Log-concave. We do not know the answer to this. Many other types of integral representations for f lead to similar computations, but the problem is always to find an analog to the positivity argument that allows us, in the Stieltjes case, to estimate the roots of q. We hope that the technique will find such applications in the future.

The index theorem here plays a role which is similar to the invariance property of the Brouwer degree of the map Φ used in [13]. This theorem can be viewed as a consequence of the Morse relations that drop out from the fact that the error function is naturally defined over certain manifolds [8].

Acknowledgments. The authors were introduced to Stieltjes functions and their special properties in rational approximation by Professor E. B. Saff, from whom they learned about the positivity argument used in Proposition 5. We are also indebted to Professor Saff who conjectured the result that we proved here. The paper also owes very much to the constructive comments of Professor D. Braess, who showed the authors how to improve their original bound.

References

1. G. A. Baker (1975): Essentials of Padé approximants. San Diego, CA: Academic Press.
2. L. Baratchart (1986): *Existence and generic properties of L^2 approximants for linear systems.* IMA J. Math. Control Inform. **3**:89–101.
3. L. Baratchart (1987): *Recent and new results in rational L^2 approximation.* In: Modelling, Robustness and Sensitivity Reduction in Control Systems (R. F. Curtain, ed.). NATO Adv. Study Inst. Ser., vol. 34. Berlin: Springer-Verlag.
4. L. Baratchart (1987): Sur l'approximation rationnelle l^2 pour les systèmes dynamiques linéaires. Thèse de doctorat d'état, Université de Nice.
5. L. Baratchart, M. Cardelli, M. Olivi (1991): *Identification and rational l^2-approximation: a gradient algorithm.* Automatica, **27**(2):413–418.
6. L. Baratchart, M. Olivi (1988): *Index of critical points in l^2-approximation.* Systems Control Lett., **10**:167–174.
7. L. Baratchart, M. Olivi, F. Wielonsky (1990): *Asymptotic properties in rational l^2-approximation.* Proceedings of the 9th International Conference on Analysis and Optimization of Systems, Antibes. Lecture Notes in Control and Information Sciences, vol. 144. Berlin: Springer-Verlag, pp. 477–486.
8. L. Baratchart, F. Wielonsky (in preparation): *Rational approximation in Hardy spaces and Morse theory.*
9. R. B. Barrar, H. L. Loeb, H. Werner (1974): *On the existence of optimal integration formulas for analytic functions.* Numer. Math., **23**:105–117.
10. B. D. Bojanov (1979): *Uniqueness of the monosplines of least deviation.* In: Numerische Integration (G. Hammerlin, ed.). International Series of Numerical Mathematics, vol. 45. Basel: Birkhauser, pp. 67–97.
11. D. Braess (1986): Nonlinear Approximation Theory. Springer Series in Computational Mathematics, vol. 7. Berlin: Springer-Verlag.
12. D. Braess, N. Dyn (1982): *On the uniqueness of monosplines and perfect splines of least L_1- and L_2-norm.* J. Analyse Math., **41**:217–233.
13. D. Braess, N. Dyn (1986): *On the uniqueness of generalized monosplines of least L_p-norm.* Constr. Approx., **2**:79–99.
14. J. Della Dora (1980): Contribution à l'approximation de fonctions de la variable complexe au sens de Hermite–Padé et de Hardy. Thèse d'état Université Scientifique et Médicale de Grenoble.
15. M. Duc-Jacquet (1973): Approximation des fonctionnelles linéaires sur les espaces Hilbertiens à noyaux reproduisants. Thèse d'état Université Scientifique et Médicale de Grenoble.
16. N. Dyn (1985): *Generalized monosplines and optimal approximation.* Constr. Approx., **1**:137–154.

17. M. FARÈS (1982): Sur l'approximation rationnelle de fonctions de la variable complexe au sens de la norme de Hardy. Thèse de 3^e cycle de l'Université Scientifique et Médicale de Grenoble.
18. J. B. GARNETT (1981): Bounded Analytic Functions. San Diego, CA: Academic Press.
19. J. ILLAN (1987): *On the rational approximation of H_p functions in the $L_p(\mu)$ metric.* Proceedings of the Approximation and Optimization, Havana. Lecture Notes in Mathematics, vol. 1354. Berlin: Springer-Verlag, pp. 155–163.
20. S. KARLIN (1972): *On a class of nonlinear approximation problems.* Bull. Amer. Math. Soc., **78**:43–49.
21. L. LJUNG (1987): System Identification. Theory for the User. Englewood Cliffs, NJ: Prentice-Hall.
22. N. RICHTER DYN (1971): *Properties of minimal integration rules.* SIAM J. Numer. Anal., **8**:379–388.
23. G. RUCKEBUSH (1978): Sur l'approximation rationnelle des filtres. Rapport No. 35, C.M.A. Ecole Polytechnique.
24. E. B. SAFF (1990): Personal communication.
25. J. L. WALSH (1969): Interpolation and Approximation by Rational Functions in the Complex Domain. American Mathematical Society Colloquium Publications, vol. XX. Providence, RI: American Mathematical Society.
26. J. C. WILLEMS (1976): *Realization of systems with internal passivity and symmetry constraints.* J. Franklin Inst., **301**:605–621.

L. Baratchart
Institut National de Recherche en Informatique et Automatique
Avenue E. Hughes
Sophia-Antipolis
06560 Valbonne
France

F. Wielonsky
Institut National de Recherche en Informatique et Automatique
Avenue E. Hughes
Sophia-Antipolis
06560 Valbonne
France