On the Convergence of Optimal Measures

T. Bloom
Department of Mathematics
University of Toronto
Toronto, Ontario
Canada M5S 2E4,

L. Bos
Department of Mathematics and Statistics
University of Calgary
Calgary, Alberta
Canada T2N 1N4,

N. Levenberg
Department of Mathematics
Indiana University
Bloomington, Indiana,
USA

and

S. Waldron
Department of Mathematics
University of Auckland
Auckland, New Zealand

June 18, 2009

Abstract

Using recent results of Berman and Boucksom [3] we show that for a non-pluripolar compact set $K \subset \mathbb{C}^d$ and an admissible weight function $w = e^{-\phi}$ any sequence of optimal measures converges weak-* to the equilibrium measure $\mu_{K,\phi}$ of (weighted) pluripotential theory for K, ϕ.

Mathematics Subject Classification: 32U20, 41A63

Keywords and phrases: weighted optimal measure, weighted transfinite diameter, weighted equilibrium measure

1 Introduction

In classical potential theory in the complex plane, given $K \subset \mathbb{C}$ compact, one minimizes the logarithmic energy

$$I(\mu) := \int_K \int_K \log \frac{1}{|z-\zeta|} d\mu(z) d\mu(\zeta)$$

over all probability measures μ supported in K. Provided K is non-polar, there exists a unique energy minimizing measure μ_K. More generally, given a nonnegative uppersemicontinuous (usc) weight function $w := e^{-\phi}$ on K with $\{z \in K : w(z) > 0\}$ non-polar (an *admissible* weight), one minimizes the weighted logarithmic energy

$$I^w(\mu) := \int_K \int_K \log \frac{1}{|z-\zeta| w(z) w(\zeta)} d\mu(z) d\mu(\zeta)$$

over all probability measures μ supported in K and one obtains a unique minimizer $\mu_{K,\phi}$. Finding $\mu_{K,\phi}$ explicitly is usually difficult; thus one looks for good approximations to $\mu_{K,\phi}$. One approach is simply discretizing the (weighted) energy; this leads to the notion of (weighted) Fekete points (cf., section 2.2 and the proof of Proposition 3.4). Another approach, which we take in this paper, is to utilize L^2-methods, leading to the notion of *optimal measures*. We show this approach is successful in higher dimensions as well.

Pluripotential theory in several complex variables ($\mathbb{C}^d$ for $d > 1$) is the study of plurisubharmonic functions. In this setting, we have analogues of equilibrium measures μ_K and $\mu_{K,\phi}$, but there are no related energy notions. We recall that a function $u : \mathbb{C}^d \to [-\infty, \infty)$ is said to be plurisubharmonic (psh) if it is usc and, when restricted to any complex line, is either subharmonic or identically $-\infty$. A set $E \subset \mathbb{C}^d$ is pluripolar if $E \subset \{z \in \mathbb{C}^d : u(z) = -\infty\}$ for some psh u (with $u \not\equiv -\infty$).

Suppose that $K \subset \mathbb{C}^d$ is compact and non-pluripolar. As in the univariate setting, we call a nonnegative usc weight function $w := e^{-\phi}$ on K with $\{z \in K : w(z) > 0\}$ non-pluripolar an *admissible* weight, and we proceed to describe a higher-dimensional generalization of $\mu_{K,\phi}$. First, the class of psh functions of at most logarithmic growth at infinity is denoted by

$$\mathcal{L} := \{u \,:\, u \text{ is psh and } u(z) \le \log^+ |z| + C\}.$$

We define

$$V_{K,\phi}(z) := \sup \{u(z) \,:\, u \in \mathcal{L},\ u \le \phi \text{ on } K\}. \tag{1}$$

The function $V^*_{K,\phi}(z)$ which is the usc regularization of $V_{K,\phi}$, will be called the weighted extremal function of K, ϕ. Associated to this extremal function is the *weighted equilibrium measure*,

$$\mu_{K,\phi} := \frac{1}{(2\pi)^d}(dd^c V^*_{K,\phi})^d.$$

Here $(dd^c v)^d$ is notation for the Monge-Ampere operator (applied to v). That $\mu_{K,\phi}$ exists and is a probability measure can be found in Appendix B of [17] (see also [15]). We simply write μ_K in the unweighted case, i.e., $w \equiv 1$ and $\phi \equiv 0$. We remark, that in one variable, for $K = [-1, 1] \subset \mathbb{C}$,

$$\mu_K = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}dx.$$

For each $n = 1, 2, ...$ we let $\mathcal{P}_n$ denote the holomorphic polynomials of degree at most n. Given μ a probability measure on K and an admissible weight w on K, for each $n = 1, 2, ...$ we form a *weighted inner product of degree* n by

$$\langle f, g\rangle_{\mu,w} := \int_K f(z)\overline{g(z)}w(z)^{2n}d\mu. \tag{2}$$

Provided $\langle p, p\rangle_{\mu,w} = 0$ for $p \in \mathcal{P}_n$ implies that $p = 0$, $\mathcal{P}_n$ equipped with the inner-product (2) is a finite dimensional Hilbert space of dimension

$$N = N(n) := \binom{d+n}{n}. \tag{3}$$

For a fixed basis $B_n = \{p_1, p_2, \cdots, p_N\}$ of $\mathcal{P}_n$ we form the Gram matrix

$$G_n^{\mu,w} = G_n^{\mu,w}(B_n) := [\langle p_i, p_j\rangle_{\mu,w}] \in \mathbb{C}^{N\times N}.$$

Definition 1.1 *Suppose that w is an admissible weight on K. If a probability measure μ has the property that*

$$(a)\ \det(G_n^{\mu',w}) \le \det(G_n^{\mu,w})$$

for all other probability measures μ' on K then μ is said to be an optimal measure of degree n for K and w.

Our main result is the following.

Main Theorem. *Suppose that $K \subset \mathbb{C}^d$ is compact and that w is an admissible weight function. Suppose further that μ_n is an optimal measure of degree n for K and w. Then*

$$\lim_{n\to\infty} \mu_n = \mu_{K,\phi}$$

where the limit is in the weak$-*$ sense.

In the next section, we provide background and motivation for the study of (unweighted) optimal measures from several perspectives. In section 3 we discuss weighted optimal measures and various properties. Then in section 4 we prove our main theorem which utilizes recent deep results of Berman and Boucksom.

2 Introduction to Optimal Measures

Here we give a motivational introduction to optimal measures in the unweighted case ($w \equiv 1$). Suppose that $K \subset \mathbb{C}^d$ is compact and non-pluripolar and that μ is a probability measure on K. We assume that μ is non-degenerate on $\mathcal{P}_n$. This means that with the associated inner-product

$$\langle f, g\rangle_\mu := \int_K f\bar{g}\,d\mu \tag{4}$$

and $L^2(\mu)$ norm, $\|f\|_{L^2(\mu)} = \sqrt{\langle f, f\rangle_\mu}$, we have $\|p\|_{L^2(\mu)} = 0$ for $p \in \mathcal{P}_n$ implies that $p = 0$. For the rest of the paper, we assume all of our measures are non-degenerate. It follows from the reasoning used in Proposition 3.5 of [7] that μ is non-degenerate on $\mathcal{P}_n$ if and only if $\mathrm{supp}(\mu)$ is not contained in an algebraic variety of degree n. Then $\mathcal{P}_n$ equipped with the inner-product (4) is a finite dimensional Hilbert space of dimension N (see (3)). We may also consider the uniform norm on K,

$$\|f\|_K := \max_{z\in K} |f(z)|$$

and it is natural to compare the two norms for $p \in \mathcal{P}_n$.

Since μ is a probability measure we always have

$$\|p\|_{L^2(\mu)} \le \|p\|_K.$$

Moreover since $\mathcal{P}_n$ is finite dimensional there is always a constant $C = C(n, \mu, K)$ such that the reverse inequality holds,

$$\|p\|_K \le C\|p\|_{L^2(\mu)}.$$

In fact, as is well known and easy to verify, the *best* constant C (sometimes called the Bernstein-Markov factor) is given by

$$C = \sup_{p\in\mathcal{P}_n,\, p\neq 0} \frac{\|p\|_K}{\|p\|_{L^2(\mu)}} = \max_{z\in K} \sqrt{K_n^\mu(z)}$$

where

$$K_n^\mu(z) := \sum_{j=1}^N |q_j(z)|^2$$

is the diagonal of the reproducing (Bergman) kernel for $\mathcal{P}_n$, sometimes called the (reciprocal of the) Christoffel function, and $Q_n = \{q_1, q_2, \cdots, q_N\}$ is an orthonormal basis for $\mathcal{P}_n$.

It is natural to ask among all probability measures on K, which one provides the smallest such factor, and this leads to our first

Motivational Definition *Suppose that the probability measure μ has the property that*

$$\max_{z\in K} \sqrt{K_n^\mu(z)} \le \max_{z\in K} \sqrt{K_n^{\mu'}(z)}$$

for all other probability measures μ' on K. Then we say that μ is an optimal measure of degree n for K.

Note that for *any* probability measure μ, $\int_K K_n^\mu(z)d\mu = N$, so that

$$\max_{z\in K} K_n^\mu(z) \geq N.$$

It turns out that for an optimal measure according to Definition 1.1 with $w \equiv 1$,

$$\max_{z\in K} K_n^\mu(z) = N \tag{5}$$

(see Proposition 3.1). We remark that optimal measures need not be discrete.

2.1 A Second Optimality Property

We show that a measure satisfying (5) also satisfies the extremal property in Definition 1.1 with $w \equiv 1$. To see this let

$$B_n = \{p_1, p_2, \cdots, p_N\}$$

be a basis for $\mathcal{P}_n$ and consider the associated Gram matrix

$$G_n^\mu(B_n) := [\langle p_i, p_j\rangle_\mu] \in \mathbb{C}^{N\times N}.$$

Note that $G_n^\mu(B_n)$ is a positive definite Hermitian matrix. If we expand p_i in the orthonormal basis Q_n we obtain

$$p_i = \sum_{k=1}^N \langle p_i, q_k\rangle_\mu q_k \tag{6}$$

so that

$$\begin{aligned} \langle p_i, p_j\rangle_\mu &= \sum_{k=1}^N \langle p_i, q_k\rangle_\mu \langle q_k, p_j\rangle_\mu \\ &= \sum_{k=1}^N \langle p_i, q_k\rangle_\mu \overline{\langle p_j, q_k\rangle_\mu}. \end{aligned}$$

It follows that we have the factorization

$$G_n^\mu(B_n) = V_n^\mu (V_n^\mu)^* \tag{7}$$

where "*" denotes conjugate transpose and

$$V_n^\mu = V_n^\mu(B_n, Q_n) := [\langle p_i, q_j\rangle_\mu] \in \mathbb{C}^{N\times N}. \tag{8}$$

If now μ' is another probability measure on K with associated inner-product $\langle f, g\rangle_{\mu'}$ and orthonormal basis $Q_n' = \{q_1', q_2', \cdots, q_N'\}$, then from the expansion (6) we obtain

$$\begin{aligned}
(V_n^{\mu'})_{ij} &= \langle p_i, q_j'\rangle_{\mu'} \\
&= \sum_{k=1}^N \langle p_i, q_k\rangle_\mu \langle q_k, q_j'\rangle_{\mu'} \\
&= \sum_{k=1}^N (V_n^\mu)_{ik} A_{kj}
\end{aligned}$$

where

$$A = A(Q_n, Q_n', \mu, \mu') := [\langle q_i, q_j'\rangle_{\mu'}] \in \mathbb{C}^{N\times N}.$$

Hence we have the transition

$$V_n^{\mu'} = V_n^\mu A. \tag{9}$$

Now, the transition matrix has the property that

$$\begin{aligned}
\sum_{i=1}^N \sum_{j=1}^N |A_{ij}|^2 &= \sum_{i=1}^N \left\{ \sum_{j=1}^N |\langle q_i, q_j'\rangle_{\mu'}|^2 \right\} \\
&= \sum_{i=1}^N |\langle q_i, q_i\rangle_{\mu'}|^2 \quad \text{(by Parseval)} \\
&= \sum_{i=1}^N \int_K |q_i(z)|^2 d\mu' \\
&= \int_K K_n^\mu(z) d\mu'.
\end{aligned}$$

Hence if μ is a measure satisfying (5), we have

$$\mathrm{tr}(A^*A) = \sum_{i=1}^N \sum_{j=1}^N |A_{ij}|^2 \le N$$

for *any* other probability measure μ'. From this it follows that the sum of the eigenvalues

$$\sum_{k=1}^N \lambda_k(A^*A) = \mathrm{tr}(A^*A) \le N$$

and hence, by the Arithmetic-Geometric Mean inequality,

$$\det(A^*A) = \prod_{k=1}^N \lambda_k(A^*A) \le \left(\frac{1}{N}\sum_{k=1}^N \lambda_k(A^*A)\right)^N \le 1,$$

i.e., if μ is a measure satisfying (5) and μ' is any other probability measure, then the determinant of the transition matrix A satisfies

$$|\det(A)| \le 1.$$

Consequently, by (9),

$$|\det(V_n^{\mu'})| \le |\det(V_n^{\mu})|$$

and by the factorization (7)

$$|\det(G_n^{\mu'}(B_n))| \le |\det(G_n^{\mu}(B_n))|,$$

i.e., a measure μ satisfying (5) also maximizes the determinant of the associated Gram matrix as in Definition 1.1 with $w \equiv 1$.

We end this subsection with an observation which will be useful. If we write

$$P(x) = \begin{bmatrix} p_1(x) \\ p_2(x) \\ \cdot \\ \cdot \\ p_N(x) \end{bmatrix} \in \mathbb{C}^N \tag{10}$$

then it is not difficult to see that

$$P(x)^*(G_n^{\mu}(B_n))^{-1}P(x) = K_n^{\mu}(x). \tag{11}$$

For $G := G_n^\mu(B_n)$ and G^{-1} are positive definite, Hermitian matrices; hence $G^{1/2},\ G^{-1/2} := (G^{-1})^{1/2}$ exist; writing $P := P(x)$, we have

$$P^*G^{-1}P = P^*G^{-1/2}G^{-1/2}P = (G^{-1/2}P)^*G^{-1/2}P.$$

To see that the right-hand-side yields $K_n^\mu(x)$, we first observe that since $G = \int_K PP^*d\mu$ the polynomials $\{\tilde{p}_1, \tilde{p}_2, \cdots, \tilde{p}_N\}$ defined by

$$G^{-1/2}P := \begin{bmatrix} \tilde{p}_1(x) \\ \tilde{p}_2(x) \\ \cdot \\ \cdot \\ \tilde{p}_N(x) \end{bmatrix} \in \mathbb{C}^N \tag{12}$$

form an *orthonormal* basis for $\mathcal{P}_n$ in $L^2(\mu)$: for

$$\int_K G^{-1/2}P \cdot (G^{-1/2}P)^*d\mu = G^{-1/2}[\int_K PP^*d\mu]G^{-1/2} = G^{-1/2}GG^{1/2} = I,$$

the $N \times N$ identity matrix. Thus

$$K_n^\mu(x) = \sum_{j=1}^N |\tilde{p}_j(x)|^2 = (G^{-1/2}P)^*G^{-1/2}P.$$

2.2 Optimal Polynomial Interpolation

There is a close connection between optimal measures and Fekete points of polynomial interpolation. Indeed, suppose that μ is a discretely supported (probability) measure of the form

$$\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}, \quad x_i \in K. \tag{13}$$

Then if μ is non-degenerate on the polynomials of degree n, it is easy to see that $q_i = \sqrt{N}\ell_i$, $1 \le i \le N$, where ℓ_i is the ith fundamental Lagrange polynomial for the points $\{x_i\}$, form an orthonormal set with respect to

$\langle\cdot,\cdot\rangle_\mu$. Hence

$$\begin{aligned}(V_n^\mu)_{ij} &= \langle p_i, q_j\rangle_\mu \\ &= \sqrt{N}\langle p_i, \ell_j\rangle_\mu \\ &= \sqrt{N}\frac{1}{N}\sum_{k=1}^N p_i(x_k)\ell_j(x_k) \\ &= \frac{1}{\sqrt{N}}p_i(x_j)\end{aligned}$$

so that V_n^μ is in this case (a multiple of) the Vandermonde matrix for the basis B_n and the points $\{x_i\}$. Hence maximizing $|\det(G_n^\mu)|$ over all discrete probability measures of the form (13) is equivalent to maximizing the modulus of the Vandermonde determinant. A set of N points which do this are called Fekete points of order n for K and the corresponding discrete measure is said to be a Fekete measure of order n. In general, Fekete points are not unique.

With regard to the Christoffel function, we have

$$K_n(z) = \sum_{k=1}^N |q_i(z)|^2 = N\sum_{k=1}^N |\ell_k(z)|^2$$

so that minimizing $\max_{z\in K} K_n(z)$ over discrete measures of the form (13) is equivalent to finding N points for which $\max_{z\in K}\sum_{k=1}^N |\ell_k(z)|^2$ is as small as possible. This problem (for the interval $K=[-1,1]$) was first studied by Fejér [12] and hence we refer to solution points as Fejér points of order n and the corresponding measure as a Fejér measure. We remark that, in general, Fekete measures and Fejér measures need not coincide nor be unique (although they do coincide and are unique for each order n in the univariate case of $K=[-1,1]$), cf. [10].

Further, if we regard the projection π_μ from $C(K)$ to $\mathcal{P}_n$

$$\pi_\mu(f) := \sum_{j=1}^N \langle f, q_j\rangle_\mu q_j = \sum_{j=1}^N f(x_j)\ell_j$$

as a map from $C(K)\to C(K)$, with both spaces equipped with the *uniform* norm, then it is easy to see that

$$\|\pi_\mu\| = \Lambda_n := \max_{z\in K}\sum_{k=1}^N |\ell_k(z)|,$$

the Lebesgue constant for the interpolation process. Points for which Λ_n is as small as possible are called Lebesgue points of order n and will in general be different from both Fekete and Fejér points. We return to Lebesgue constants in a remark at the end of the paper.

2.3 Optimal Experimental Designs

Consider a polynomial $p \in \mathcal{P}_n$ which we write in the form

$$p = \sum_{k=1}^{N} \theta_k p_k$$

for a fixed basis $\{p_1, ..., p_N\}$ of $\mathcal{P}_n$. Suppose that we observe the values of p at $M \geq N$ points $x_j \in K$ with some random errors, i.e., we observe

$$y_j = p(x_j) + \epsilon_j, \quad 1 \leq j \leq N$$

where we assume that the errors ϵ_j are independent normal random variables with mean 0 and variance σ^2. In matrix form this becomes

$$y = X\theta + \epsilon$$

where $y, \theta, \epsilon \in \mathbb{C}^N$ and

$$X = \begin{bmatrix} p_1(x_1) & p_2(x_1) & \cdot & \cdot & \cdot & p_N(x_1) \\ p_1(x_2) & p_2(x_2) & \cdot & \cdot & \cdot & p_N(x_2) \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ p_1(x_M) & p_2(x_M) & \cdot & \cdot & \cdot & p_N(x_M) \end{bmatrix} \in \mathbb{C}^{M \times N}.$$

Our assumption on the error vector ϵ means that

$$\mathrm{cov}(\epsilon) = \sigma^2 I_N \in \mathbb{R}^{N \times N}.$$

Now, the least squares estimate of θ is

$$\widehat{\theta} := (X^*X)^{-1}X^*y$$

and we may compute the covariance matrix

$$\mathrm{cov}(\widehat{\theta}) = \sigma^2 (X^*X)^{-1}.$$

Hence the confidence region of level t for θ is the set

$$\begin{aligned} & \{\theta \in \mathbb{C}^N \,:\, (\theta - \widehat{\theta})^* [\mathrm{cov}(\widehat{\theta})]^{-1} (\theta - \widehat{\theta}) \le t\} \\ = \; & \{\theta \in \mathbb{C}^N \,:\, \sigma^{-2} (\theta - \widehat{\theta})^* (X^*X)(\theta - \widehat{\theta}) \le t\}. \end{aligned}$$

The volume of such a set is proportional to $1/\sqrt{\det(X^*X)}$ and hence maximizing $\det(X^*X)$ is equivalent to choosing the observation points $x_i \in K$ so as to have the most "concentrated" confidence region for the parameter to be estimated.

Note however that the entries of $\dfrac{1}{M}X^*X$ are the discrete inner products of the p_i with respect to the measure

$$\mu = \frac{1}{M} \sum_{k=1}^{M} \delta_{x_k}, \tag{14}$$

i.e., $\dfrac{1}{M}X^*X$ is the Gram matrix associated to this μ. Hence we may think, heuristically, of an optimal measure as that which gives the confidence region of greatest concentration.

There is also a second statistical interpretation of optimal measures. Taking $P(x)$ as in (10), the least squares estimate of the observed polynomial is

$$P(x)^t \widehat{\theta}.$$

We may compute its variance to be

$$\begin{aligned} \mathrm{var}(P(x)^t \widehat{\theta}) &= \sigma^2 P(x)^* (X^*X)^{-1} P(x) \\ &= \frac{1}{M} \sigma^2 P(x)^* (G_n^\mu)^{-1} P(x) \end{aligned}$$

with μ given by (14). From (11)

$$P(x)^* (G_n^\mu)^{-1} P(x) = K_n^\mu(x)$$

so that

$$\mathrm{var}(P(x)^t \widehat{\theta}) = \frac{1}{M} \sigma^2 K_n^\mu(x)$$

and the experiment that minimizes the maximum variance of the estimate of the observed polynomial is exactly the one that minimizes the maximum of K_n^μ.

We hope that the reader is convinced that optimal measures are interesting and worthy of further study. More about optimal experimental design may be found in the monographs [13] and [11]. In the next section we discuss a weighted version of optimal measures.

3 Weighted Optimal Measures

Let $K \subset \mathbb{C}^d$ be compact and non-pluripolar. Fix μ a probability measure on K and w an admissible weight on K. We recall the notation from the introduction. For each n we have the weighted inner product of degree n

$$\langle f, g\rangle_{\mu,w} := \int_K f(z)\overline{g(z)}w(z)^{2n}d\mu.$$

Fixing a basis $B_n = \{p_1, p_2, \cdots, p_N\}$ of $\mathcal{P}_n$ we form the Gram matrix

$$G_n^{\mu,w} = G_n^{\mu,w}(B_n) := [\langle p_i, p_j\rangle_{\mu,w}] \in \mathbb{C}^{N\times N}$$

and the associated weighted Christoffel function

$$K_n^{\mu,w}(z) := \sum_{j=1}^{N} |q_j(z)|^2 w(z)^{2n}$$

where $Q_n = \{q_1, q_2, \cdots, q_N\}$ is an orthonormal basis for $\mathcal{P}_n$ with respect to the inner-product (2). If a probability measure μ has the property that

$$\det(G_n^{\mu',w}) \le \det(G_n^{\mu,w})$$

for all other probability measures μ' on K then μ is said to be an optimal measure of degree n for K and w.

By (the proof of) Lemma 2.1 of [13], Chapter X], the set of matrices

$$\{G_n^{\mu,w} : \mu \text{ is a probability measure on } K\}$$

is compact (and convex). Hence an optimal measure of degree n for K and w always exists. They need not be unique. An equivalent characterization of optimality is given by the Kiefer-Wolfowitz Equivalence Theorem [14].

Proposition 3.1 *Let w be an admissible weight on K. A probability measure μ is an optimal measure of degree n for K and w; i.e.,*

$$(a)\ \det(G_n^{\mu',w}) \le \det(G_n^{\mu,w})$$

for all other probability measures μ' on K, if and only if

$$(b)\ \max_{z\in K} K_n^{\mu,w}(z) = N.$$

We sketch a proof of the equivalence of conditions (a) and (b) following [10] (but see also [13], Theorem 2.1, Chapter X]). These references prove this theorem only in the unweighted case, but the generalization to the weighted case is straightforward. First, with P defined as in (10), the proof of (11) gives

$$w^{2n}P^*(G_n^{\mu,w})^{-1}P = K_n^{\mu,w}. \tag{15}$$

A computation shows that

$$\mu \to \log\det G_n^{\mu,w}$$

is concave on the space of probability measures; i.e., if

$$h(t) := \log\det G_n^{t\mu_1+(1-t)\mu_2,w}$$

for two probability measures μ_1 and μ_2, then $h''(t) \le 0$. Hence μ_1 is optimal in the sense of (a) if and only if $h'(t) \le 0$ for all μ_2. Computing this derivative one sees that μ_1 is optimal in the sense of (a) if and only if

$$\mathrm{trace}[(G_n^{\mu_1,w})^{-1}G_n^{\mu_2,w}] = \int_K w^{2n}P^*(G_n^{\mu_1,w})^{-1}P d\mu_2 = \int K_n^{\mu_1,w} d\mu_2 \le N \tag{16}$$

for all μ_2. Here we use (15) and the fact that, for an $N\times N$ matrix A, an $N\times 1$ matrix B, and a $1\times N$ matrix C,

$$\mathrm{trace}(ABC) = \mathrm{trace}(CAB) = CAB;$$

thus, writing $G_j := G_n^{\mu_j,w}$ and using $G_2 = \int_K w^{2n} PP^* d\mu_2$,

$$\mathrm{trace}[(G_n^{\mu_1,w})^{-1} G_n^{\mu_2,w}] = \mathrm{trace}[G_1^{-1} \int_K w^{2n} PP^* d\mu_2] = \int_K w^{2n} P^* G_1^{-1} P d\mu_2.$$

Taking μ_2 to be a point mass at a point $z \in K$ in (16) gives $K_n^{\mu_1,w}(z) \le N$; then taking $\mu_2 = \mu_1$ gives $\int K_n^{\mu_1,w} d\mu_1 = N$ by orthonormality. This proves the equivalence of (a) and (b).

Indeed, the end of this argument yields the following key property of optimal measures.

Lemma 3.2 *Suppose that μ is optimal for K and w. Then*

$$K_n^{\mu,w}(z) = N, \quad a.e.\ [\mu].$$

Proof. On the one hand

$$\max_{z \in K} K_n^{\mu,w}(z) = N$$

while on the other hand, again by orthonormality of the q_j,

$$\int_K K_n^{\mu,w}\, d\mu = \int_K \sum_{j=1}^{N} |q_j(z)|^2 w(z)^{2n}\, d\mu(z) = N,$$

and the result follows. ■

We recall that for a basis B_n and a set of points $Z_n = \{z_i \,:\, 1 \le i \le N\} \subset K$ the matrix

$$V_n = V_n(B_n, Z_n) = [p_i(z_j)] \in \mathbb{C}^{N \times N}$$

is called the Vandermonde matrix of the system. In case that the basis B_n is the *standard* monomial basis for $\mathcal{P}_n$ then we will write

$$VDM(z_1, z_2, \cdots, z_N) := \det(V_n).$$

Of fundamental importance for us will be

Definition 3.3 *Suppose that $K \subset \mathbb{C}^d$ is compact and that w is an admissible weight function on K. We set*

$$\delta_n^w(K) := \left(\max_{z_i \in K} |VDM(z_1, \cdots, z_N)| w^n(z_1) w^n(z_2) \cdots w^n(z_N)\right)^{1/m_n}$$

where $m_n = dnN/(d+1)$ is the sum of the degrees of the N monomials of degree at most n. Then

$$\delta^w(K) = \lim_{n\to\infty} \delta_n^w(K)$$

is called the weighted transfinite diameter of K. We refer to $\delta_n^w(K)$ as the weighted nth order diameter of K.

A proof that this limit exists may be found in [9] or [2]; it was first proved in the unweighted case ($w \equiv 1$; i.e., $\delta^1(K)$) by Zaharjuta [18].

Given the close connection between Vandermonde matrices and Gram matrices, as explained in the Introduction, it is perhaps not suprising that we have

Proposition 3.4 *Suppose that K is compact and that w is an admissible weight function. Suppose further that μ_n is an optimal measure of degree n for K and w. Take the basis B_n to be the standard basis of monomials for $\mathcal{P}_n$. Then*

$$\lim_{n\to\infty} \det(G_n^{\mu_n, w})^{1/(2m_n)} = \delta^w(K).$$

Proof. We first note the formula (cf. formula (3.3) of [9])

$$\int_{K^N} |VDM(z_1, \cdots, z_N)|^2 w(z_1)^{2n} \cdots w(z_N)^{2n} d\mu_n(z_1) \cdots d\mu_n(z_N)$$
$$= N! \det(G_n^{\mu_n, w}). \tag{17}$$

It follows immediately, since μ_n is a probability measure, that

$$\det(G_n^{\mu_n, w}) \le \frac{1}{N!} (\delta_n^w(K))^{2m_n}. \tag{18}$$

Secondly, note that if $f_1, f_2, \cdots, f_N \in K$ are weighted Fekete points of degree n for K, i.e., points in K for which

$$|VDM(z_1, \cdots, z_N)| w^n(z_1) w^n(z_2) \cdots w^n(z_N)$$

is maximal, then the discrete measure

$$\nu_n = \frac{1}{N}\sum_{k=1}^{N} \delta_{f_k} \tag{19}$$

based on these points is a candidate probability measure for property (a) of Definition 1.1. Hence

$$\det(G_n^{\nu_n,w}) \le \det(G_n^{\mu_n,w}).$$

But, as is easy to see,

$$\begin{aligned}
\det(G_n^{\nu_n,w}) &= \frac{1}{N^N}|VDM(f_1,\cdots,f_N)|^2 w(f_1)^{2n} w(f_2)^{2n}\cdots w(f_N)^{2n}\\
&= \frac{1}{N^N}\left(\max_{z_i\in K}|VDM(z_1,\cdots,z_N)|w^n(z_1)w^n(z_2)\cdots w^n(z_N)\right)^2\\
&= \frac{1}{N^N}(\delta_n^w(K))^{2m_n}.
\end{aligned}$$

Hence,

$$\frac{1}{N^N}(\delta_n^w(K))^{2m_n} \le \det(G_n^{\mu_n,w}) \le \frac{1}{N!}(\delta_n^w(K))^{2m_n}$$

by combining the lower bound with the upper bound (18). ■

Of course, it then follows that

$$\lim_{n\to\infty}\frac{1}{2m_n}\log\,\det(G_n^{\mu_n,w}) = \log(\delta^w(K)). \tag{20}$$

Now, suppose that $u \in C(K)$ and that w is an admissible weight function. Following the ideas in [1], [2], [3], [4], [5] we consider the weight $w_t(z) := w(z)\exp(-tu(z))$, $t \in \mathbb{R}$, and let μ_n be an optimal measure of degree n for K and w. We set

$$f_n(t) := -\frac{1}{2m_n}\log\,\det(G_n^{\mu_n,w_t}). \tag{21}$$

For $t = 0$, $w_0 = w$ and (20) says

$$\lim_{n\to\infty} f_n(0) = -\log(\delta^w(K)).$$

We have the following (see Lemma 6.4 in [1]).

Lemma 3.5 *We have*

$$f_n'(t) = \frac{d+1}{dN}\int_K u(z)K_n^{\mu_n,w_t}(z)d\mu_n.$$

In particular,

$$\begin{aligned} f_n'(0) &= \frac{d+1}{dN}\int_K u(z)K_n^{\mu_n,w}(z)d\mu_n \\ &= \frac{d+1}{d}\int_K u(z)d\mu_n \quad \textit{(by Lemma 3.2).} \end{aligned} \tag{22}$$

Proof. Recall that $G_n^{\mu_n,w_t}$ is a positive definite Hermitian matrix; hence it can be diagonalized by a unitary matrix and we can define $\log(G_n^{\mu_n,w_t})$. Using $\log\det(G_n^{\mu_n,w_t}) = \operatorname{trace}\log(G_n^{\mu_n,w_t})$, we calculate

$$\begin{aligned} 2m_nf_n'(t) &= -\frac{d}{dt}\operatorname{trace}\left(\log(G_n^{\mu_n,w_t})\right) \\ &= -\operatorname{trace}\left(\frac{d}{dt}\log(G_n^{\mu_n,w_t})\right) \\ &= -\operatorname{trace}\left((G_n^{\mu_n,w_t})^{-1}\frac{d}{dt}G_n^{\mu_n,w_t}\right) \\ &= 2n\operatorname{trace}\left((G_n^{\mu_n,w_t})^{-1}\left[\int_K p_i(z)\overline{p_j(z)}u(z)w(z)^{2n}\exp(-2ntu(z))d\mu_n\right]\right) \end{aligned}$$

As in the proof of Proposition 3.1 we use

$$\operatorname{trace}(ABC) = \operatorname{trace}(CAB) = CAB$$

to write the previous line as

$$\begin{aligned} &= 2n\int_K P^*(z)(G_n^{\mu_n,w_t})^{-1}P(z)u(z)w(z)^{2n}\exp(-2ntu(z))d\mu_n \\ &= 2n\int_K u(z)P^*(z)(G_n^{\mu_n,w_t})^{-1}P(z)w_t(z)^{2n}d\mu_n \\ &= 2n\int_K u(z)K_n^{\mu_n,w_t}(z)d\mu_n \end{aligned}$$

where the last equality follows from the remark (15).

The result follows from the fact that $m_n = dnN/(d+1)$. ■

The next result was proved in a slightly different way in [5], Lemma 2.2.

Lemma 3.6 *The functions* $f_n(t)$ *are concave, i.e.,* $f_n''(t) \le 0$.

Proof. First, let
$$g_n(h) := 2m_n f_n(t+h)$$
so that $f_n''(t) = \dfrac{1}{2m_n} g_n''(0)$. Also, note that if we change the basis $B_n = \{p_1, \cdots, p_N\}$ to $C_n := \{q_1, \cdots, q_N\}$ by $p_i = \sum_{j=1}^N a_{ij} q_j$, then the Gram matrices transform (see e.g. [D, §8.7]) by

$$G_n^{\mu_n, w_t}(B_n) = A G_n^{\mu_n, w_t}(C_n) A^*$$

where $A = [a_{ij}] \in \mathbb{C}^{N\times N}$. Hence,

$$g_n(h) = -\log(\det(G_n^{\mu_n, w_{t+h}}(B_n))) = -\log(\det(G_n^{\mu_n, w_{t+h}}(C_n))) - \log(|\det(A)|^2)$$

and we see that the derivatives of g_n are independent of the basis chosen.

Let us choose C_n to be an orthonormal basis for $\mathcal{P}_n$ with respect to the inner-product $\langle \cdot, \cdot \rangle_{\mu_n, w} = \langle \cdot, \cdot \rangle_{\mu_n, w_t}$.

Now, for convenience, write $G(h) = G_n^{\mu_n, w_{t+h}}$ and set $F(h) = \log(G(h))$ so that $G(h) = \exp(F(h))$. By our choice of basis C_n we have $G(0) = I \in \mathbb{C}^{N\times N}$, the identity matrix, and $F(0) = [0] \in \mathbb{C}^{N\times N}$, the zero matrix. Then, (see e.g. [6], p. 311]),

$$\frac{dG}{dh} = \frac{d}{dh}\exp(F(h)) = \int_0^1 e^{(1-s)F(h)} \frac{dF}{dh} e^{sF(h)} ds.$$

In particular
$$\frac{dG}{dh}(0) = \frac{dF}{dh}(0).$$

Further,

$$\begin{aligned} \frac{d^2G}{dh^2} &= \int_0^1 \left\{ \left[\frac{d}{dh} e^{(1-s)F(h)}\right] \frac{dF}{dh} e^{sF(h)} + e^{(1-s)F(h)} \frac{d^2F}{dh^2} e^{sF(h)} \right. \\ &\qquad \left. + e^{(1-s)F(h)} \frac{dF}{dh} \left[\frac{d}{dh} e^{sF(h)}\right] \right\} ds. \end{aligned}$$

Evaluating at $h=0$, using the fact that $F(0)=[0]$, we obtain

$$\begin{aligned}
\frac{d^2G}{dh^2}(0) &= \int_0^1 \left\{(1-s)\frac{dF}{dh}(0)\times\frac{dF}{dh}(0)\times I + I\times\frac{d^2F}{dh^2}(0)\times I \right.\\
&\quad \left. +I\times\frac{dF}{dh}(0)\times s\frac{dF}{dh}(0)\right\} ds\\
&= \int_0^1 \left\{(1-s+s)\left(\frac{dF}{dh}(0)\right)^2 + \frac{d^2F}{dh^2}(0)\right\} ds\\
&= \left(\frac{dF}{dh}(0)\right)^2 + \frac{d^2F}{dh^2}(0).
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{d^2F}{dh^2}(0) &= \frac{d^2G}{dh^2}(0) - \left(\frac{dF}{dh}(0)\right)^2\\
&= [\int_K q_i(z)\overline{q_j(z)}(-2nu(z))^2 w_t(z)^{2n} d\mu_n] - [\int_K q_i(z)\overline{q_j(z)}(-2nu(z))w_t(z)^{2n}d\mu_n]^2.
\end{aligned}$$

Since $g_n'(h) = \frac{d}{dh}[-\log(\det(G(h))]$ and $\log(\det(G(h)) = \mathrm{trace}(\log(G(h))) = \mathrm{trace}(F(h))$ it follows that

$$\begin{aligned}
g_n''(0) &= -\mathrm{trace}\left(\left[\int_K q_i(z)\overline{q_j(z)}(-2nu(z))^2 w_t(z)^{2n} d\mu_n\right]\right)\\
&\quad +\mathrm{trace}\left(\left[\int_K q_i(z)\overline{q_j(z)}(-2nu(z))w_t(z)^{2n} d\mu_n\right]^2\right)\\
&= -\sum_{i=1}^N \int_K |q_i(z)|^2 w_t(z)^{2n}(2nu(z))^2 d\mu_n\\
&\quad +\sum_{i=1}^N\sum_{j=1}^N \left|\int_K q_i(z)\overline{q_j(z)} w_t(z)^{2n}(2nu(z))d\mu_n\right|^2\\
&= -\sum_{i=1}^N \left\{\int_K |q_i(z)|^2 w_t(z)^{2n}(2nu(z))^2 d\mu_n - \right.\\
&\quad \left. \sum_{j=1}^N \left|\int_K q_i(z)\overline{q_j(z)} w_t(z)^{2n}(2nu(z))d\mu_n\right|^2\right\}.
\end{aligned}$$

But notice that $\int_K q_i(z)\overline{q_j(z)}w_t(z)^{2n}(2nu(z))d\mu_n$ is the jth Fourier coefficient of the function $2nu(z)q_i(z)$ with respect to the orthonormal basis C_n, and also that $\int_K |q_i(z)|^2 w_t(z)^{2n}(2nu(z))^2 d\mu_n$ is the L^2 norm squared of this same function. Hence, by Parseval's inequality, $g_n''(0) \leq 0$. ■

4 The Limit of Optimal Measures

In this section we prove the main theorem. Let $K \subset \mathbb{C}^d$ be compact with admissible weight function $w := e^{-\phi}$. Recall from the introduction that the weighted extremal function $V_{K,\phi}^*(z)$ is the usc regularization of $V_{K,\phi}$ in (1), and the weighted equilibrium measure is

$$\mu_{K,\phi} := \frac{1}{(2\pi)^d}(dd^c V_{K,\phi}^*)^d.$$

Berman and Boucksom [3] have recently shown that the discrete probability measures based on the weighted Fekete points (19) tend weak$-*$ to $\mu_{K,\phi}$. This is based on a remarkable sequence of papers (see [1], [2], [3], [4]). Indeed, the argument in [3] shows that if for each n, we take points $x_1^{(n)}, x_2^{(n)}, \cdots, x_N^{(n)} \in K$ for which

$$\lim_{n\to\infty} [|VDM(x_1^{(n)}, \cdots, x_N^{(n)})| w(x_1^{(n)})^n w(x_2^{(n)})^n \cdots w(x_N^{(n)})^n]^{1/m_n} = \delta^w(K) \tag{23}$$

(*asymptotically* weighted Fekete points), then the discrete measures

$$\nu_n = \frac{1}{N}\sum_{k=1}^N \delta_{x_k^{(n)}}$$

converge weak$-*$ to $\mu_{K,\phi}$. The main point of this note is to remark that their proof may be extended to also give the limit of optimal measures. For completeness we give the details of the proof, but we emphasize that it is their same argument as for the Fekete measure case (see also [4]).

Main Theorem. *Suppose that $K \subset \mathbb{C}^d$ is compact and that w is an admissible weight function. We again set $\phi := -\log(w)$. Suppose further* *that μ_n is an optimal measure of degree n for K and w. Then*

$$\lim_{n\to\infty} \mu_n = \mu_{K,\phi}$$

where the limit is in the weak$-*$ sense.

Proof. For $u \in C^2(K)$ we again set $w_t(z) := w(z)\exp(-tu(z))$ which corresponds to $\phi_t := \phi + tu$ and $f_n(t)$ as in (21). As mentioned in (20),

$$\lim_{n\to\infty} f_n(0) = -\log(\delta^w(K)).$$

Of fundamental importance is the Rumely formula for the transfinite diameter ([16], [1], [2]):

$$-\log(\delta^w(K)) = \frac{1}{d(2\pi)^d}\mathcal{E}(V_{K,\phi}^*, V_T). \tag{24}$$

Here V_T is the (unweighted) extremal function for a polydisc that contains K and $\mathcal{E}$ is a certain "mixed energy" whose exact formula is not important here. What is important is the derivative formula of Berman and Boucksom [1], [2]:

$$\frac{d}{dt}\mathcal{E}(V_{K,\phi+tu}, V_T)\bigg|_{t=0} = (d+1)\int_K u(dd^cV_{K,\phi}^*)^d. \tag{25}$$

In other words, setting $g(t) = -\log(\delta^{w_t}(K))$,

$$g'(0) = \frac{d+1}{d(2\pi)^d}\int_K u(z)(dd^cV_{K,\phi}^*)^d. \tag{26}$$

Now note that for each fixed t, the measure μ_n, being optimal for K and $w = w_0$, is a candidate for the optimal measure for K and w_t. If follows from property (a) of Definition 1.1 that

$$\det(G_n^{\mu_n,w_t}) \le \det(G_n^{\mu_n^t,w_t})$$

where we denote an optimal measure for K and w_t by μ_n^t. Hence (see (21))

$$f_n(t) \ge -\frac{1}{2m_n}\log(\det(G_n^{\mu_n^t,w_t}))$$

and consequently from Proposition 3.4 that

$$\liminf_{n\to\infty} f_n(t) \ge -\log(\delta^{w_t}(K)) = g(t). \tag{27}$$

It now follows from Lemma 4.1 in [3] (see Lemma 4.1 below) that

$$\lim_{n\to\infty} f_n'(0) = g'(0).$$

In other words, by Lemma 3.5,

$$\begin{aligned}\lim_{n\to\infty} \frac{d+1}{d}\int_K u(z)d\mu_n &= \frac{d+1}{d(2\pi)^d}\int_K u(z)(dd^cV_{K,\phi}^*)^d \\ &= \frac{d+1}{d}\int_K u(z)d\mu_{K,\phi},\end{aligned}$$

and hence $\mu_n \to \mu_{K,\phi}$ weak$-*$. ■

Lemma 4.1 *(Berman and Boucksom [3]) Let $f_n(t)$ be a sequence of concave functions on $\mathbb{R}$ and $g(t)$ a function on $\mathbb{R}$. Suppose that*

$$\liminf_{n\to\infty} f_n(t) \geq g(t), \quad \forall t \in \mathbb{R}$$

and that

$$\lim_{n\to\infty} f_n(0) = g(0).$$

Suppose further that the f_n and g are differentiable at $t = 0$. Then

$$\lim_{n\to\infty} f_n'(0) = g'(0).$$

Remark. In [1], [2] the derivative formula (25) is proved in a very general setting under the assumption that ϕ is continuous. However, in our setting, their proof remains valid for lowersemicontinuous ϕ and hence our main theorem remains true for general usc weights w.

Remark. The reader will note that the key properties of optimal measures used here are Lemma 3.2, used in the proof of Lemma 3.5, and Proposition 3.4, which is used in the proof of (27): if μ is optimal for K and w then

$$K_n^{\mu,w}(z) = N, \quad a.e.\ [\mu]$$

and

$$\lim_{n\to\infty} \det(G_n^{\mu_n,w})^{1/(2m_n)} = \delta^w(K).$$

These properties are also satisfied for asymptotically weighted Fekete measures (measures associated to points satisfying (23))

$$\nu_n = \frac{1}{N} \sum_{k=1}^{N} \delta_{f_k}.$$

Thus weak-* convergence to $\mu_{K,\phi}$ for both sequences $\{\mu_n\}$ and $\{\nu_n\}$ follows from (24) and (25).

There exist many other natural sequences of measures $\{\mu_n\}$ which converge weak-* to $\mu_{K,\phi}$. For simplicity, we discuss the unweighted case ($\phi = 0$). Recall from subsection 1.2 that if $x_1, ..., x_N \in K$, then $\Lambda_n := \max_{z \in K} \sum_{k=1}^{N} |\ell_k(z)|$ is the so-called Lebesgue constant associated to polynomial interpolation at these points. Suppose for each $n = 1, 2, ...$ we have $N = N(n)$ points $x_1^{(n)}, ..., x_N^{(n)} \in K$ with Lebesgue constant Λ_n. An elementary argument in [8] shows that if $\limsup_{n\to\infty} \Lambda_n^{1/n} \leq 1$, then

$$\lim_{n\to\infty} |VDM(x_1^{(n)}, ..., x_N^{(n)})|^{1/m_n} = \delta^1(K), \tag{28}$$

i.e., subexponential growth of the Lebesgue constants implies the array of points is asymptotically Fekete ((23) holds with $w \equiv 1$). By the main result of [4], this asymptotic Fekete property (28) implies that the discrete measures

$$\mu_n := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^{(n)}}$$

converge weak-* to μ_K. It is easy to see that the Lebesgue constants for either the Lebesgue or Fejer points satisfy the subexponential growth of the Lebesgue constants so the weak-* convergence to the equilibrium measure holds for these arrays. Furthermore, in Proposition 3.7 of [8] it was shown that for a *Leja sequence* $\{x_1, x_2, ...\} \subset K$,

$$\lim_{n\to\infty} |VDM(x_1, ..., x_N)|^{1/m_n} = \delta^1(K).$$

Thus the asymptotic Fekete property (28) holds for this *sequence* of points; so, again from [4], it follows that the discrete measures

$$\mu_n := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$$

converge weak-* to μ_K. Such a sequence is defined inductively as follows. Take the standard monomial basis $\{p_1, p_2, ...\}$ for $\cup_{n=0}^{\infty}\mathcal{P}_n$ ordered so that $\deg p_i \leq \deg p_j$ if $i \leq j$. Given m points $z_1, ..., z_m$ in $\mathbb{C}^d$, as before we write

$$VDM(z_1, ..., z_m) = \det[p_i(z_j)]_{i,j=1,...,m}.$$

Starting with any point $x_1 \in K$, having chosen $x_1, ..., x_m \in K$ we choose $x_{m+1} \in K$ so that

$$|VDM(x_1, ..., x_m, x_{m+1})| = \max_{x\in K} |VDM(x_1, ..., x_m, x)|.$$

We remark that despite possessing the desirable property that $\mu_n \to \mu_K$ weak-*, it is unknown if $\limsup_{n\to\infty} \Lambda_n^{1/n} \leq 1$ always holds for a Leja sequence, even in the univariate case ($d = 1$).

REFERENCES

[1] Berman, R. and Boucksom, S., *Growth of balls of holomorphic sections and energy at equilibrium*, preprint.

[2] Berman, R. and Boucksom, S., *Capacities and Weighted Volumes of Line Bundles*, preprint arXiv:0803.1950.

[3] Berman, R. and Boucksom, S., *Equidistribution of Fekete Points on Complex Manifolds*, preprint arXiv:0807.0035.

[4] Berman, R., Boucksom, S. and Nystrom, D. W., *Convergence towards equilibrium on complex manifolds*, preprint.

[5] Berman, R. and Nystrom, D. W., *Convergence of Bergman measures for high powers of a line bundle*, preprint arXiv:0805.2846.

[6] Bhatia, R., **Matrix Analysis**, GTM 169, Springer, New York, 1997.

[7] Bloom, T., *Orthogonal polynomials in* $\mathbb{C}^n$, Indiana Univ. Math. J., Vol. 46, No. 2 (1997), 427 – 452.

[8] Bloom, T., Bos, L., Christensen, C. and Levenberg, N., *Polynomial interpolation of holomorphic functions in* $\mathbb{C}$ *and* $\mathbb{C}^n$, Rocky Mtn. J. Math., 22 (1992), 441 – 470.

[9] Bloom, T. and Levenberg, N., *Transfinite diameter notions in* $\mathbb{C}^N$ *and integrals of Vandermonde determinants*, to appear in Arkiv för Matematik, DOI: 10.1007/s11512-009-0101-9.

[10] Bos, L., *Some Remarks on the Fejér Problem for Lagrange Interpolation in Several Variables*, J. Approx. Theory, Vol. 60, No. 2 (1990), 133 – 140.

[11] Dette, H. and Studden, W.J., **The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis**, Wiley Interscience, New York, 1997.

[12] Fejér, L., *Bestimmung derjenigen Abszissen eines Intervalles, für welche die Quadratsumme der Grundfunktionen der Lagrangeschen Interpolation im Intervalle ein moglichst kleines Maximum besitzt*, Ann. Scuoal Norm. Sup. Pisa (2) **1** (1932), 263 – 276.

[13] Karlin, S. and Studden, W.J., **Tchebycheff Systems: With Applications in Analysis and Statistics**, Wiley Interscience, New York, 1966.

[14] Kiefer, J. and Wolfowitz, J., *The equivalence of two extremum problems*, Canad. J. Math. **12** (1960), 363 – 366.

[15] Klimek, M., **Pluripotential Theory**, Oxford Univ. Press, 1991.

[16] Rumely, R., *A Robin Formula for the Fekete-Leja Transfinite Diameter*, Math. Ann. **337** no. 4 (2007), 729 – 738.

[17] Saff, E. and Totik, V., **Logarithmic Potentials with External Fields**, Springer, 1997.

[18] V. P. Zaharjuta, *Transfinite diameter, Chebyshev constants, and capacity for compacta in* $\mathbb{C}^n$, Math. USSR Sbornik, **25** (1975), no. 3, 350 – 364.