BERNSTEIN-SZEGŐ MEASURES ON THE TWO DIMENSIONAL TORUS

GREG KNESE

ABSTRACT. We present a new viewpoint (namely, reproducing kernels) and new proofs for several recent results of J. Geronimo and H. Woerdeman on orthogonal polynomials on the two dimensional torus (and related subjects). In addition, we show how their results give a new proof of Andô's inequality via an equivalent version proven by Cole and Wermer. A major theme is the use of so-called Bernstein-Szegő measures. A simple necessary and sufficient condition for two variable polynomial stability is also given.

1. INTRODUCTION

1.1. **Prelude.** In several recent papers, J. Geronimo and H. Woerdeman have presented a number of important generalizations of classical theorems revolving around such topics as orthogonal polynomials on the unit circle, Fejér-Riesz factorization, and autoregressive filter design to the context of two variables; namely, generalizations to the two dimensional torus or just the *2-torus* $\mathbb{T}^2 = (\partial\mathbb{D})^2 \subset \mathbb{C}^2$ (also known as the distinguished boundary of the bidisk $\mathbb{D}^2$). See [5], [7], [6]. One of the goals of this paper is to present and prove many of their results from a different viewpoint[1] so that a larger audience may appreciate what they have done.

A major theme in this paper is the use of Bernstein-Szegő measures, which are nothing more than absolutely continuous measures on the 2-torus (or even the *n*-torus for that matter) with weight given by the reciprocal of the squared modulus of a polynomial. Studying such measures allows us to give a new proof of Andô's inequality from operator

[1]The perspective of Geronimo and Woerdeman may best be described as matrix polynomial theory, while our own perspective may be described as reproducing kernel theory. Our approach is partially inspired by the book [9].

theory via an equivalent version proven by B. Cole and J. Wermer. The
proof here is somewhat more elementary than other proofs (although
not necessarily shorter). As is well-known, Andô's inequality is inti-
mately related to finite interpolation problems for bounded analytic
functions on the bidisk (Pick interpolation), and therefore this work
provides a new connection between ideas in orthogonal polynomials
and interpolation.

All of this may sound quite specialized; however, the results in this
paper have consequences that anyone can appreciate. Let us whet the
reader's appetite with the following interesting by-product of our work
(which will also serve to introduce the terms *degree* (n, m), *reflection*,
and *stable*).

Theorem 1.1. *Let q be a polynomial in two complex variables of de-
gree (n, m), i.e.*

$$q\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right) = \sum_{j=0}^{n} \sum_{k=0}^{m} a_{j,k} z^j w^k$$

and define the reflection $\vec{q}$ *of q to be*

$$\vec{q}\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right) = z^n w^m \overline{q\left(\begin{smallmatrix}1/\bar{z}\\1/\bar{w}\end{smallmatrix}\right)} = \sum_{j=0}^{n} \sum_{k=0}^{m} \bar{a}_{(n-j),(m-k)} z^j w^k$$

*The polynomial q has no zeros in the closed bidisk $\overline{\mathbb{D}}^2$ (i.e. q is
stable) if and only if there exists a $c > 0$ such that*

(1.1) $$\left|q\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)\right|^2 - \left|\vec{q}\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)\right|^2 \geq c(1 - |z|^2)(1 - |w|^2),$$

in which case we can take

$$\frac{1}{c} = \frac{1}{(2\pi i)^2} \int_{\mathbb{T}^2} \frac{1}{\left|q\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)\right|^2} \frac{dz}{z} \frac{dw}{w}.$$

See Section 7 for a proof. The significance of this theorem is that it
says when a polynomial has no zeros on the *boundary* of the bidisk. At
the same time this says something very specific about two variables:
the analogue of this theorem in three variables is probably not true,
but, while the analogue of this theorem in one variable *is* true, the
condition (1.1) can be weakened to remove the term $|\vec{q}|^2$.

Interestingly enough, the one variable version of this theorem can be
used to prove the fundamental theorem of algebra. This two variable
version can be used to prove the following: any two variable polynomial
with no zeros on $\overline{\mathbb{D}}^2 \cup (\mathbb{C} \setminus \mathbb{D})^2$ must be constant.

1.2. **A key to notations and conventions.** All of the following notations will be introduced throughout the paper, but we collect them here for the benefit of the reader.

Notations/Definitions:

$$\mathbb{C} = \text{ complex plane}$$

$$\mathbb{D} = \text{ open unit disk in } \mathbb{C}$$

$$\overline{\mathbb{D}} = \text{closed unit disk}$$

$$\mathbb{T} = \partial\mathbb{D} = \text{boundary of the disk in } \mathbb{C}$$

$$\mathbb{D}^2 = \mathbb{D} \times \mathbb{D} = \text{ the bidisk}$$

$$\overline{\mathbb{D}}^2 = \text{ the closed bidisk}$$

$$\mathbb{T}^2 = \mathbb{T} \times \mathbb{T} = \text{ the 2-torus}$$

$$\vec{q}\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right) = z^n w^m \overline{q\left(\begin{smallmatrix} 1/\bar{z} \\ 1/\bar{w} \end{smallmatrix}\right)} = \text{ reflection of } q$$

$$V \vee W = \text{ the join or span of two vector spaces}$$

$$(f \cdot V) = \{fP : P \in V\} \text{ for a vector space } V$$

$$q \text{ is stable} = q \text{ has no zeros in } \overline{\mathbb{D}}^2$$

$$\square = \text{span}\{z^j w^j : 0 \le j \le n, 0 \le k \le m\}$$

$$= \text{ polynomials of degree } (n, m)$$

$$K_\rho \square = \text{ reproducing kernel for } \square \text{ with respect to a measure } \rho$$

These last two somewhat strange looking notations will make more sense later. Many other important notations are defined in Section 2 but we do not present them here.

Conventions:

n, m are positive integers, fixed throughout the paper

p is a holomorphic polynomial in two variables of particular importance

q, P, Q are generic holomorphic polynomials of two variables

f is a typical element of H^2 or L^2

z, w are the typical holomorphic variables

Z, W are the typical anti-holomorphic variables

ρ is a probability measure on $\mathbb{T}^2$

μ is a Bernstein-Szegő measure on $\mathbb{T}^2$ (see (4.1))

$\binom{z}{w}$ is an element of $\mathbb{C}^2$ and *not* a binomial coefficient

$\dfrac{1}{2\pi i}\dfrac{dz}{z}$ is normalized Lebesgue measure on $\mathbb{T}$

$\langle, \rangle$ is an inner product

$\langle, \rangle_\rho$ is an inner product given by ρ

$\perp_\rho$ is orthogonality with respect to ρ

1.3. **Probability measures on $\mathbb{T}^2$.** The story begins with a probability measure ρ on the 2-torus $(\partial\mathbb{D})^2 = \mathbb{T}^2$. Let $\langle f, g \rangle_\rho$ denote the standard inner product on $L^2(\rho)$ given by $\int_{\mathbb{T}^2} f\bar{g}d\rho$ and let $||f||$ as usual denote $\sqrt{\langle f, f \rangle_\rho}$. Fix positive integers m, n throughout the paper, and define the complex polynomials of degree (n, m)

$$\square := \operatorname{span}\{z^j w^k : 0 \le j \le n, 0 \le k \le m\}.$$

The following nondegeneracy condition will often be imposed on ρ:

Definition 1.2. The measure ρ is said to be *nondegenerate* (at the (n, m) level) if a polynomial $P \in \square$ can have norm $||P|| = 0$ only if $P \equiv 0$.

This is just another way of saying that the Toeplitz moment matrix for ρ corresponding to the product index set $\{0, \ldots, n\} \times \{0, \ldots, m\}$ is positive definite.

Let us now define a particular kind of measure that will be of the utmost importance in this paper. Recall that a polynomial in two variables is *stable* if it has no zeros on the closed bidisk.

Definition 1.3. Given a stable polynomial q the *Bernstein-Szegő* measure corresponding to q is the following probability measure on $\mathbb{T}^2$:

$$(1.2) \qquad d\mu := \frac{c^2}{(2\pi i)^2 |q\binom{z}{w}|^2} \frac{dz}{z} \frac{dw}{w}$$

where $c > 0$ is chosen to make μ a bona fide probability measure.

To see why such measures are of interest, let us present a classical one-variable theorem found in [9] (page 95) for instance.

Theorem 1.4 (Bernstein-Szegő Approximation). *Let $d\nu$ be a nontrivial[2] probability measure on $\partial\mathbb{D}$ and let p be the polynomial (in one complex variable z) of degree n satisfying*

$$p \perp_\nu z, z^2, \ldots, z^n$$
$$p(0) > 0$$
$$||p||_\nu = 1.$$

Then, p has no zeros in the closed disk $\overline{\mathbb{D}}$ and the measure on $\mathbb{T}$

$$d\nu_n := \frac{1}{2\pi i |p(z)|^2} \frac{dz}{z}$$

defines the same inner product on polynomials of degree less than or equal to n as $d\nu$.

Moreover, $d\nu_n \to d\nu$ in the weak* topology.

Obvious analogues of the above theorem do not hold in two variables. The following theorem of Geronimo and Woerdeman says exactly when a measure on $\mathbb{T}^2$ has a "Bernstein-Szegő" approximation. (Actually this is a slight weakening of their Theorem 1.1.2 in [5], but will suffice for our purposes. In Section 6, we state a group of equivalences that help translate between our paper and [5].)

Theorem 1.5 (Geronimo-Woerdeman). *Suppose ρ is nondegenerate (as in Definition 1.2) and consider the polynomial $p \in \square$ defined by the conditions*

$$p \perp_\rho z^j w^k \text{ for } 0 \le j \le n, 0 \le k \le m, (i,j) \ne (0,0)$$
$$p\binom{0}{0} > 0$$
$$||p|| = 1.$$

Then, p is stable (i.e. has no zeros in the closed bidisk $\overline{\mathbb{D}}^2$) and the measure on $\mathbb{T}^2$

$$d\mu := \frac{1}{(2\pi i)^2 |p\binom{z}{w}|^2} \frac{dz}{z} \frac{dw}{w}$$

defines the same inner product on $\square$ as ρ if and only if ρ satisfies the following "automatic orthogonality condition": Every polynomial q in the span of $\{z^i w^j : 0 \le i \le n, 0 \le j \le m-1\}$ satisfying

$$q \perp_\rho z^j w^k \text{ for } 0 \le j \le n-1, 0 \le k \le m-1$$

---

[2]Meaning, not a finite number of point masses.

is automatically orthogonal to more monomials:

$$q \perp_\rho z^j w^m \text{ for } 0 \leq j \leq n-1.$$

In Section 2 we shall introduce a new notation and in Sections 3 and 4 we present several results which we believe provide insight into this somewhat technical sounding, yet important, theorem. Before that, we turn to an interesting connection between this work, interpolation for bounded analytic functions on the bidisk, and two-variable operator theory.

1.4. **Andô, Agler, and Christoffel-Darboux.** An important part of the proof of the Geronimo-Woerdeman Theorem 1.5 is a "Christoffel-Darboux like formula" for two variables. To set the stage and for later use, let us present the Christoffel-Darboux formula for orthogonal polynomials on the unit circle (see [9] Theorem 2.2.7 page 124).

Theorem 1.6 (Christoffel-Darboux formula). *Let ν be a nontrivial probability measure on $\mathbb{T}$, and let p be the unit norm polynomial (of one complex variable) of degree n orthogonal to $z, z^2, \ldots, z^n$ with $p(0) > 0$. Also, let K_{n-1} be the reproducing kernel (with respect to ν) for polynomials of degree at most $n-1$ with respect to ν. Then,*

$$p(z)\overline{p(Z)} - \vec{p}(z)\overline{\vec{p}(Z)} = (1 - z\bar{Z})K_{n-1}(z, Z)$$

where $\vec{p}(z) := z^n \overline{p(1/\bar{z})}$.

The two variable formula proven by Geronimo-Woerdeman has the following flavor. For p and ρ as in the conclusion of Theorem 1.5, there exist polynomials $P_j \in \square$, $j = 0, 1, \ldots n-1$, $Q_k \in \square$, $k = 0, \ldots, m-1$ such that

$$p\begin{pmatrix}z\\w\end{pmatrix}\overline{p\begin{pmatrix}Z\\W\end{pmatrix}} - \vec{p}\begin{pmatrix}z\\w\end{pmatrix}\overline{\vec{p}\begin{pmatrix}Z\\W\end{pmatrix}} = (1 - z\bar{Z})\sum_j P_j\begin{pmatrix}z\\w\end{pmatrix}\overline{P_j\begin{pmatrix}Z\\W\end{pmatrix}}$$

$$+ (1 - w\bar{W})\sum_k Q_k\begin{pmatrix}z\\w\end{pmatrix}\overline{Q_k\begin{pmatrix}Z\\W\end{pmatrix}}$$

where $\vec{p}\begin{pmatrix}z\\w\end{pmatrix} = z^n w^m \overline{p\begin{pmatrix}1/\bar{z}\\1/\bar{w}\end{pmatrix}}$.

For the moment, we do not need to say what P_j and Q_k are (although we can and will later). It turns out this Christoffel-Darboux formula is closely related to Andô's theorem in operator theory and Agler's theorem on finite interpolation for bounded analytic functions on the bidisk.

Andô's theorem [2] says given a polynomial P in two complex variables and two commuting operators S, T on a Hilbert space with operator norms $||S||, ||T|| \leq 1$, the following holds:

$$||P\binom{S}{T}|| \leq \sup_{\mathbb{T}^2} |P|.$$

The Pick interpolation theorem on the bidisk, namely Agler's theorem (see [1] page 180), says given N distinct points $\lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{D}^2$ (where we write $\lambda_i = (z_i, w_i)$) and N points $c_1, \ldots, c_n \in \mathbb{C}$, there exists a holomorphic function f on $\mathbb{D}^2$ with $\sup_{\mathbb{D}^2} |f| \leq 1$ which interpolates $f(\lambda_i) = c_i$ if and only if there exist positive semidefinite $N \times N$ matrices (A_{jk}) and (B_{jk}) so that

$$1 - c_j \bar{c}_k = (1 - z_j \bar{z}_k) A_{jk} + (1 - w_j \bar{w}_k) B_{jk}$$

for $1 \leq j, k \leq N$.

It turns out that both Andô's theorem and Agler's theorem are equivalent to a result about polynomials that looks suspiciously like the Geronimo-Woerdeman formula.

Theorem 1.7 (Cole-Wermer [4]). *Let P and Q be polynomials in two complex variables satisfying*

(1.3) $$|P\binom{z}{w}| \geq |Q\binom{z}{w}| \text{ for all } \binom{z}{w} \in \mathbb{D}^2$$

and

(1.4) $$|P\binom{z}{w}| = |Q\binom{z}{w}| \text{ for all } \binom{z}{w} \in \mathbb{T}^2.$$

Then, there exist polynomials A_j, B_j in two variables, $j = 1, \ldots, N$, such that we have

(1.5) $$|P\binom{z}{w}|^2 - |Q\binom{z}{w}|^2 = (1-|z|^2) \sum_{j=1}^{N} |A_j\binom{z}{w}|^2 + (1-|w|^2) \sum_{j=1}^{N} |B_j\binom{z}{w}|^2$$

for all $\binom{z}{w} \in \mathbb{C}^2$.

The conditions (1.3) and (1.4) are just another way of saying that Q/P is a rational inner function on the bidisk. Rational inner functions on the bidisk have a quite simple form: Q as above has to equal $\vec{P}$ (where we have to perform the "reflection" at the right degree). In Section 5 we shall present the simple proof that the Geronimo-Woerdeman formula actually proves the above theorem. More importantly, the approach of Geronimo-Woerdeman and of this paper actually gives more information about the decomposition in (1.5) than previous proofs of Andô's theorem. The decomposition (1.5) is usually proven with a finite dimensional Hahn-Banach theorem and as such is not explicit about what A_j and B_j are. The approach here provides very specific

information about A_j and B_j in (1.5) (at least in the case where P is stable). The key to our proof is the study of Bernstein-Szegő measures.

2. Notation and Theorem Rephrasing

This section is devoted to a new notation which we believe will make many of the earlier results more conceptually clear and easier to digest. Like earlier, we shall fix a nondegenerate probability measure ρ on $\mathbb{T}^2$ and corresponding inner product $\langle , \rangle$. Let V be a finite dimensional subspace of two variable polynomials. For each $\binom{Z}{W} \in \mathbb{C}^2$, evaluation at $\binom{Z}{W}$ is a bounded linear functional on V (bounded, of course, because V is finite dimensional), and can therefore be represented as the inner product against the reproducing kernel for V which we denote as $K_\rho V_{\binom{Z}{W}}$. Specifically, if $P \in V$, then $\langle P, K_\rho V_{\binom{Z}{W}} \rangle = P\binom{Z}{W}$. We define $K_\rho V : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ without the subscript $\binom{Z}{W}$ to be

$$K_\rho V \left(\binom{z}{w}, \binom{Z}{W} \right) := K_\rho V_{\binom{Z}{W}} \binom{z}{w} = \langle K_\rho V_{\binom{Z}{W}}, K_\rho V_{\binom{z}{w}} \rangle,$$

which happens to be a holomorphic polynomial in z, w, an anti-holomorphic polynomial in Z, W, and conjugate symmetric in the first and second vectors. Essentially we have defined a map depending on ρ

$K_\rho : \{\text{finite dimensional subspaces of } \square\} \to \{ \text{reproducing kernels}\}$

$$V \mapsto \text{ reproducing kernel for } V$$

which, as we shall see in Section 3, turns orthogonal direct sums into sums, vector space shifts into multiplication shifts, and "reflection" of subspaces into "double reflection" of reproducing kernels.

The following subspaces will be useful throughout. Once the definitions are digested, the use of these strange symbols will become apparent. Essentially we are defining subspaces of polynomials based on the support of their Fourier coefficients (and the symbols represent the Fourier domain).

- Let $\square$ denote the span of $\{z^j w^k : 0 \le j \le n, 0 \le k \le m, (j,k) \ne (n,m)\}$
- Let $\square$ denote the span of $\{z^j w^k : 0 \le j \le n, 0 \le k \le m, (j,k) \ne (0,0)\}$.
- Let $\square$ denote the span of $\{z^j w^k : 0 \le j \le n, 0 \le k \le m-1\}$.
- Let $\square$ denote the span of $\{z^j w^k : 0 \le j \le n-1, 0 \le k \le m\}$.
- Let $\square$ denote the span of $\{z^j w^k : 0 \le j \le n-1, 0 \le k \le m-1\}$.

We hope now the definitions of the subspaces $\square$ and $\square$ will be clear. Next, we define many more subspaces in terms of orthogonal complements using the inner product $\langle , \rangle$.

Define

$$\boxdot := \square \ominus \blacksquare$$

$$\boxdot := \square \ominus \blacksquare$$

$$\boxdot := \square \ominus \blacksquare$$

$$\boxdot := \square \ominus \square$$

and again we hope the definitions of many other such symbols will be intuitively clear (in fact, that is the prime reason for this new notation). When there is more than one measure present, we will add an additional subscript to make it clear which measure we are referring to; e.g. $\boxdot_\rho$ refers to the subspace defined above using the inner product $\langle,\rangle_\rho$ defined by ρ. However, if we are looking at a reproducing kernel $K_\rho\boxdot_\rho$ we will leave off the second ρ and just write $K_\rho\boxdot$.

Since the subspace $\boxdot$ is one dimensional, it is not difficult to see that the corresponding reproducing kernel can be identified with

$$(2.1) \qquad K_\rho\boxdot\left(\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right), \left(\begin{smallmatrix}Z\\W\end{smallmatrix}\right)\right) = p\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)\overline{p\left(\begin{smallmatrix}Z\\W\end{smallmatrix}\right)}$$

for some unit norm polynomial p (which is unique up to multiplication by a unimodular constant).

With all of this notation in place let us restate the Geronimo-Woerdeman Theorem.

Theorem 2.1 (Geronimo-Woerdeman). *Given a nondegenerate probability measure ρ on $\mathbb{T}^2$, any unit norm polynomial p in $\boxdot$ is stable and the measure on $\mathbb{T}^2$*

$$d\mu := \frac{1}{(2\pi i)^2|p\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)|^2}\frac{dz}{z}\frac{dw}{w}$$

has reproducing kernel $K_\mu\square$ equal to $K_\rho\square$ if and only if

$$\boxdot_\rho = \boxdot_\rho.$$

3. Reproducing Kernel Calculus

There are several advantages to working with various reproducing kernels when studying polynomials with respect to a measure on $\mathbb{T}^2$. One advantage is the interface between subspaces and algebraic operations on kernels demonstrated in the following three fundamental propositions. Another advantage is that Bergman identity (below) provides an interface between orthogonal polynomials and reproducing kernels. In this way, choices of ordering polynomials and orthogonal bases are avoided and absorbed into reproducing kernels.

Like before we fix a nondegenerate probability measure ρ on $\mathbb{T}^2$ and corresponding inner product $\langle,\rangle$.

Proposition 3.1. *If V, W are finite dimensional subspaces of polynomials, then V and W are orthogonal if and only if the join of V and W, denoted $(V \vee W)$, satisfies the following formula for reproducing kernels*

$$K_\rho(V \vee W) = K_\rho V + K_\rho W$$

Proof. Suppose V and W are orthogonal. Then, for any $P \in V \oplus W = V \vee W$, which we may write as $P = Q + R$ for $Q \in V$ and $R \in W$, we have

$$\langle P, K_\rho V_{\binom{z}{w}} + K_\rho W_{\binom{z}{w}} \rangle = \langle Q, K_\rho V_{\binom{z}{w}} \rangle + \langle R, K_\rho W_{\binom{z}{w}} \rangle$$
$$= Q\binom{z}{w} + R\binom{z}{w} = P\binom{z}{w}$$

and therefore, $K_\rho(V \oplus W) = K_\rho V + K_\rho W$.

For the other direction, first note that the reproducing kernels $K_\rho V_{\binom{z}{w}}$ span V, since any polynomial in V orthogonal to all reproducing kernels is identically zero. So, if $K_\rho(V \vee W) = K_\rho V + K_\rho W$ as reproducing kernels, then $Q \in V$ implies

$$\langle Q, K_\rho W_{\binom{z}{w}} \rangle = \langle Q, K_\rho(V \vee W)_{\binom{z}{w}} - K_\rho V_{\binom{z}{w}} \rangle = Q\binom{z}{w} - Q\binom{z}{w} = 0$$

i.e. Q is orthogonal to the reproducing kernel for W and hence all of W. □

For example, $K_\rho \square = K_\rho(\square \oplus_\rho \boxdot) = K_\rho \square + K_\rho \boxdot$.

Proposition 3.2. *Let V be a finite dimensional subspace of two variable polynomials and define $(z \cdot V) := \{zP : P \in V\}$ and likewise for $(w \cdot V)$. Then,*[3]

$$K_\rho(z \cdot V)\left(\binom{z}{w}, \binom{Z}{W}\right) = z\bar{Z} K_\rho V\left(\binom{z}{w}, \binom{Z}{W}\right)$$

and likewise for $(w \cdot V)$.

Proof. Observe that for any $Q \in V$

$$\langle zQ, z\bar{Z} K_\rho V_{\binom{Z}{W}} \rangle = \langle Q, \bar{Z} K_\rho V_{\binom{Z}{W}} \rangle = Z Q\binom{Z}{W}$$

since multiplication by z is a unitary. So, $z\bar{Z} K_\rho V$ reproduces a generic element zQ of $(z \cdot V)$. The claim follows by uniqueness of reproducing kernels. □

For example, $K_\rho \square = K_\rho(z \cdot \square) = z\bar{Z} K_\rho \square$ and $K_\rho \boxplus = K_\rho(w \cdot \boxdot) = w\bar{W} K_\rho \boxdot$.

For the third proposition, we recall one definition and introduce another related one.

---

[3]We warn the reader not to confuse the function z and the variable z in the following formula.

Definition 3.3. Given any $P \in \square$, the *reflection* of P (at the (n, m) level) is another polynomial $\vec{P} \in \square$ which is defined by

$$\vec{P}\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right) := z^n w^m \overline{P\left(\begin{smallmatrix} 1/\bar{z} \\ 1/\bar{w} \end{smallmatrix}\right)}$$

Remark 3.4. The map $P \mapsto \vec{P}$ is anti-unitary; i.e.

$$\langle \vec{P}, \vec{Q} \rangle = \langle Q, P \rangle$$

and can be described in concrete terms as the anti-linear map which sends a monomial $z^j w^k \mapsto z^{n-j} w^{m-k}$.

Definition 3.5. Given $\Delta : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ a holomorphic polynomial in the first two variables of degree (n, m) and an anti-holomorphic polynomial in the second two variables of degree (n, m), the *double reflection* of Δ shall be denoted $\overleftrightarrow{\Delta}$ and defined by

$$\overleftrightarrow{\Delta}\left(\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right), \left(\begin{smallmatrix} Z \\ W \end{smallmatrix}\right)\right) := (z\bar{Z})^n (w\bar{W})^m \Delta\left(\left(\begin{smallmatrix} 1/\bar{Z} \\ 1/\bar{W} \end{smallmatrix}\right), \left(\begin{smallmatrix} 1/\bar{z} \\ 1/\bar{w} \end{smallmatrix}\right)\right)$$

Remark 3.6. If Δ is conjugate symmetric, then

$$\overleftrightarrow{\Delta}\left(\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right), \left(\begin{smallmatrix} Z \\ W \end{smallmatrix}\right)\right) = \bar{Z}^n \bar{W}^m \overline{\overrightarrow{\Delta_{\left(\begin{smallmatrix} 1/\bar{Z} \\ 1/\bar{W} \end{smallmatrix}\right)}}}\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right)$$

where $\Delta_{\left(\begin{smallmatrix} 1/\bar{Z} \\ 1/\bar{W} \end{smallmatrix}\right)}\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right) = \Delta\left(\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right), \left(\begin{smallmatrix} 1/\bar{Z} \\ 1/\bar{W} \end{smallmatrix}\right)\right)$.

Proposition 3.7. *Let V be a subspace of $\square$ and define $\vec{V} := \{\vec{P} : P \in V\}$. Then,*

$$K_\rho \vec{V} = \overleftrightarrow{K_\rho V}$$

Proof. Let $P \in V$. Then, $\vec{P} \in \vec{V}$ and by Remarks 3.4 and 3.6

$$\langle \vec{P}, \overleftrightarrow{K_\rho V}_{\left(\begin{smallmatrix} Z \\ W \end{smallmatrix}\right)} \rangle = \langle \vec{P}, \bar{Z}^n \bar{W}^m \overline{\overrightarrow{K_\rho V_{\left(\begin{smallmatrix} 1/\bar{Z} \\ 1/\bar{W} \end{smallmatrix}\right)}}} \rangle = Z^n W^m \langle K_\rho V_{\left(\begin{smallmatrix} 1/\bar{Z} \\ 1/\bar{W} \end{smallmatrix}\right)}, P \rangle = \vec{P}\left(\begin{smallmatrix} Z \\ W \end{smallmatrix}\right)$$

and the claim follows, again by uniqueness. $\qquad\square$

Proposition 3.8 (Bergman Identity). *Let V be a finite dimensional subspace of two variable polynomials and let $q_1, q_2, \ldots, q_N$ be an orthonormal basis for V. Then,*

$$K_\rho V\left(\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right), \left(\begin{smallmatrix} Z \\ W \end{smallmatrix}\right)\right) = \sum_{j=1}^{N} q_j\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right) \overline{q_j\left(\begin{smallmatrix} Z \\ W \end{smallmatrix}\right)}$$

Proof. The expression on the right reproduces each orthonormal basis element and hence everything in V. $\qquad\square$

With the fundamental properties of these reproducing kernels out of the way let us present a formula that holds in complete generality and sheds some light on the Geronimo-Woerdeman theorem.

Since ⊡ is one dimensional, the Bergman identity tells us there is a unit norm polynomial p, unique up to multiplication by a unimodular constant, such that

$$K_\rho\boxdot\left(\binom{z}{w},\binom{Z}{W}\right) = p\binom{z}{w}\overline{p\binom{Z}{W}}.$$

Allow us to abbreviate this as $K_\rho\boxdot = p\bar{p}$. By Proposition 3.7 above $K_\rho\boxdot = \vec{p}\,\vec{\bar{p}}$.

Theorem 3.9. *Let ρ be a nondegenerate probability measure on $\mathbb{T}^2$, and let p be any unit norm polynomial in ⊡. Then, suppressing the argument $\left(\binom{z}{w},\binom{Z}{W}\right)$ we have*

$$\begin{aligned}
p\bar{p} - \vec{p}\,\vec{\bar{p}} =&(1-z\bar{Z})K_\rho\boxminus + (1-w\bar{W})K_\rho\boxminus + (1-z\bar{Z})(1-w\bar{W})K_\rho\square \\
&+ \left[\left(K_\rho\boxminus - K_\rho\boxminus\right) - \left(K_\rho\boxminus - K_\rho\boxminus\right)\right].
\end{aligned}$$

Proof. Throughout the proof we will suppress K_ρ and the argument $\left(\binom{z}{w},\binom{Z}{W}\right)$. Observe that

$$\begin{aligned}
\boxdot - \boxdot &= \left(\square - \blacksquare\right) - \left(\square - \square\right) \\
&= -z\bar{Z}\square + \square \\
&= (1-z\bar{Z})\square
\end{aligned}$$

and similarly

$$(3.1) \qquad\qquad \boxdot - \boxdot = (1-z\bar{Z})\square$$

from which we can write

$$\begin{aligned}
(1-z\bar{Z})\boxdot &= \left(\boxdot - \boxdot\right) - \left(\boxdot - \boxdot\right) \\
&= \left(\boxdot - \boxdot\right) - (1-w\bar{W})\boxdot - \left(\boxdot - \boxdot\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
(3.2) \quad \left(\boxdot - \boxdot\right) - \left(\boxdot - \boxdot\right) &= (1-z\bar{Z})\boxdot &&+(1-w\bar{W})\boxdot \\
&= (1-z\bar{Z})\boxdot &&+(1-w\bar{W})\boxdot \\
& &&+(1-z\bar{Z})(1-w\bar{W})\square
\end{aligned}$$

where the last equality follows from (3.1). Finally, if we note that

$$\boxdot = \boxdot + \boxdot = p\bar{p} + \boxdot$$

and

$$\boxdot = \vec{p}\,\vec{\bar{p}} + \boxdot$$

the theorem follows from (3.2) after a little rearrangement. $\square$

This easily yields the following corollary.

Corollary 3.10. *With the same setup as Theorem 3.9, if* $\square = \square$, *then*

$$|p\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)|^2 - |\vec{p}\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)|^2 \geq (1 - |z|^2)(1 - |w|^2)$$

Proof. When $\square = \square$, it is also true that $\square = \square$ by reflection. Therefore, on the diagonal $z = Z$, $w = W$

$$|p\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)|^2 - |\vec{p}\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)|^2 \geq (1 - |z|^2)(1 - |w|^2)K_\rho \square \geq (1 - |z|^2)(1 - |w|^2)$$

since ρ is a probability measure (and $\square$ can be split into 1 and the polynomials orthogonal to 1). $\square$

This corollary combines with the following lemma about polynomials in two variables to yield one part of the Geronimo-Woerdeman theorem.

Lemma 3.11. *Let* $q \in \square$ *and suppose there is a* $c > 0$ *such that*

$$|q\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)|^2 - |\vec{q}\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right)|^2 \geq c(1 - |z|^2)(1 - |w|^2).$$

Then, q is stable.

Remark 3.12. The significance of this lemma is that it tells us when q has no zeros on the boundary of the bidisk (as it is obvious that the above inequality implies q has no zeros on the *open* bidisk). Also, compare this proof to the "fifth proof of theorem 1.7.1" on page 103 in [9].

Proof. First, suppose q has a zero, say $\left(\begin{smallmatrix}z_0\\w_0\end{smallmatrix}\right)$, on $\mathbb{T} \times \mathbb{D}$. Then, $|q\left(\begin{smallmatrix}rz_0\\w_0\end{smallmatrix}\right)|^2 = O(1 - r)^2$ but

$$|q\left(\begin{smallmatrix}rz_0\\w_0\end{smallmatrix}\right)|^2 \geq c(1 - |w_0|^2)(1 - r^2)$$

which fails as $r \nearrow 1$. Similarly, q has no zeros on $\mathbb{D} \times \mathbb{T}$.

Second, suppose q has a zero, $\left(\begin{smallmatrix}z_0\\w_0\end{smallmatrix}\right)$, on $\mathbb{T}^2$. Writing $q\left(\begin{smallmatrix}rz_0\\rw_0\end{smallmatrix}\right) = a(1 - r) + O(1 - r)^2$ we have

$$\vec{q}\left(\begin{smallmatrix}rz_0\\rw_0\end{smallmatrix}\right) = r^{n+m}z_0^n w_0^m \overline{q\left(\begin{smallmatrix}z_0/r\\w_0/r\end{smallmatrix}\right)} = r^{n+m}z_0^n w_0^m \left(\bar{a}(1 - \frac{1}{r}) + O(1 - r)^2\right).$$

Therefore,

$$
\begin{aligned}
|q\left(\begin{smallmatrix}rz_0\\rw_0\end{smallmatrix}\right)|^2 - |\vec{q}\left(\begin{smallmatrix}rz_0\\rw_0\end{smallmatrix}\right)|^2 =& |a(1 - r) + O(1 - r)^2|^2 \\
& - r^{2(n+m-1)}|a(1 - r) + O(1 - r)^2|^2 \\
=& |a|^2(1 - r)^2(1 - r^{2(n+m-1)}) + O(1 - r)^3 \\
=& O(1 - r)^3 \geq c(1 - r^2)^2
\end{aligned}
$$

which also fails as $r \nearrow 1$. $\square$

Corollary 3.13. *With the setup of Theorem 3.9, if ⊡ = ⊟, then p is stable.*

This proves one part of the reverse implication of Theorem 2.1. For the rest, we refer the reader to the original paper [5].

4. BERNSTEIN-SZEGŐ MEASURES ON $\mathbb{T}^2$

Next, we would like to indicate why the forward implication of the Geronimo-Woerdeman theorem holds. This involves studying Bernstein-Szegő measures:

$$(4.1) \qquad d\mu := \frac{c^2}{(2\pi i)^2 |q\binom{z}{w}|^2} \frac{dz}{z} \frac{dw}{w}$$

where q is a polynomial in $\square$ with no zeros in the closed bidisk and $c > 0$ is chosen to make μ a probability measure.

The first thing to notice about such a measure is that the norm it provides for $L^2(\mathbb{T}^2, \mu)$ is equivalent to the norm on $L^2(\mathbb{T}^2)$ with Lebesgue measure (because q is bounded above and below on $\mathbb{T}^2$). So, all closures taken with respect to Lebesgue measure are equal to closures taken with respect to the norm determined by μ. For instance, $H^2(\mathbb{T}^2, \mu)$, the closure of the polynomials in $L^2(\mathbb{T}^2, \mu)$, is equal to the usual Hardy space $H^2(\mathbb{T}^2)$.

Another important fact about Bernstein-Szegő measures is given in the following proposition.

Proposition 4.1. *For μ and q as in (4.1),*

$$q \in \boxdot \text{ and } K_\mu \boxdot \left(\binom{z}{w}, \binom{Z}{W}\right) = \frac{1}{c^2} q\binom{z}{w} \overline{q\binom{Z}{W}}$$

Proof. Observe that

$$\langle z^j w^k, q \rangle = \frac{c^2}{(2\pi i)^2} \int_{\mathbb{T}^2} \frac{z^j w^k \overline{q\binom{z}{w}}}{|q\binom{z}{w}|^2} \frac{dz}{z} \frac{dw}{w}$$

$$= \frac{c^2}{(2\pi i)^2} \int_{\mathbb{T}^2} \frac{z^j w^k}{q\binom{z}{w}} \frac{dz}{z} \frac{dw}{w}$$

and this equals zero when $j > 0$ or $k > 0$. So, $q \in \boxdot$ and since $||q|| = c$, it follows that $K_\mu \boxdot \left(\binom{z}{w}, \binom{Z}{W}\right) = \frac{1}{c^2} q\binom{z}{w} \overline{q\binom{Z}{W}}$ by the Bergman identity. $\square$

Let us define several important closed subspaces of $H^2(\mathbb{T}^2)$.

Definition 4.2. All closed spans below are taken with respect to $H^2(\mathbb{T}^2)$.

(1) The *border subspace* **B** is defined to be

B $:= \overline{\operatorname{span}}\{z^j w^k : j, k \geq 0 \text{ and either } j \leq n - 1 \text{ or } k \leq m - 1\}.$

(2) The *bottom border subspace* **BB** is defined to be

$\qquad$ **BB** $:= \overline{\operatorname{span}}\{z^j w^k : j \geq 0 \text{ and } 0 \leq k \leq m - 1\}.$

(3) The *left border subspace* **LB** is defined to be

$\qquad$ **LB** $:= \overline{\operatorname{span}}\{z^j w^k : k \geq 0 \text{ and } 0 \leq j \leq n - 1\}.$

Lemma 4.3. *Define*

$$L_{\binom{Z}{W}}\binom{z}{w} := \frac{(z\bar{Z})^n [q\binom{z}{w}\overline{q\binom{1/\bar{z}}{W}} - \vec{q}\binom{z}{w}\overline{\vec{q}\binom{1/\bar{z}}{W}}]}{(1 - z\bar{Z})(1 - w\bar{W})}$$

The functions $L_{\binom{Z}{W}}$ *for* $Z, W \in \mathbb{D}$ *have the following properties.*

(1) $L_{\binom{Z}{W}}$ *is orthogonal (with respect to* μ*) to* **LB**.

(2) $L_{\binom{Z}{W}}$ *is an element of* **BB**.

(3) $L_{\binom{Z}{W}}$ *reproduces* $(z^n \cdot$ **BB**$)$*; i.e.* $f\binom{z}{w} = \langle f, L_{\binom{z}{w}}\rangle_\mu$ *for* f *in the closed span of* $\{z^j w^k : j \geq n, 0 \leq k \leq m - 1\}.$

Remark 4.4. In essence, the proposition says that $L_{\binom{z}{w}}$ reproduces the $H^2(\mathbb{T}^2)$ projection of $f \in$ **B** to $(z^n \cdot$ **BB**$)$.

Before proving this let us show how it proves the following important fact for Bernstein-Szegő measures.

Theorem 4.5. *Let* q *be a polynomial with no zeros on* $\overline{\mathbb{D}}^2$ *and define its Bernstein-Szegő measure* μ *as in* (4.1). *Then,*

$$\mathbf{B} \ominus_\mu \mathbf{BB} = \mathbf{LB} \ominus_\mu \boxminus,$$
$$\mathbf{B} \ominus_\mu \mathbf{LB} = \mathbf{BB} \ominus_\mu \boxminus$$

and

$$\boxplus_\mu = \boxminus_\mu \text{ and } \boxplus_\mu = \boxminus_\mu$$

Proof. If $f \in \mathbf{B}\ominus_\mu\mathbf{BB}$ then by Remark 4.4, $f \in \mathbf{LB}$ because $f \perp_\mu L_{\binom{z}{w}}$ for all $\binom{z}{w} \in \mathbb{D}^2$. So, $\mathbf{B} \ominus_\mu \mathbf{BB} \subseteq \mathbf{LB} \ominus_\mu \boxminus$. On the other hand, if $f \in \mathbf{LB} \ominus_\mu \boxminus$ and if $f \perp_\mu \mathbf{B} \ominus_\mu \mathbf{BB}$, then $f \in \boxminus$ yet at the same time $f \perp_\mu \boxminus$; i.e. $f \equiv 0$. Hence, $\mathbf{B}\ominus_\mu\mathbf{BB} = \mathbf{LB}\ominus_\mu\boxminus$. It follows easily from this that $\boxplus_\mu = \boxminus_\mu$. The rest of the theorem follows by symmetry. $\square$

Now, we prove Lemma 4.3 via a series of propositions, which are interesting in their own right.

Proposition 4.6. *Define* $F_{\binom{Z}{W}} \in H^2(\mathbb{T}^2)$ *by*

$$F_{\binom{Z}{W}}\binom{z}{w} = \frac{(z\bar{Z})^n q\binom{z}{w}\overline{q\binom{1/\bar{z}}{W}}}{c^2(1 - z\bar{Z})(1 - w\bar{W})}$$

Then, for any $f \in H^2(\mathbb{T}^2)$, *the function*

$$g\binom{z}{w} := \overline{\langle f, F_{\binom{z}{w}}\rangle_\mu}$$

is the $H^2(\mathbb{T}^2)$-*orthogonal projection of* f *to the subspace* $(z^n \cdot H^2(\mathbb{T}^2))$. *(So, a projection in* H^2 *with Lebesgue measure is achieved via the* μ-*inner product.)*

Proof. Observe that

$$g\binom{Z}{W} = \frac{1}{(2\pi i)^2}\int_{\mathbb{T}^2} f\binom{z}{w}\frac{(\bar{z}Z)^n\overline{q\binom{z}{w}}q\binom{z}{W}}{(1 - \bar{z}Z)(1 - \bar{w}W)|q\binom{z}{w}|^2}\frac{dw}{w}\frac{dz}{z}$$

$$= \frac{1}{2\pi i}\int_{\mathbb{T}} f\binom{z}{W}\frac{(\bar{z}Z)^n}{1 - \bar{z}Z}\frac{dz}{z}$$

and the last line does indeed equal the projection of f to the subspace $(z^n \cdot H^2(\mathbb{T}^2))$. $\square$

Proposition 4.7. *The closed subspace* $(\vec{q} \cdot H^2(\mathbb{T}^2))$ *equals* $H^2(\mathbb{T}^2) \ominus_\mu \mathbf{B}$.

Proof. If $f \in H^2(\mathbb{T}^2)$ then

$$\langle f\vec{q}, z^j w^k\rangle_\mu = \langle f z^{n-j}w^{m-j}, q\rangle_\mu$$

$$= \frac{c}{(2\pi i)^2}\int_{\mathbb{T}^2} f\binom{z}{w}\frac{z^{n-j}w^{m-k}}{q\binom{z}{w}}\frac{dz}{z}\frac{dw}{w}$$

equals zero if $j < n$ or $k < m$. So, $(\vec{q} \cdot H^2(\mathbb{T}^2)) \subseteq H^2(\mathbb{T}^2) \ominus_\mu \mathbf{B}$. On the other hand, if $f \in H^2(\mathbb{T}^2) \ominus_\mu \mathbf{B}$ and $f \perp_\mu (\vec{q} \cdot H^2(\mathbb{T}^2))$, then since $\vec{q} \in \square$ and the coefficient of $z^n w^m$ in $\vec{q}$ is necessarily nonzero, it follows that $f \perp_\mu z^n w^m$. But, since $f \perp_\mu w\vec{q}$, this implies $f \perp_\mu z^n w^{m+1}$. Continuing like this we see that $f \perp_\mu z^j w^k$ for all $j, k \geq 0$; i.e. $f \equiv 0$. So, $(\vec{q} \cdot H^2(\mathbb{T}^2)) = H^2(\mathbb{T}^2) \ominus_\mu \mathbf{B}$. $\square$

Proposition 4.8. *The function*

$$G\left(\binom{z}{w}, \binom{Z}{W}\right) := \frac{q\binom{z}{w}\overline{q\binom{Z}{W}} - (z\bar{Z})^n \vec{q}\binom{1/\bar{Z}}{w}\overline{\vec{q}\binom{1/\bar{z}}{W}}}{1 - w\bar{W}}$$

is a polynomial of degree $m - 1$ *in* w *and* $\bar{W}$.

Proof. The proposition follows for algebraic reasons, but we can actually give meaning to G so we provide the following more explicit proof.

First, notice that if the claim is true when $z = Z$ then it is true in general. This is because of the polarization theorem for holomorphic functions (i.e. $f(z, \bar{z}) \equiv 0$ implies $f \equiv 0$) applied to coefficients of $w^j \bar{W}^k$ in the expansion of G above.

To prove the result for $z = Z$ we can appeal to the one variable Christoffel-Darboux formula. Namely, for each $z \in \overline{\mathbb{D}}$ let

$$d\sigma_z(w) := \frac{c_z^2}{2\pi i |q\binom{z}{w}|^2} \frac{dw}{w}$$

where $c_z > 0$ is chosen to make σ_z a probability measure on $\mathbb{T}$. Then, it is not difficult to check the polynomial in w given by $q\binom{z}{w}/c_z$ satisfies the conditions of the Christoffel-Darboux theorem (up to multiplication by a unimodular which does not matter). Therefore, if we denote the reproducing kernel for σ_z for polynomials in w with degree up to $m-1$ by K_{m-1}^z, then

$$c_z^2(1 - w\bar{W})K_{m-1}^z(w, W) = q\binom{z}{w}\overline{q\binom{z}{W}} - (w\bar{W})^m\overline{q\binom{z}{1/\bar{w}}}q\binom{z}{1/\bar{W}}$$

$$= q\binom{z}{w}\overline{q\binom{z}{W}} - (z\bar{Z})^n\vec{q}\binom{1/\bar{z}}{w}\overline{\vec{q}\binom{1/\bar{z}}{W}}$$

and this proves G is a polynomial in w and $\bar{W}$ of degree $m-1$ for $z = Z \in \overline{\mathbb{D}}$. □

These three propositions allow us to prove Lemma 4.3 with ease.

Proof of Lemma 4.3.   (1) The functions $L_{\binom{z}{w}}$ equal a sum of a functions orthogonal to **LB** by Propositions 4.6 and 4.7.

(2) The function $L_{\binom{Z}{W}}$ belongs to **BB** by Proposition 4.8 because

$$L_{\binom{Z}{W}}\binom{z}{w} = \frac{(z\bar{Z})^n}{(1 - z\bar{Z})}G\left(\binom{z}{w}, \binom{1/\bar{z}}{W}\right)$$

is holomorphic in $z, \bar{Z}, w, \bar{W}$ and a polynomial of degree $m-1$ in $w, \bar{W}$.

(3) Finally, $L_{\binom{z}{w}}$ is the sum of a function that reproduces $(z^n \cdot H^2(\mathbb{T}^2))$ and a function orthogonal to $(z^n \cdot \textbf{BB})$. So, $L_{\binom{z}{w}}$ reproduces functions in $(z^n \cdot \textbf{BB})$.

□

Finally, Theorem 4.5 allows us to prove the forward implication of Theorem 2.1, which we state as a corollary.

Corollary 4.9. *Given a nondegenerate probability measure ρ on $\mathbb{T}^2$, if a unit norm polynomial $p \in \boxdot$ is stable and the measure on $\mathbb{T}^2$*

$$d\mu := \frac{1}{(2\pi i)^2 |p\binom{z}{w}|^2} \frac{dz}{z} \frac{dw}{w}$$

has reproducing kernel $K_\mu \square$ equal to $K_\rho \square$, then

$$\boxdot_\rho = \boxdot_\rho.$$

Proof. If $K_\rho \square$ equals $K_\mu \square$, then Theorem 4.5 proves

$$\boxdot_\rho = \boxdot_\mu = \boxdot_\mu = \boxdot_\rho.$$

$\square$

5. Andô, Agler, and Christoffel-Darboux Redux

Let us present and prove the Geronimo-Woerdeman Christoffel-Darboux formula (plus another identity) in our own language.

Theorem 5.1. *If ρ is a nondegenerate probability measure on $\mathbb{T}^2$ and $\boxdot = \boxdot$, then writing, as usual, $p\bar{p} = K_\rho \boxdot$ and suppressing $\left(\binom{z}{w}, \binom{Z}{W}\right)$*

$$
\begin{aligned}
p\bar{p} - \vec{p}\bar{\vec{p}} &= (1 - z\bar{Z})K_\rho\boxdot + (1 - w\bar{W})K_\rho\boxdot + (1 - z\bar{Z})(1 - w\bar{W})K_\rho\square \\
&= (1 - z\bar{Z})K_\rho\boxdot + (1 - w\bar{W})K_\rho\boxdot \\
&= (1 - z\bar{Z})K_\rho\boxdot + (1 - w\bar{W})K_\rho\boxdot.
\end{aligned}
$$

Proof. These all follow from Theorem 3.9 using the fact that $\boxdot = \boxdot$ implies $\boxdot = \boxdot$ (by reflection) and the identities

$$K_\rho\boxdot - K_\rho\boxdot = (1 - z\bar{Z})K_\rho\square$$
$$K_\rho\boxdot - K_\rho\boxdot = (1 - w\bar{W})K_\rho\square.$$

$\square$

Let us now prove the Cole-Wermer equivalent of Andô's theorem.

Proof of Theorem 1.7. Let us recall the setup. We have two-variable polynomials P and Q satisfying

$$|P| = |Q| \text{ on } \mathbb{T}^2$$

and

$$|P| \geq |Q| \text{ on } \mathbb{D}^2$$

There is no loss in assuming P and Q are relatively prime (just multiply the formula we prove by any common factors if necessary). Then, Q/P is holomorphic on $\mathbb{D}^2$ because it is bounded by 1 wherever it is defined. In fact, P has no zeros on $\mathbb{D}^2$ because otherwise any zeros of P in $\mathbb{D}^2$ necessarily coincide with zeros of Q, in which case P and Q would

have a common factor. This implies Q/P is a rational inner function. It is a fact that $Q = \vec{P}$ where we have to perform the "reflection" at the appropriate level, which we assume to be (n, m) (see [8] Theorem 5.2.5).

If it were the case that P had no zeros on $\overline{\mathbb{D}}^2$, then the theorem would be proved by defining a Bernstein-Szegő measure using P and applying the above Christoffel-Darboux formula. To handle the case where P might have zeros on $\mathbb{T}^2$ we look at $P_r\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right) := P\left(\begin{smallmatrix} rz \\ rw \end{smallmatrix}\right)$ and let $r \nearrow 1$.

Define μ_r to be the Bernstein-Szegő measure for P_r, i.e.

$$d\mu_r := \frac{c_r^2}{(2\pi i)^2 |P_r\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right)|^2} \frac{dz}{z} \frac{dw}{w}$$

where c_r is chosen to make μ_r a probability measure. Then, by Proposition 4.1 and the Christoffel-Darboux formula on the diagonal $z = Z$, $w = W$

$$|P_r|^2 - |\vec{P}_r|^2 = c_r^2(1 - |z|^2)K_{\mu_r}\boxminus + c_r^2(1 - |w|^2)K_{\mu_r}\boxdot.$$

Also, by the Bergman identity (evaluating $K_{\mu_r}\boxminus$ at $\left(\left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right), \left(\begin{smallmatrix} z \\ w \end{smallmatrix}\right)\right)$)

$$\frac{c_r^2}{(2\pi i)^2} \int_{\mathbb{T}^2} K_{\mu_r}\boxminus \frac{dz}{z} \frac{dw}{w} \leq \left(\sup_{\mathbb{T}^2} |P_r|^2\right) \frac{c_r^2}{(2\pi i)^2} \int_{\mathbb{T}^2} \frac{K_{\mu_r}\boxminus}{|P_r|^2} \frac{dz}{z} \frac{dw}{w}$$
$$\leq n \sup_{\mathbb{T}^2} |P|^2$$

since the dimension of $\boxminus$ is n. Similarly,

$$\frac{c_r^2}{(2\pi i)^2} \int_{\mathbb{T}^2} K_{\mu_r}\boxdot \frac{dz}{z} \frac{dw}{w} \leq m \sup_{\mathbb{T}^2} |P|^2.$$

Hence, the polynomials forming $c_r^2 K_{\mu_r}\boxminus$ as in the Bergman identity are bounded in L^2 norm and are all of bounded degree (likewise for $c_r^2 K_{\mu_r}\boxdot$). Taking subsequences if necessary, said polynomials will converge to polynomials as $r \nearrow 1$. Also, $\vec{P}_r$ tends to $\vec{P}$, and that proves the theorem: i.e.

$$|P|^2 - |\vec{P}|^2 = (1 - |z|^2) \sum_{j=0}^{n-1} |A_j|^2 + (1 - |w|^2) \sum_{k=0}^{m-1} |B_k|^2$$

for some polynomials A_j of degree $(n - 1, m)$ and B_k of degree $(n, m - 1)$. $\square$

6. Equivalences

For those interested in referring back to the original paper of Geronimo and Woerdeman [5] we have included this section devoted to demonstrating equivalences between their theorem statements and ours. As usual, we let ρ be a probability measure on the 2-torus.

Let $\Lambda_+ = \{0, \ldots, n\} \times \{0, \ldots, m\}$ and define the doubly Toeplitz matrix corresponding to ρ by

$$c_{u-v}^{\rho} = \langle z^{u_1} w^{u_2}, z^{v_1} w^{v_2} \rangle_{\rho}$$

where $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are elements of Λ_+.

First, let us explicitly mention that our nondegeneracy condition is the same as the one in [5].

Proposition 6.1. *The following are equivalent.*

(1) *ρ is nondegenerate (at the (n, m) level).*
(2) *The doubly Toeplitz matrix corresponding to ρ is positive definite:*
$$(c_{u-v}^{\rho})_{u,v \in \Lambda_+} > 0.$$

Proof. Both conditions are just another way of saying that for any complex numbers a_{jk}, $j = 0, \ldots, n$, $k = 0, \ldots, m$

$$\langle \sum_{j=0}^{n} \sum_{k=0}^{m} a_{jk} z^j w^k, \sum_{j=0}^{n} \sum_{k=0}^{m} a_{jk} z^j w^k \rangle_{\rho} > 0$$

as long as some a_{jk} is nonzero. $\qquad\square$

Next, our "automatic orthogonality condition" is equivalent to a couple of different conditions on Toeplitz matrices that are used in [5]. The automatic orthogonality condition is also used in [7].

Proposition 6.2. *The following are equivalent.*

(1) *Automatic orthogonality:*
$$\boxed{\cdot}_{\rho} = \boxed{\cdot}_{\rho}$$

(2) *A low rank condition on the doubly Toeplitz matrix:*

$$rank(c_{u-v}^{\rho})_{\substack{u \in \{1, \ldots, n\} \times \{0, \ldots, m\} \\ v \in \{0, \ldots, n\} \times \{1, \ldots, m\}}} = nm$$

(3) *The vanishing of certain blocks of the inverse of the doubly Toeplitz matrix:*

$$\left[(c_{u-v}^{\rho})_{u,v \in \Lambda_+ \setminus \{(0,0)\}} \right]^{-1}_{\substack{\{1, \ldots, n\} \times \{0\} \\ \{0\} \times \{1, \ldots, m\}}} = 0$$

Proof. The equivalence of the second and third statements follows from Theorem 2.4.1 in [5]. We shall prove (1) is equivalent to (3). Let $b_{u,v}$ be the (u,v) entry of the inverse of the matrix

$$(c^\rho_{u-v})_{u,v\in\Lambda_+\setminus\{(0,0)\}}.$$

Warning: the rows and columns of $(b_{u,v})$ and (c^ρ_{u-v}) are each indexed by the set $\Lambda_+\setminus\{(0,0)\}$. For each $u\in\Lambda_+\setminus\{(0,0)\}$ define

$$\Phi_u\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right) := \sum_{v\in\Lambda_+\setminus\{(0,0)\}} b_{u,v}z^{v_1}w^{v_2}$$

where we write $v=(v_1,v_2)$. The set $\{\Phi_u\}_{u\in\Lambda_+\setminus\{(0,0)\}}$ is a dual basis for the monomials in $\square$, because for $t=(t_1,t_2)\in\Lambda_+\setminus\{(0,0)\}$

$$\langle\Phi_u, z^{t_1}w^{t_2}\rangle_\rho = \sum_{v\in\Lambda_+\setminus\{(0,0)\}} b_{u,v}c^\rho_{v-t} = \delta_{u,t}.$$

Therefore, the set $\{\Phi_{(j,0)}\}_{j=1,\dots,n}$ forms a basis for the subspace $\boxdot$. The condition $b_{(j,0),(0,k)}=0$ for $j=1,\dots,n$ and $k=1,\dots,m$ is equivalent to saying each $\Phi_{(j,0)}\in\boxplus$ and since these polynomials are linearly independent, this is equivalent to saying $\boxdot=\boxplus$. By reflection, this is equivalent to $\boxbar=\boxtimes$. It is not difficult to see that in general we have

$$\boxbar\oplus_\rho\boxslash = \boxslash\oplus_\rho\boxtimes$$

and so $\boxbar=\boxtimes$ if and only if $\boxslash=\boxslash$ by taking orthogonal complements. Hence, conditions (1) and (3) are equivalent.

We conclude by remarking that the above proof shows that the conditions obtained by switching the roles of z and w are equivalent to the conditions (1)-(3). $\qquad\square$

The following is probably too easy to be a proposition, but we state it for emphasis.

Proposition 6.3. *Let ρ and σ be two probability measures on $\mathbb{T}^2$. The following are equivalent.*

(1) *The Toeplitz matrices agree:*

$$c^\rho_{u-v} = c^\sigma_{u-v} \text{ for } u,v\in\Lambda_+.$$

(2) *The inner products agree on $\square$:*

$$\langle P,Q\rangle_\rho = \langle P,Q\rangle_\sigma \text{ for } P,Q\in\square$$

(3) *The reproducing kernels agree:*

$$K_\rho\square\left(\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right),\left(\begin{smallmatrix}Z\\W\end{smallmatrix}\right)\right) = K_\sigma\square\left(\left(\begin{smallmatrix}z\\w\end{smallmatrix}\right),\left(\begin{smallmatrix}Z\\W\end{smallmatrix}\right)\right).$$

Proof. The first two equivalences are easy. It is also clear that (2) implies (3). If the reproducing kernels agree then the inner products agree on linear combinations of kernel functions. Since the kernel functions span □, it follows that the inner products must agree on all of □.   □

With all of this out of the way, let us state the important theorem (namely Theorem 2.1) in this paper in its original language. As mentioned earlier, this theorem is a slight weakening of Theorem 1.1.2 in [5].

Theorem 6.4 (Geronimo-Woerdeman)**.** *Given is a probability measure ρ on $\mathbb{T}^2$ with*

$$(c^{\rho}_{u-v})_{u,v\in\Lambda_+} > 0$$

A unit norm polynomial p in $\boxdot$ is stable and the measure on $\mathbb{T}^2$

$$d\mu := \frac{1}{(2\pi i)^2 |p\binom{z}{w}|^2} \frac{dz}{z} \frac{dw}{w}$$

has Fourier coefficients $c^{\mu}_u = c^{\rho}_u$ for $u \in \{-n,\dots,n\} \times \{-m,\dots,m\}$ if and only if

$$rank(c^{\rho}_{u-v})_{\substack{u \in \{1,\dots,n\} \times \{0,\dots,m\} \\ v \in \{0,\dots,n\} \times \{1,\dots,m\}}} = nm.$$

7. Conclusions and Commentary

7.1. **Three Variables?** There are several proofs of Andô's inequality (see [2] and [3] for instance). The work of Geronimo and Woerdeman and this paper provide two more. Such a simple looking inequality has garnered our interest partially because of its connection to bounded analytic functions and interpolation. It is also interesting because the analogous statement for three commuting contractions fails. This phenomenon is not well understood. We hope that the notation and approach of this paper can play some small part in addressing this problem. Admittedly, our notation has certain drawbacks for generalizations to three variables, but we do not think they are insurmountable.

7.2. **Formulas for reproducing kernels.** It would be interesting to have explicit formulas for the reproducing kernels involved in the two variable Christoffel-Darboux formula in the case of Bernstein-Szegő measures. This might allow for a slicker proof of the Cole-Wermer theorem (i.e. passing to subsequences might be avoided). It would also be interesting to study the probability measures μ_r in our proof of the Cole-Wermer theorem.

7.3. **Proof of the stability result.** Let us conclude the paper with a proof of the stability result from the introduction.

Proof of Theorem 1.1. The reverse implication is given by Lemma 3.11. To prove the forward implication, define a Bernstein-Szegő measure using q. The constant c is defined by

$$\frac{1}{c^2} = \frac{1}{(2\pi i)^2} \int_{\mathbb{T}^2} \frac{1}{|q\binom{z}{w}|^2} \frac{dz}{z} \frac{dw}{w}.$$

By Theorem 4.5, Corollary 3.10, and Proposition 4.1, it follows that

$$|q\binom{z}{w}|^2 - |\vec{q}\binom{z}{w}|^2 \geq c^2(1 - |z|^2)(1 - |w|^2).$$

$\square$

Symbols and acknowledgments

References

[1] Jim Agler and John E. McCarthy. *Pick interpolation and Hilbert function spaces*, volume 44 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[2] T. Andô. On a pair of commutative contractions. *Acta Sci. Math. (Szeged)*, 24:88–90, 1963.

[3] Joseph A. Ball, Cora Sadosky, and Victor Vinnikov. Scattering systems with several evolutions and multidimensional input/state/output systems. *Integral Equations Operator Theory*, 52(3):323–393, 2005.

[4] Brian J. Cole and John Wermer. Ando's theorem and sums of squares. *Indiana Univ. Math. J.*, 48(3):767–791, 1999.

[5] Jeffrey S. Geronimo and Hugo J. Woerdeman. Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables. *Ann. of Math. (2)*, 160(3):839–906, 2004.

[6] Jeffrey S. Geronimo and Hugo J. Woerdeman. Two Variable Orthogonal Polynomials on the Bi-circle and Structured Matrices. *preprint*, 2006.

[7] Jeffrey S. Geronimo and Hugo J. Woerdeman. Two-variable polynomials: intersecting zeros and stability. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 53(5):1130–1139, 2006.

[8] Walter Rudin. *Function theory in polydiscs*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[9] Barry Simon. *Orthogonal polynomials on the unit circle. Part 1*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005. Classical theory.

Department of Mathematics, Washington University in St. Louis, St. Louis, MO, 63130

E-mail address: geknese@wustl.edu

URL: http://www.math.wustl.edu/~geknese