Acta Math., 194 (2005), 133–153

Distinguished varieties

by

JIM AGLER and JOHN E. McCARTHY

University of California, San Diego
La Jolla, CA, U.S.A.

Washington University
St. Louis, MO, U.S.A.

0. Introduction

In this paper, we shall be looking at a special class of bordered (algebraic) varieties that are contained in the bidisk $\mathbf{D}^2$ in $\mathbf{C}^2$.

Definition 0.1. A non-empty set V in $\mathbf{C}^2$ is a *distinguished variety* if there is a polynomial p in $\mathbf{C}[z,w]$ such that

$$V=\{(z,w)\in\mathbf{D}^2:p(z,w)=0\}$$

and such that

$$\overline{V}\cap\partial(\mathbf{D}^2)=\overline{V}\cap(\partial\mathbf{D})^2. \tag{0.2}$$

Condition (0.2) means that the variety exits the bidisk through the distinguished boundary of the bidisk, the torus. We shall use ∂V to denote the set given by (0.2): topologically, it is the boundary of V within Z_p, the zero set of p, rather than in all of $\mathbf{C}^2$. We shall always assume that p is chosen to be minimal, i.e. so that no irreducible component of Z_p is disjoint from $\mathbf{D}^2$ and so that p has no repeated irreducible factors. Why should one single out distinguished varieties from other bordered varieties?

One of the most important results in operator theory is T. Andô's inequality [7] (see also [12] and [24]). This says that if T_1 and T_2 are commuting operators, and both of them are of norm 1 or less, then for any polynomial p in two variables, the inequality

$$\|p(T_1,T_2)\|\leqslant\|p\|_{\mathbf{D}^2} \tag{0.3}$$

holds. Andô's inequality is essentially equivalent to the commutant lifting theorem of B. Sz.-Nagy and C. Foiaş [23]—see, e.g., [20] for a discussion of this.

The first author was partially supported by the National Science Foundation. The second author was partially supported by National Science Foundation Grant DMS 0070639.

Our first main result, Theorem 3.1, is that if T_1 and T_2 are matrices, then the inequality (0.3) can be improved to

$$\|p(T_1, T_2)\| \leqslant \|p\|_V,$$

where V is some distinguished variety depending on T_1 and T_2. Indeed, in the proof of the theorem, we construct co-isometric extensions of the matrices that naturally live on this distinguished variety. So when studying bivariable matrix theory, rather than operator theory, one is led inexorably to study distinguished varieties.

Conversely, in Theorem 1.12, we show that all distinguished varieties can be represented as

$$\{(z, w) \in \mathbf{D}^2 : \det(\Psi(z) - wI) = 0\}$$

for some analytic matrix-valued function Ψ on the disk that is unitary on $\partial\mathbf{D}$. This shows that the study of distinguished varieties leads back to operator theory. Consider the natural notion of isomorphism of two distinguished varieties, namely that there is a biholomorphic bijection between them.

Definition 0.4. A function Φ is *holomorphic* on a set V in $\mathbf{C}^2$ if, at every point λ in V, there is a non-empty ball $B(\lambda, \varepsilon)$ centered at λ and an analytic mapping of two variables defined on $B(\lambda, \varepsilon)$ that agrees with Φ on $B(\lambda, \varepsilon) \cap V$.

Definition 0.5. Two distinguished varieties V_1 and V_2 are *isomorphic* if there is a function Φ that is holomorphic on V_1 and continuous on $\overline{V}_1$ such that Φ is a bijection from $\overline{V}_1$ onto $\overline{V}_2$ and such that Φ^{-1} is holomorphic on V_2.

(The requirement that Φ^{-1} be holomorphic does not follow automatically from the holomorphicity of Φ—consider, e.g., $V_1 = \{(z, z) : z \in \mathbf{D}\}$ and $V_2 = \{(z^2, z^3) : z \in \mathbf{D}\}$, which are not isomorphic.)

By the maximum modulus principle, Φ must map the boundary of V_1 onto the boundary of V_2. It follows that $\Phi = (\phi_1, \phi_2)$ is a pair of *inner functions*, i.e. a pair of holomorphic scalar-valued functions that each have modulus one on ∂V_1. So studying isomorphism classes of distinguished varieties is closely connected to the rich structure of inner functions.

W. Rudin has studied when an arbitrary finite Riemann surface R is isomorphic to a distinguished variety, in the sense that there is an unramified pair of separating inner functions on R that are continuous on $\overline{R}$ [22]. His results show, for example, that a finitely connected planar domain is isomorphic to a distinguished variety if and only if the domain is either a disk or an annulus. He also showed that for every $n \geqslant 1$, there is a finite Riemann surface R that is topologically an n-holed torus minus one disk, and such that R is isomorphic to a distinguished variety.

In §2 we show that, under fairly general conditions, a pair of "inner" functions (ϕ_1, ϕ_2) on a set X must map X into a distinguished variety (i.e. the algebraic relation on the ϕ_i's comes for free).

Another reason to study distinguished varieties comes from considering the Pick problem on the bidisk. This is the problem of deciding, given points $\lambda_1, ..., \lambda_N$ in $\mathbf{D}^2$, and values $w_1, ..., w_N$ in $\mathbf{C}$, whether there is a function in $H^\infty(\mathbf{D}^2)$, the bounded analytic functions on $\mathbf{D}^2$, that interpolates the data and is of norm at most 1. The problem is called *extremal* if there is an interpolating function of norm exactly 1, but not less.

If an extremal Pick problem is given, the solution may or may not be unique (see §4 for an example). Our second main result is Theorem 4.1, where we show that there is always a distinguished variety on which the solution is unique.

One can think then of the Pick problem as having two parts:

(a) Solve the problem on the distinguished variety where the solution is unique.

(b) Parametrize all the extensions of the solution to the whole bidisk.

We give a formula (4.10) for problem (a). The extension problem (b) is non-trivial: unless the distinguished variety is isomorphic to a disk, there will always be some functions that cannot be extended to the whole bidisk without increasing the norm [5]. Obviously a function arising from a Pick problem will be extendable, but what distinguishes such functions remains mysterious.

If one starts with an inner function on V and wants to extend this to a rational inner function on $\mathbf{D}^2$, there can be more than one extension. However, there is a restriction on the degree, given by Theorem 2.8. If the variety is of rank (n_1, n_2), i.e. there are generically n_1 sheets above every first coordinate and n_2 above every second coordinate, then any regular rational inner extension of degree (d_1, d_2) must have $d_1 n_1 + d_2 n_2$ equal to the number of zeroes that the original function had on V.

An *admissible kernel* K on a set $\{\lambda_1, ..., \lambda_N\}$ in $\mathbf{D}^2$ is a positive definite $N \times N$-matrix such that

$$[(1-\lambda_i^r \bar{\lambda}_j^r) K_{ij}] \geqslant 0, \quad r = 1, 2.$$

It is known [1], [4] that studying all the admissible kernels on a set is essential to understanding the Pick problem. A key idea in the proof of Theorem 4.1 is that every admissible kernel automatically extends to a distinguished variety.

Distinguished varieties have been studied in a somewhat more abstract and general setting by J. Ball and V. Vinnikov [9]. They have a determinantal representation that is analogous to Theorem 1.12.

We would like to thank the referees for many valuable remarks.

1. Representing distinguished varieties

Let V be a distinguished variety. We say that a function f is *holomorphic on* V if, for every point of V, there is an open ball B in $\mathbf{C}^2$ containing the point, and a holomorphic function ϕ of two variables on B, such that $\phi|_{B\cap V} = f|_{B\cap V}$. We shall use $A(V)$ to denote the Banach algebra of functions that are holomorphic on V and continuous on $\overline{V}$. This is a uniform algebra on ∂V, i.e. a closed unital subalgebra of $C(\partial V)$ that separates points. The maximal ideal space of $A(V)$ is $\overline{V}$.

If μ is a finite measure on a distinguished variety V, let $H^2(\mu)$ denote the closure in $L^2(\mu)$ of the polynomials. If Ω is an open subset of a Riemann surface S, and ν is a finite measure on $\overline{\Omega}$, let $\mathcal{A}^2(\nu)$ denote the closure in $L^2(\nu)$ of $A(\Omega)$, the functions that are holomorphic on Ω and continuous on $\overline{\Omega}$. We say that a point λ is a bounded point evaluation for $H^2(\mu)$ (or $\mathcal{A}^2(\nu)$) if evaluation at λ, *a priori* defined only for a dense set of analytic functions, extends continuously to the whole Hilbert space. If λ is a bounded point evaluation, we call the function k_λ that has the property

$$\langle f, k_\lambda \rangle = f(\lambda)$$

the *evaluation functional at* λ.

The following lemma is well known. It is valid in much greater generality, but this will suffice for our purposes. If the boundary of Ω consists of closed analytic curves, the lemma follows from J. Wermer's proof [25] that $A(\Omega)$ is hypo-Dirichlet, and the description of representing measures for hypo-Dirichlet algebras given by P. Ahern and D. Sarason in [6]. (Actually Wermer's proof extends without difficulty to the case where the boundary is just piecewise C^2, but we shall not need this fact). For a detailed description of the measures in this case, see K. Clancey's paper [10].

LEMMA 1.1. *Let S be a compact Riemann surface. Let $\Omega \subseteq S$ be a domain whose boundary is a finite union of piecewise smooth Jordan curves. Then there exists a measure ν on $\partial\Omega$ such that every λ in Ω is a bounded point evaluation for $\mathcal{A}^2(\nu)$, and such that the linear span of the evaluation functionals is dense in $\mathcal{A}^2(\nu)$.*

Proof. Because its boundary is nice, Ω is regular for the Dirichlet problem (see, e.g., [14, §IV.2]). Let ν be harmonic measure for Ω with respect to some fixed base-point. Then by Harnack's inequality, harmonic measure for any other point in the domain is boundedly absolutely continuous with respect to ν. As harmonic evaluation functionals are *a fortiori* analytic evaluation functionals, we get that every point of Ω is a bounded point evaluation (with an L^∞ evaluation functional) for $\mathcal{A}^2(\nu)$.

Ahern and Sarason [6, p. 159] proved that the span of the evaluation functionals is dense. Their argument, in brief, was to find an exhaustion Ω_n of Ω, i.e. an increasing family of open sets, each contained compactly in the next, whose union was Ω. Let ν_n be harmonic measure for each Ω_n, with respect to the same fixed base-point. Then they showed that for every u in $L^1(\partial\Omega, \nu)$, its norm was equal to

$$\lim_{n\to\infty} \int_{\partial\Omega_n} |\hat{u}|\, d\nu_n,$$

where $\hat{u}$ is the harmonic extension of u to Ω. In particular, any function in $\mathcal{A}^2(\nu)$ that vanishes identically on Ω must be the zero function. $\square$

LEMMA 1.2. *Let V be a distinguished variety. There is a measure μ on ∂V such that every point in V is a bounded point evaluation for $H^2(\mu)$, and such that the span of the evaluation functionals is dense in $H^2(\mu)$.*

Proof. Let p be the minimal polynomial such that V is the intersection of Z_p with $\mathbf{D}^2$. Let C be the projective closure of Z_p in $\mathbf{CP}^2$. Let S be the desingularization of C. This means that S is a compact Riemann surface (not connected if C is not irreducible) and there is a holomorphic function $\phi: S \to C$ that is biholomorphic from S' onto C' and finite-to-one from $S\backslash S'$ onto $C\backslash C'$. Here C' is the set of non-singular points in C, and S' is the preimage of C'. See, e.g., [15] or [17] for details of the desingularization.

Let $\Omega = \phi^{-1}(V)$. Then $\partial\Omega$ is a finite union of disjoint curves, each of which is analytic except possibly at a finite number of cusps. Let ν be the measure from Lemma 1.1 (or the sum of these if Ω is not connected).

The desired measure μ is the push-forward of ν by ϕ, i.e. it is defined by $\mu(E) = \nu(\phi^{-1}(E))$. Indeed, if λ is in V and $\phi(\zeta) = \lambda$, let $k_\zeta \nu$ be a representing measure for ζ in $A(\Omega)$. Then the function $k_\zeta \circ \phi^{-1}$ is defined μ-a.e., and satisfies

$$\int_{\partial V} p(k_\zeta \circ \phi^{-1})\, d\mu = \int_{\partial\Omega} (p\circ\phi) k_\zeta\, d\nu = p\circ\phi(\zeta) = p(\lambda). \qquad \square$$

Note that $\{g\circ\phi : g \in A(V)\}$ is a finite-codimensional subalgebra of $A(\Omega)$. For a description of what finite-codimensional subalgebras look like, see Gamelin's paper [16].

For positive integers m and n, let

$$U = \begin{pmatrix} A & B \\ C & D \end{pmatrix} : \mathbf{C}^m \oplus \mathbf{C}^n \longrightarrow \mathbf{C}^m \oplus \mathbf{C}^n \tag{1.3}$$

be a unitary $(m+n)\times(m+n)$-matrix. Let

$$\Psi(z) = A + zB(I - zD)^{-1}C \tag{1.4}$$

be the $m\times m$-matrix-valued function defined on the unit disk $\mathbf{D}$ by the entries of U. This is called the *transfer function* of U. Because $U^*U=I$, a calculation yields

$$I-\Psi(z)^*\Psi(z)=(1-|z|^2)C^*(I-\bar{z}D^*)^{-1}(I-zD)^{-1}C, \tag{1.5}$$

so $\Psi(z)$ is a rational matrix-valued function that is unitary on the unit circle and contractive on the unit disk. Such functions are called rational matrix inner functions, and it is well known that all rational matrix inner functions have the form (1.4) for some unitary matrix decomposed as in (1.3)—see, e.g., [4] for a proof.

Let V be the set

$$V=\{(z,w)\in\mathbf{D}^2:\det(\Psi(z)-wI)=0\}. \tag{1.6}$$

We shall show that V is a distinguished variety, and that every distinguished variety arises in this way.

LEMMA 1.7. *Let*

$$U'=\begin{pmatrix} D^* & B^* \\ C^* & A^* \end{pmatrix}:\mathbf{C}^n\oplus\mathbf{C}^m\longrightarrow\mathbf{C}^n\oplus\mathbf{C}^m,$$

let

$$\Psi'(z)=D^*+zB^*(I-zA^*)^{-1}C^*$$

and let

$$V'=\{(z,w)\in\mathbf{D}^2:\det(\Psi'(w)-zI)=0\}.$$

Then $V=V'$.

Proof. The point $(z,w)\in\mathbf{D}^2$ is in V if and only if there is a non-zero vector v_1 in $\mathbf{C}^m$ such that

$$[A+zB(1-zD)^{-1}C]v_1=wv_1. \tag{1.8}$$

CLAIM. (1.8) *holds if and only if there is a non-zero vector* v_2 *in* $\mathbf{C}^n$ *such that*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} v_1 \\ zv_2 \end{pmatrix}=\begin{pmatrix} wv_1 \\ v_2 \end{pmatrix}. \tag{1.9}$$

Proof of the claim. If (1.9) holds, then solving gives (1.8). Conversely, if (1.8) holds, define

$$v_2=(I-zD)^{-1}Cv_1.$$

Then (1.9) holds. Moreover, if v_2 were 0, then v_1 would be in the kernel of C and be a w-eigenvector of A. As $A^*A+C^*C=I$, this would force $|w|=1$, contradicting the fact that $(z,w)\in\mathbf{D}^2$. $\square$

Given the claim, the point (z, w) is in V' if and only if there are non-zero vectors v_1 and v_2 such that

$$\begin{pmatrix} D^* & B^* \\ C^* & A^* \end{pmatrix}\begin{pmatrix} v_2 \\ wv_1 \end{pmatrix} = \begin{pmatrix} zv_2 \\ v_1 \end{pmatrix}. \tag{1.10}$$

Interchanging coordinates, (1.10) becomes

$$\begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix}\begin{pmatrix} wv_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ zv_2 \end{pmatrix}. \tag{1.11}$$

Clearly, (1.9) and (1.11) are equivalent. □

Note that if C has a non-trivial kernel $\mathcal{N}$, then (1.5) shows that $\Psi(z)$ is isometric on $\mathcal{N}$ for all z, and so by the maximum principle is equal to a constant isometry with initial space $\mathcal{N}$. If C has a trivial kernel, we say that Ψ is *pure*. Every rational inner function decomposes into the direct sum of a pure rational inner function and a unitary matrix—see, e.g., [24]. Since $A^*A+C^*C=I$, we see that C has no kernel if and only if $\|A\|<1$. Since $AA^*+BB^*=I$, this in turn is equivalent to B^* having no kernel. Therefore Ψ is pure if and only if Ψ' is pure.

THEOREM 1.12. *The set V, defined by* (1.6) *for some rational matrix inner function Ψ, is a distinguished variety. Moreover, every distinguished variety can be represented in this form.*

Proof. Suppose that V is given by (1.6), and that (z, w) is in $\overline{V}$. Without loss of generality, we can assume that Ψ is pure. Indeed, any unitary summand of Ψ would add sheets to the variety $\det(\Psi(z)-wI)=0$ of the type $\mathbf{C}\times\{w_0\}$, for some unimodular w_0. These sheets are all disjoint from the open bidisk $\mathbf{D}^2$.

If $|z|<1$, equation (1.5) then shows that $\Psi(z)$ is a strict contraction, so all its eigenvalues must have modulus less than 1, and so $|w|<1$ also. To prove that $|w|<1$ implies $|z|<1$, just apply the same argument to V'. Therefore (0.2) holds, and V is a distinguished variety.

To prove that all distinguished varieties arise in this way, let V be a distinguished variety. Let μ be the measure from Lemma 1.2, and let $H^2(\mu)$ be the closure of the polynomials in $L^2(\mu)$. The set of bounded point evaluations for $H^2(\mu)$ is precisely V. (It cannot be larger, because $\overline{V}$ is polynomially convex, and Lemma 1.2 ensures that it is not smaller).

Let $T=(T_1, T_2)$ be the pair of operators on $H^2(\mu)$ given by multiplication by the coordinate functions. They are pure commuting isometries[1] because the span of the

[1] A *pure isometry* S is one that has no unitary summand; this is the same as requiring that $\bigcap_{i=1}^{\infty}\operatorname{ran}(S^i)=\{0\}$.

evaluation functionals is dense. The joint eigenfunctions of their adjoints are the evaluation functionals.

By the Sz.-Nagy–Foiaş model theory [24], T_1 can be modelled as M_z, multiplication by the independent variable z on $H^2 \otimes \mathbf{C}^m$, a vector-valued Hardy space on the unit circle. In this model, T_2 can be modelled as M_Ψ, multiplication by $\Psi(z)$ for some pure rational matrix inner function Ψ. A point (z,w) in $\mathbf{D}^2$ is a bounded point evaluation for $H^2(\mu)$ if and only if $(\bar{z}, \bar{w})$ is a joint eigenvalue for (T_1^*, T_2^*). In terms of the unitarily equivalent Sz.-Nagy–Foiaş model, this is equivalent to $\bar{w}$ being an eigenvalue of $\Psi(z)^*$.

Therefore

$$V = \{(z,w) \in \mathbf{D}^2 : \det(\Psi(z) - wI) = 0\},$$

as desired. □

If Ψ is the transfer function of a unitary U as in (1.3), and Ψ is pure, we shall say that V is of *rank* (m,n). This means that generically there are m sheets above each z, and n sheets above each w.

2. Inner functions

Rudin's results [22] show that planar annuli can be mapped isomorphically into distinguished varieties by a pair of inner functions. The advantage of doing this is that the coordinate functions are then easier to deal with than the original inner functions. Inner functions on a finite bordered Riemann surface can be shown to satisfy an algebraic equation. In this section, we show that even without the Riemann surface structure, inner functions must satisfy an algebraic equation. The result is reminiscent of Livšic's Cayley–Hamilton theorem for a pair of commuting operators with finite-rank imaginary parts—see, e.g., the book [19].

Let X be a set. By a *kernel* on X we mean a self-adjoint map $k: X \times X \to \mathbf{C}$ that is positive definite, in the sense that for any finite set $\{\lambda_1, ..., \lambda_N\}$ of distinct points in X, the self-adjoint matrix $k(\lambda_j, \lambda_i)$ is positive definite. Given any kernel k, there is a Hilbert space $\mathcal{H}_k$ of functions on X for which k is the reproducing kernel, i.e.

$$\langle f(\,\cdot\,), k(\,\cdot\,, \lambda) \rangle = f(\lambda) \quad \text{for all } f \in \mathcal{H}_k \text{ and } \lambda \in X.$$

(For details of the passage between a kernel and a Hilbert function space, see, e.g., [4].)

Let ϕ_1 and ϕ_2 be functions on X with modulus less than one at every point. Assume that we can find some kernel k on X so that multiplication by each ϕ_i is a pure isometry on $\mathcal{H}_k$ with finite-dimensional cokernel. For example, X could be a distinguished variety, the ϕ_i's could be the coordinate functions, and $\mathcal{H}_k$ could be the closure of the polynomials in $L^2(\partial X)$. Or, X could be a smoothly bounded planar domain, the ϕ_i's could be inner functions that are continuous on $\overline{X}$ and have finitely many zeroes, and $\mathcal{H}_k$ could be the closure in $L^2(\partial X)$ of the rational functions with poles off $\overline{X}$.

Let $e_1, ..., e_m$ be an orthonormal basis for $(\phi_1\mathcal{H}_k)^\perp$. Then

$$\{\phi_1^i e_j : i\in\mathbf{N},\, 1\leqslant j\leqslant m\}$$

is an orthonormal basis for $\mathcal{H}_k$. So by Bergman's formula [4, Proposition 2.18],

$$k(\zeta,\lambda)=\sum_{i=0}^{\infty}\sum_{j=1}^{m}\phi_1^i(\zeta)e_j(\zeta)\overline{\phi_1^i(\lambda)e_j(\lambda)}=\frac{\sum_{j=1}^m e_j(\zeta)\overline{e_j(\lambda)}}{1-\phi_1(\zeta)\overline{\phi_1(\lambda)}}. \tag{2.1}$$

Similarly, if $f_1, ..., f_n$ is an orthonormal basis for $(\phi_2\mathcal{H}_k)^\perp$, we get

$$k(\zeta,\lambda)=\frac{\sum_{j=1}^n f_j(\zeta)\overline{f_j(\lambda)}}{1-\phi_2(\zeta)\overline{\phi_2(\lambda)}}. \tag{2.2}$$

Equating the right-hand sides of (2.1) and (2.2), and cross-multiplying, we get

$$\sum_{j=1}^{m}e_j(\zeta)\overline{e_j(\lambda)}+\sum_{i=1}^{n}\phi_1(\zeta)f_i(\zeta)\overline{\phi_1(\lambda)f_i(\lambda)}=\sum_{j=1}^{m}\phi_2(\zeta)e_j(\zeta)\overline{\phi_2(\lambda)e_j(\lambda)}+\sum_{i=1}^{n}f_i(\zeta)\overline{f_i(\lambda)}. \tag{2.3}$$

Let $f(\zeta)$ be the vector in $\mathbf{C}^n$ with components $f_1(\zeta), ..., f_n(\zeta)$, and let

$$e(\zeta)=(e_1(\zeta), ..., e_m(\zeta))^t.$$

Then (2.3) can be rewritten as saying that the map

$$U=\begin{pmatrix} A & B\\ C & D\end{pmatrix}: \mathbf{C}^m\oplus\mathbf{C}^n\longrightarrow\mathbf{C}^m\oplus\mathbf{C}^n,$$

$$\begin{pmatrix} e(\zeta)\\ \phi_1(\zeta)f(\zeta)\end{pmatrix}\longmapsto\begin{pmatrix}\phi_2(\zeta)e(\zeta)\\ f(\zeta)\end{pmatrix}$$

is an isometry on the linear span of the vectors

$$\left\{\begin{pmatrix} e(\zeta)\\ \phi_1(\zeta)f(\zeta)\end{pmatrix}:\zeta\in X\right\}.$$

Even if these vectors do not span all of $\mathbf{C}^m\oplus\mathbf{C}^n$, we can always extend U to be unitary from $\mathbf{C}^m\oplus\mathbf{C}^n$ onto $\mathbf{C}^m\oplus\mathbf{C}^n$, and we shall assume that we have done this.

Let

$$\Psi(z)=A+zB(I-zD)^{-1}C \tag{2.4}$$

be the $m\times m$-matrix-valued function defined on the unit disk $\mathbf{D}$ that is the transfer function of U. Moreover, we have

$$\Psi(\phi_1(\zeta))e(\zeta)=\phi_2(\zeta)e(\zeta).$$

Therefore the points $(\phi_1(\zeta),\phi_2(\zeta))$ all lie in the set

$$V=\{(z,w)\in\mathbf{D}^2:\det(\Psi(z)-wI)=0\}, \tag{2.5}$$

which we know from Theorem 1.12 is a distinguished variety. Thus we have proved the following theorem:

THEOREM 2.6. *Let $\mathcal{H}_k$ be a reproducing kernel Hilbert space on a set X. Let ϕ_1 and ϕ_2 be multipliers of $\mathcal{H}_k$ such that multiplication by each ϕ_i is a pure isometry with finite-dimensional cokernel, and such that $|\phi_i(\zeta)|<1$ for all $\zeta\in X$. With notation as above, the function*

$$\zeta\longmapsto(\phi_1(\zeta),\phi_2(\zeta))$$

maps X into the distinguished variety V given by (2.5).

Note that applying Theorem 2.6 to $H^2(\mu)$, the space in Lemma 1.2, we get the second part of Theorem 1.12.

If V is a distinguished variety, an inner function on V may or may not extend to an inner function on $\mathbf{D}^2$. If it does extend, the extension may not be unique. It is curious, however, that there is a rigidity in the degree of this extension. Let ϕ be a rational inner function on $\mathbf{D}^2$. Then it can be represented as

$$\phi(\zeta)=\frac{\zeta^d\,\overline{p(1/\bar{\zeta})}}{p(\zeta)} \tag{2.7}$$

for some polynomial p that does not vanish on $\mathbf{D}^2$ [21], where $\zeta=(\zeta_1,\zeta_2)$ and d is a multi-index. The representation is not unique—e.g. taking $p(z,w)=i(z^2-w^2)$ and $d=(2,2)$, one gets the constant function 1. The representation will be unique if p is restricted so that $Z_p\cap\mathbf{T}^2$ is finite. In this event, we shall call $d=(d_1,d_2)$ the degree of ϕ.

If ϕ is an inner function in $A(\mathbf{D}^2)$, then it is rational and, moreover, the function p will not vanish on $\overline{\mathbf{D}}^2$ [21, Theorem 5.2.5]; we shall call such a function *regular.*

THEOREM 2.8. *Let V be a variety of rank $n=(n_1,n_2)$, and let ϕ be a regular rational inner function on $\mathbf{D}^2$ of degree d. Then ϕ restricted to V has exactly $n\cdot d=n_1d_1+n_2d_2$ zeroes, counting multiplicities.*

Proof. By applying an automorphism of $\mathbf{D}^2$, we can assume that $(0,0)$ is not in V and that all points with first or second coordinate 0 are regular.

Consider first the case $\phi(z,w)=z^{d_1}w^{d_2}$, i.e. $p\equiv 1$ in (2.7). Then each of the n_1 points in V with second coordinate 0 has a zero of multiplicity d_1, and each of the n_2 points in V with first coordinate 0 has a zero of multiplicity d_2.

Now let p be an arbitrary polynomial that does not vanish on $\overline{\mathbf{D}}^2$, normalized so that $p(0,0)=1$. Let $p_r(\zeta)=p(r\zeta)$ and

$$\phi_r(\zeta)=\frac{\zeta^d\,\overline{p_r(1/\bar\zeta)}}{p_r(\zeta)}.$$

As r increases from 0 to 1, the function ϕ_r changes continuously from ζ^d to ϕ. As each ϕ_r is in $A(V)$ and is inner, the number of zeroes must remain constant. □

Example. Let V be the distinguished variety $\{(z,w):z^2=w^3\}$, of rank $(3,2)$. The inner function $\phi(z,w)=z^2$ can be extended to either the function z^2 of degree $(2,0)$, or w^3 of degree $(0,3)$. In either event, $n\cdot d=6$.

3. A sharpening of Andô's inequality

THEOREM 3.1. *Let T_1 and T_2 be commuting contractive matrices, neither of which has eigenvalues of modulus 1. Then there is a distinguished variety V such that, for any polynomial p in two variables, the inequality*

$$\|p(T_1,T_2)\|\leqslant\|p\|_V \tag{3.2}$$

holds.

Proof. Let the dimension of the space on which the matrices act be N.

(i) First, let us assume that each T_r has N linearly independent unit eigenvectors, $\{v_j\}_{j=1}^N$. So we have

$$T_rv_j=\lambda_j^rv_j,\quad r=1,2,\ 1\leqslant j\leqslant N,$$

for some set of scalars $\{\lambda_j^r\}$. As each T_r is a contraction, we have that $I-T_r^*T_r$ is positive semi-definite, so

$$\langle (I-T_r^*T_r)v_j,v_i\rangle=(1-\bar\lambda_i^r\lambda_j^r)\langle v_j,v_i\rangle\geqslant 0. \tag{3.3}$$

As the matrix in (3.3) is positive semi-definite, it can be represented as the Grammian of vectors u_j^r, which can be chosen to lie in a Hilbert space of dimension d_r equal to the defect of T_r (the defect of T_r is the rank of $I-T_r^*T_r$). So we have

$$(1-\bar\lambda_i^1\lambda_j^1)\langle v_j,v_i\rangle=\langle u_j^1,u_i^1\rangle, \tag{3.4}$$

$$(1-\bar\lambda_i^2\lambda_j^2)\langle v_j,v_i\rangle=\langle u_j^2,u_i^2\rangle. \tag{3.5}$$

Multiplying the first equation by $1-\bar{\lambda}_i^2\lambda_j^2$ and the second equation by $1-\bar{\lambda}_i^1\lambda_j^1$, we see that they are equal. Therefore

$$(1-\bar{\lambda}_i^1\lambda_j^1)\langle u_j^2, u_i^2\rangle = (1-\bar{\lambda}_i^2\lambda_j^2)\langle u_j^1, u_i^1\rangle. \tag{3.6}$$

Reordering equation (3.6), we get

$$\langle u_j^1, u_i^1\rangle + \bar{\lambda}_i^1\lambda_j^1\langle u_j^2, u_i^2\rangle = \langle u_j^2, u_i^2\rangle + \bar{\lambda}_i^2\lambda_j^2\langle u_j^1, u_i^1\rangle. \tag{3.7}$$

Equation (3.7) says that there is some unitary matrix

$$U = \begin{pmatrix} A & B \\ C & D \end{pmatrix} : \mathbf{C}^{d_1}\oplus\mathbf{C}^{d_2} \longrightarrow \mathbf{C}^{d_1}\oplus\mathbf{C}^{d_2} \tag{3.8}$$

such that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} u_j^1 \\ \lambda_j^1 u_j^2 \end{pmatrix} = \begin{pmatrix} \lambda_j^2 u_j^1 \\ u_j^2 \end{pmatrix}. \tag{3.9}$$

If the linear span of the vectors $u_j^1\oplus\lambda_j^1 u_j^2$ is not all of $\mathbf{C}^{d_1}\oplus\mathbf{C}^{d_2}$, then U will not be unique. In this event, we just choose one such U. Define the $d_1\times d_1$-matrix-valued analytic function Ψ by

$$\Psi(z) = A + zB(1-zD)^{-1}C. \tag{3.10}$$

For any function Θ of two variables, scalar- or matrix-valued, let

$$\Theta^{\cup}(Z,W) := [\Theta(Z^*, W^*)]^*.$$

Let $\Phi = \Psi^{\cup}$, so that

$$\Phi(z) = A^* + zC^*(1-zD^*)^{-1}B^*.$$

Equation (3.9) implies that

$$\Psi(\lambda_j^1)u_j^1 = [\Phi(\bar{\lambda}_j^1)]^* u_j^1 = \lambda_j^2 u_j^1. \tag{3.11}$$

Let s be the Szegő kernel in the Hardy space H^2 of the unit disk, so that

$$s_\lambda(z) = \frac{1}{1-\bar{\lambda}z}. \tag{3.12}$$

Let k_j be the vector in $H^2\otimes\mathbf{C}^{d_1}$ given by

$$k_j := s_{\bar{\lambda}_j^1}\otimes u_j^1.$$

Consider the pair of isometries (M_z, M_Φ) on $H^2\otimes\mathbf{C}^{d_1}$, where M_z is multiplication by the coordinate function (times the identity matrix on $\mathbf{C}^{d_1}$) and M_Φ is multiplication by the matrix function Φ. Then

$$\begin{aligned} M_z^*&: k_j \longmapsto \lambda_j^1 k_j,\\ M_\Phi^*&: k_j \longmapsto \lambda_j^2 k_j. \end{aligned}$$

Therefore the map that sends each v_j to k_j gives a unitary equivalence between (T_1,T_2) and the pair (M_z^*, M_Φ^*) restricted to the span of the vectors $\{k_j\}_{j=1}^N$. Therefore the pair (M_z^*, M_Φ^*), acting on the full space $H^2\otimes\mathbf{C}^{d_1}$, is a co-isometric extension of (T_1,T_2).

Let p be any polynomial (scalar- or matrix-valued) in two variables. We have

$$\begin{aligned} \|p(T_1,T_2)\| &= \|p(M_z^*, M_\Phi^*)|_{\bigvee\{k_j\}}\| \leqslant \|p(M_z^*, M_\Phi^*)\|_{H^2\otimes\mathbf{C}^{d_1}}\\ &= \|p^\cup(M_z, M_\Phi)\|_{H^2\otimes\mathbf{C}^{d_1}} \leqslant \|p^\cup(M_z, M_\Phi)\|_{L^2\otimes\mathbf{C}^{d_1}} = \|p^\cup\|_{\partial V^\cup}, \end{aligned} \tag{3.13}$$

where $V^\cup$ and V are the sets

$$\begin{aligned} V^\cup &= \{(z,w)\in\mathbf{D}^2 : \det(\Phi(z)-wI)=0\},\\ V &= \{(z,w)\in\mathbf{D}^2 : \det(\Psi(z)-wI)=0\}. \end{aligned} \tag{3.14}$$

Equality (3.13) follows from the observation that

$$\|p^\cup(M_z, M_\Phi)\|_{L^2\otimes\mathbf{C}^{d_1}} = \sup_\theta \|p^\cup(e^{i\theta}I, \Phi(e^{i\theta}))\|, \tag{3.15}$$

where the norm on the right is the operator norm on the $d_1\times d_1$-matrices. Equation (1.5) shows that, except possibly for the finite set $\sigma(D)\cap\mathbf{T}$, the matrix $\Phi(e^{i\theta})$ is unitary, and so the norm of any polynomial applied to $\Phi(e^{i\theta})$ is just the maximum value of the norm of the polynomial on the spectrum of $\Phi(e^{i\theta})$. By continuity, we obtain (3.13). Taking complex conjugates, (3.13) gives

$$\|p(T_1,T_2)\| \leqslant \|p\|_V,$$

the desired inequality.

By Theorem 1.12, we see that V and $V^\cup$ are distinguished varieties, and by construction, V contains the points $\{(\lambda_j^1,\lambda_j^2) : 1\leqslant j\leqslant N\}$.

(ii) Now, we drop the assumption that $T=(T_1,T_2)$ be diagonizable. J. Holbrook proved that the set of diagonizable commuting matrices is dense in the set of all commuting matrices [18]. So we can assume that there is a sequence $T^{(n)}=(T_1^{(n)},T_2^{(n)})$ of commuting matrices that converges to T in norm and such that each pair satisfies the hypotheses of (i), i.e. each $T^{(n)}$ is a pair of commuting contractions that have N linearly independent eigenvectors and no unimodular eigenvalues. Each $T^{(n)}$ has a unitary U_n associated to it as in (3.8). By passing to a subsequence if necessary, we can assume that the defects d_1 and d_2 are constant, and that the matrices U_n converge to a unitary U. The corresponding functions Ψ_n from (3.10) will converge to some function Ψ. Let $q_n(z,w)=\det(\Psi_n(z)-wI)$ and $q(z,w)=\det(\Psi(z)-wI)$. Let V be defined by (3.14) for this Ψ, and V_n be the variety corresponding to Ψ_n. Notice that the degrees of q_n are uniformly bounded.

CLAIM. *V is non-empty.*

Indeed, otherwise it would contain no points of the form $(0,w)$ for $w\in\mathbf{D}$. That would mean that $\sigma(A)\subseteq\mathbf{T}$, and so B and C would be zero. That in turn would mean that the submatrices A_n in U_n would have all their eigenvalues tending to $\mathbf{T}$, and hence by (3.9), the eigenvalues of $T_2^{(n)}$ would all tend to $\mathbf{T}$. Therefore T_2 would have a unimodular eigenvalue, contradicting the hypotheses.

CLAIM. *V is a distinguished variety.*

This follows from Theorem 1.12.

CLAIM. *The inequality* (3.2) *holds.*

This follows from continuity. Indeed, fix some polynomial p. For every $\varepsilon>0$, and for every $n\geqslant n(\varepsilon)$, we have

$$\|p(T)\|\leqslant\varepsilon+\|p(T^{(n)})\|\leqslant\varepsilon+\|p\|_{V_n}.$$

We wish to show that

$$\lim_{n\to\infty}\|p\|_{V_n}\leqslant\|p\|_V.$$

Suppose the contrary. Then there is some sequence (z_n,w_n) in V_n such that

$$|p(z_n,w_n)|\geqslant\|p\|_V+\varepsilon \tag{3.16}$$

for some $\varepsilon>0$. Moreover, we can assume that (z_n,w_n) converges to some point (z_0,w_0) in $\overline{\mathbf{D}}^2$. The point (z_0,w_0) is in the zero set of q, so if it were in $\mathbf{D}^2$, then it would be in V. Otherwise, (z_0,w_0) must be in $\mathbf{T}^2$. To ensure that (z_0,w_0) is in $\overline{V}$, we must rule out the possibility that some sheet of the zero set of q just grazes the boundary of $\mathbf{D}^2$ without ever coming inside.

But this cannot happen. For every z in $\mathbf{D}$, there are d_1 roots of $\det(\Psi(z)-wI)=0$, and *all* of these occur in $\mathbf{D}$. So as z tends to z_0 from inside $\mathbf{D}$, one of the d_1 branches of w must tend to w_0 from inside the disk too. Therefore (z_0,w_0) is in the closure of V, and (3.16) cannot happen. $\square$

Remark 1. If T_1 has a unimodular eigenvalue λ, then the corresponding eigenspace $\mathcal{H}'$ will be reducing for T_2. Indeed, writing

$$T_1 = \begin{pmatrix} \lambda I & 0 \\ 0 & T_1'' \end{pmatrix} \quad \text{and} \quad T_2 = \begin{pmatrix} T_2' & X \\ 0 & T_2'' \end{pmatrix},$$

the commutativity of T_1 and T_2 means $X(T_1''-\lambda)=0$. As λ is not in the spectrum of T_1'', it follows that $X=0$.

Therefore for any polynomial p, we have

$$\|p(T_1,T_2)\| = \max\{\|p(\lambda I, T_2')\|, \|p(T_1'', T_2'')\|\}. \tag{3.17}$$

By von Neumann's inequality for one matrix, the first entry on the right-hand side of (3.17) is majorized by

$$\|p\|_{\{\lambda\}\times \mathbf{D}}.$$

So if we allow the matrices to have unimodular eigenvalues, we can still obtain (3.2) by adding to V a finite number of disks in the boundary of $\mathbf{D}^2$. The new V, however, will not be a distinguished variety.

Remark 2. Once one knows Andô's inequality for matrices, then it follows for all commuting contractions by approximating them by matrices—see [13] for an explicit construction. Of course, the set V must be replaced by the limit points of the sets that occur at each stage of the approximation, and in general this may be the whole bidisk.

Remark 3. We have actually constructed a co-isometric extension of T that is localized to V, and a unitary dilation of T with spectrum contained in ∂V.

4. The uniqueness variety

A *solvable Pick problem on* $\mathbf{D}^2$ is a set $\{\lambda_1, ..., \lambda_N\}$ of points in $\mathbf{D}^2$ and a set $\{w_1, ..., w_N\}$ of complex numbers such that there is some function ϕ of norm less than or equal to 1 in $H^\infty(\mathbf{D}^2)$ that interpolates (satisfies $\phi(\lambda_i)=w_i$ for all $1\leqslant i\leqslant N$). An *extremal Pick problem* is a solvable Pick problem for which no function of norm less than 1 interpolates. The points λ_i are called the *nodes*, and w_i are called the *values*. By *interpolating function* we mean any function in the closed unit ball of $H^\infty(\mathbf{D}^2)$ that interpolates.

Consider the following two examples, in the case $N=2$.

Example 1. Let $\lambda_1=(0,0)$, $\lambda_2=\left(\frac{1}{2},0\right)$, $w_1=0$ and $w_2=\frac{1}{2}$. Then a moment's thought reveals that the interpolating function is unique, and is given by $\phi(z,w)=z$.

Example 2. Let $\lambda_1=(0,0)$, $\lambda_2=(\frac{1}{2},\frac{1}{2})$, $w_1=0$ and $w_2=\frac{1}{2}$. Then the interpolating function is far from unique—either coordinate function will do, as will any convex combination of them. (A complete description of all solutions is given by J. Ball and T. Trent in [8]). But on the distinguished variety $\{(z,z): z\in\mathbf{D}\}$, all solutions coincide by Schwarz's lemma. For an arbitrary solvable Pick problem, let $\mathcal{U}$ be the set of points in $\mathbf{D}^2$ on which all the interpolating functions in the closed unit ball of $H^\infty(\mathbf{D}^2)$ have the same value. The preceding examples show that $\mathcal{U}$ may be either the whole bidisk or a proper subset. In the event that $\mathcal{U}$ is not the whole bidisk, it is a variety. Indeed, for any λ_{N+1} not in U, there are two distinct values w_{N+1} and w'_{N+1} so that the corresponding $(N+1)$-point Pick problem has a solution. By [8] and [2] these problems have interpolating functions that are rational, of degree bounded by $2(N+1)$. The set $\mathcal{U}$ must lie in the zero set of the difference of these rational functions. Taking the intersection over all λ_{N+1} not in $\mathcal{U}$, one gets that $\mathcal{U}$ is the intersection of the zero sets of polynomials. Therefore $\mathcal{U}$ is a variety, and indeed, by factoring these polynomials into their irreducible factors, we see that $\mathcal{U}$ is the intersection with the bidisk of the zero set of one polynomial, together with possibly a finite number of isolated points. We shall call $\mathcal{U}$ the *uniqueness variety.* (If the problem is not extremal, $\mathcal{U}$ is just the original set of nodes.)

We shall say that an N-point extremal Pick problem is *minimal* if none of the $(N-1)$-point subproblems is extremal. The main result of this section is that if the uniqueness variety is not the whole bidisk, then it at least contains a distinguished variety running through the nodes. If $N=3$, it is shown in [3] that either $\mathcal{U}=\mathbf{D}^2$ or the minimal extremal problem has a solution that is a function of one coordinate function only.

THEOREM 4.1. *Let $N\geqslant 2$, and let $\lambda_1,...,\lambda_N$ and $w_1,...,w_N$ be the data for a minimal extremal Pick problem on the bidisk. The uniqueness variety $\mathcal{U}$ contains a distinguished variety V that contains each of the nodes.*

For a point λ in $\mathbf{D}^2$, we shall write λ^1 and λ^2 for the first and second coordinates, respectively. Given a set of points $\{\lambda_1,...,\lambda_N\}$ in $\mathbf{D}^2$, an *admissible kernel* K is a positive definite $N\times N$-matrix, with all the diagonal entries 1, such that

$$[(1-\lambda_i^r\bar{\lambda}_j^r)K_{ij}]\geqslant 0,\quad r=1,2. \tag{4.2}$$

A theorem of the first author [1] asserts that a Pick problem on $\mathbf{D}^2$ is solvable if and only if, for every admissible kernel K, the matrix

$$[(1-w_i\bar{w}_j)K_{ij}] \tag{4.3}$$

is positive semi-definite (see [11], [8] and [2] for alternative proofs). We shall say that an admissible kernel is *active* if the matrix (4.3) has a non-trivial null space, i.e. if it is positive semi-definite but not positive definite.

LEMMA 4.4. *A solvable Pick problem has an active kernel if and only if it is extremal.*

Proof. ($\Rightarrow$) If the problem were not extremal, then for some $\varrho<1$ one would have

$$[(\varrho^2 I - w_i \bar{w}_j) K_{ij}] \geqslant 0 \tag{4.5}$$

for all admissible kernels. Take K to be an active kernel, with γ a non-zero vector in the null space of $[(I-w_i\bar{w}_j)K_{ij}]$. Then taking the inner product of the left-hand side of (4.5) applied to γ with γ gives $-(1-\varrho^2)\|\gamma\|^2$, which is negative.

($\Leftarrow$) As the problem is extremal, for each $\varrho<1$ there is some admissible kernel K such that $[(\varrho^2 I - w_i\bar{w}_j)K_{ij}]$ is not positive semi-definite. By compactness of the set of positive semi-definite $N\times N$-matrices with 1's down the diagonal, there therefore exists some positive semi-definite K, satisfying (4.2), and such that (4.3) is not positive definite. It just remains to show that this K is actually positive definite, and therefore a kernel.

Suppose it were not, so that for some non-zero vector $v=(v^1, ..., v^N)^t$, we have $Kv=0$. By (4.2), for each $r=1,2$, the vector $\lambda^r \cdot v$ (i.e. the vector whose ith component is $\lambda_i^r v^i$) is also in the null space of K. Iterating this observation, one gets that for any polynomial p, the vector

$$p(\lambda)\cdot v = \begin{pmatrix} p(\lambda_1)v^1 \\ \vdots \\ p(\lambda_N)v^N \end{pmatrix}$$

is in the null space of K. Taking p to be a polynomial that is 1 at λ_1 and zero on the other nodes, we get $K_{11}=0$, a contradiction. $\square$

LEMMA 4.6. *Every admissible kernel on a set $\{\lambda_1, ..., \lambda_N\}$ can be extended to a continuous admissible kernel on a distinguished variety that contains the points $\{\lambda_1, ..., \lambda_N\}$.*

Proof. Let K be an admissible kernel on the set $\{\lambda_1, ..., \lambda_N\}$. As it is positive definite, there are vectors v_i in $\mathbf{C}^N$ such that $K_{ij}=\langle v_j, v_i\rangle$. Because K is admissible, equations (3.4) and (3.5) hold. Following the proof of Theorem 3.1, one gets that for every point (z,w) in the variety V given by (3.14), one has non-zero vectors $\hat{u}^1(z,w)$ and $\hat{u}^2(z,w)$ such that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \hat{u}^1(z,w) \\ z\hat{u}^2(z,w) \end{pmatrix} = \begin{pmatrix} w\hat{u}^1(z,w) \\ \hat{u}^2(z,w) \end{pmatrix}.$$

Moreover, as the vector $(\hat{u}^1, \hat{u}^2)^t$ must just be chosen in the null space of

$$\begin{pmatrix} A-wI & zB \\ C & zD-I \end{pmatrix},$$

it can be chosen continuously. When (z,w) is one of the nodes λ_j, we choose

$$\hat{u}^1(\lambda_j^1,\lambda_j^2)=u_j^1,$$
$$\hat{u}^2(\lambda_j^1,\lambda_j^2)=u_j^2.$$

Normalize the vectors so that

$$\|\hat{u}^1(z,w)\|=\sqrt{1-|z|^2}.$$

Now let

$$k(z,w)=s_{\bar{z}}\otimes\hat{u}^1(z,w),$$

where s is the Szegő kernel on the disk as in (3.12).

The desired extension of K to V is given by

$$\widehat{K}(\zeta,\lambda)=\langle k(\lambda),k(\zeta)\rangle.$$

This is obviously a kernel that extends K, it is continuous on $V\times V$ by construction, and the fact that it is admissible follows, in the language of Theorem 3.1, from the fact that M_z and M_Ψ are contractions. □

Proof of Theorem 4.1.

Step 1. By Lemma 4.4, the problem has an extremal kernel, and by Lemma 4.6, this kernel can be extended to a distinguished variety V that contains all the nodes. Let us call the extended kernel K.

Let $\gamma=(\gamma^1,...,\gamma^N)$ be a non-zero vector in the null space of $[(1-w_i\bar{w}_j)K_{ij}]$. Let $\lambda_{N+1}=(\lambda_{N+1}^1,\lambda_{N+1}^2)$ be any point in V that is not one of the original nodes. Let w_{N+1} be some possible value that an interpolating function can take at λ_{N+1}. As the $(N+1)$-point Pick problem with nodes $\lambda_1,...,\lambda_{N+1}$ and values $w_1,...,w_{N+1}$ is solvable, and as K is admissible, we must have that

$$[(1-w_i\bar{w}_j)K_{ij}]_{i,j=1}^{N+1}\geqslant 0.$$

Therefore, for every $t\in\mathbf{C}$, we have

$$\left\langle [(1-w_i\bar{w}_j)K_{ij}]\binom{\gamma}{t},\binom{\gamma}{t}\right\rangle\geqslant 0. \tag{4.7}$$

As γ is in the null space of $[(1-w_i\bar{w}_j)K_{ij}]_{i,j=1}^N$, inequality (4.7) reduces to

$$2\,\mathrm{Re}\left[\bar{t}\sum_{j=1}^N(1-\bar{w}_jw_{N+1})K_{N+1,j}\gamma^j\right]+|t|^2(1-|w_{N+1}|^2)\geqslant 0. \tag{4.8}$$

As this holds for all t, we must have that the linear term vanishes, and so we can solve for w_{N+1} and get

$$w_{N+1}\left(\sum_{j=1}^{N} \bar{w}_j K_{N+1,j}\gamma^j\right) = \sum_{j=1}^{N} K_{N+1,j}\gamma^j, \tag{4.9}$$

$$w_{N+1} = \frac{\sum_{j=1}^{N} K_{N+1,j}\gamma^j}{\sum_{j=1}^{N} \bar{w}_j K_{N+1,j}\gamma^j}. \tag{4.10}$$

As long as both sides of (4.9) do not reduce to zero, this gives a formula for w_{N+1}, which must therefore be unique.

Step 2. So far, we have not used the minimality of the problem. Minimality ensures that no component of γ can be zero, for otherwise an $(N-1)$-point subproblem would have an active kernel.

Fix one of the nodes, λ_1 say, and consider what happens when λ_{N+1} tends to λ_1 along some sheet of V. By continuity, $K_{N+1,j}$ tends to $K_{1,j}$ for each j. If

$$\sum_{j=1}^{N} \bar{w}_j K_{1,j}\gamma^j \neq 0,$$

then by continuity

$$\sum_{j=1}^{N} \bar{w}_j K_{N+1,j}\gamma^j \neq 0$$

for λ_{N+1} in V and close to λ_1, and so formula (4.10) gives the unique value that the interpolating function must take at λ_{N+1}.

Assume instead that

$$\sum_{j=1}^{N} \bar{w}_j K_{1,j}\gamma^j = 0. \tag{4.11}$$

Consider the N-point Pick problem with nodes $\lambda_1, ..., \lambda_N$, and values $w_1+\varepsilon, w_2, ..., w_N$ for some ε in $\mathbf{C}$. If this problem were solvable, then, since K is an admissible kernel, one would have

$$[(1-w_i'\bar{w}_j')K_{ij}] \geqslant 0, \tag{4.12}$$

where

$$w_i' = \begin{cases} w_i, & i \neq 1, \\ w_1+\varepsilon, & i=1. \end{cases}$$

Take the inner product of the left-hand side of (4.12) applied to γ with γ. We get

$$\sum_{i,j=1}^{N}(1-w_i'\bar{w}_j')K_{ij}\gamma^j\bar{\gamma}^i=\sum_{i,j=1}^{N}(1-w_i\bar{w}_j)K_{ij}\gamma^j\bar{\gamma}^i \\ -2\,\mathrm{Re}\left[\varepsilon\bar{\gamma}^1\sum_{j=1}^{N}\bar{w}_jK_{1,j}\gamma^j\right]-|\varepsilon|^2K_{11}|\gamma^1|^2. \tag{4.13}$$

The first sum in (4.13) vanishes because γ is in the null space of $[(1-w_i\bar{w}_j)K_{ij}]$. The second sum vanishes by hypothesis (4.11). Therefore for any $\varepsilon\neq 0$, (4.13) is negative. This means that the value w_1 at λ_1 is uniquely determined by the choice of the other $N-1$ values at $\lambda_2, ..., \lambda_N$. Therefore this $(N-1)$-point subproblem must be extremal, contradicting the minimality hypothesis.

We therefore conclude that (4.10) gives a well-defined formula for the unique value of w_{N+1} at points λ_{N+1} in V near the nodes. As we know that some solution to the problem is given by a rational function, we therefore know that this rational function gives the unique solution near the nodes. Hence the union of the irreducible components of V that contain the nodes is a distinguished variety contained in $\mathcal{U}$. □

Question 4.14. Is the distinguished variety constructed in the proof equal to all of $\mathcal{U}$? Given any function on any subset of the bidisk, the result in [1] tells whether it can be extended to a function in the closed unit ball of $H^\infty(\mathbf{D}^2)$. If the set is a distinguished variety, and the function is analytic on it, is there a better criterion, which one might think of as solving problem (b) in the introduction?

Question 4.15. How can one tell whether a function on a distinguished variety extends to all of $\mathbf{D}^2$ without increasing its norm?

References

[1] Agler, J., Some interpolation theorems of Nevanlinna–Pick type. Preprint, 1988.
[2] Agler, J. & McCarthy, J. E., Nevanlinna–Pick interpolation on the bidisk. *J. Reine Angew. Math.*, 506 (1999), 191–204.
[3] — The three point Pick problem on the bidisk. *New York J. Math.*, 6 (2000), 227–236.
[4] — *Pick Interpolation and Hilbert Function Spaces.* Grad. Stud. Math., 44. Amer. Math. Soc., Providence, RI, 2002.
[5] — Norm preserving extensions of holomorphic functions from subvarieties of the bidisk. *Ann. of Math.*, 157 (2003), 289–312.
[6] Ahern, P. R. & Sarason, D., The H^p spaces of a class of function algebras. *Acta Math.*, 117 (1967), 123–163.
[7] Andô, T., On a pair of commutative contractions. *Acta Sci. Math.* (*Szeged*), 24 (1963), 88–90.

[8] BALL, J. A. & TRENT, T. T., Unitary colligations, reproducing kernel Hilbert spaces, and Nevanlinna–Pick interpolation in several variables. *J. Funct. Anal.*, 157 (1998), 1–61.

[9] BALL, J. A. & VINNIKOV, V., Hardy spaces on a finite bordered Riemann surface, multivariable operator model theory and Fourier analysis along a unimodular curve, in *Systems, Approximation, Singular Integral Operators, and Related Topics* (Bordeaux, 2000), pp. 37–56. Oper. Theory Adv. Appl., 129. Birkhäuser, Basel, 2001.

[10] CLANCEY, K. F., Representing measures on multiply connected planar domains. *Illinois J. Math.*, 35 (1991), 286–311.

[11] COLE, B. J. & WERMER, J., Pick interpolation, von Neumann inequalities, and hyperconvex sets, in *Complex Potential Theory* (Montreal, PQ, 1993), pp. 89–129. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 439. Kluwer, Dordrecht, 1994.

[12] — Andô's theorem and sums of squares. *Indiana Univ. Math. J.*, 48 (1999), 767–791.

[13] DRURY, S. W., Remarks on von Neumann's inequality, in *Banach spaces, Harmonic Analysis, and Probability Theory* (Storrs, CT, 1980/81), pp. 14–32. Lecture Notes in Math., 995. Springer, Berlin, 1983.

[14] FARKAS, H. M. & KRA, I., *Riemann Surfaces*, 2nd edition. Graduate Texts in Math., 71. Springer, New York, 1992.

[15] FISCHER, G., *Plane Algebraic Curves.* Stud. Math. Libr., 15. Amer. Math. Soc., Providence, RI, 2001.

[16] GAMELIN, T. W., Embedding Riemann surfaces in maximal ideal spaces. *J. Funct. Anal.*, 2 (1968), 123–146.

[17] GRIFFITHS, P. & HARRIS, J., *Principles of Algebraic Geometry.* Wiley, New York, 1978.

[18] HOLBROOK, J. A. R., Polynomials in a matrix and its commutant. *Linear Algebra Appl.*, 48 (1982), 293–301.

[19] LIVŠIC, M. S., KRAVITSKY, N., MARKUS, A. S. & VINNIKOV, V., *Theory of Commuting Nonselfadjoint Operators.* Math. Appl., 332. Kluwer, Dordrecht, 1995.

[20] NIKOLSKI, N. & VASYUNIN, V., Elements of spectral theory in terms of the free function model, I: Basic constructions, in *Holomorphic Spaces* (Berkeley, CA, 1995), pp. 211–302. Math. Sci. Res. Inst. Publ., 33. Cambridge Univ. Press, Cambridge, 1998.

[21] RUDIN, W., *Function Theory in Polydiscs.* Benjamin, New York–Amsterdam, 1969.

[22] — Pairs of inner functions on finite Riemann surfaces. *Trans. Amer. Math. Soc.*, 140 (1969), 423–434.

[23] SZ.-NAGY, B. & FOIAŞ, C., Commutants de certains opérateurs. *Acta Sci. Math. (Szeged)*, 29 (1968), 1–17.

[24] — *Harmonic Analysis of Operators on Hilbert Space.* North-Holland, Amsterdam–London, 1970.

[25] WERMER, J., Analytic disks in maximal ideal spaces. *Amer. J. Math.*, 86 (1964), 161–170.

JIM AGLER
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093
U.S.A.
jagler@euclid.ucsd.edu

JOHN E. MCCARTHY
Department of Mathematics
Washington University
St. Louis, MO 63130
U.S.A.
mccarthy@wustl.edu

Received August 31, 2004
Received in revised form March 31, 2005