Operator Monotone Functions and Löwner Functions of Several Variables

Jim Agler * John E. McCarthy † N. J. Young ‡

September 21, 2010

Abstract

We prove generalizations of Löwner's results on matrix monotone functions to several variables. We give a characterization of when a function of d variables is locally monotone on d-tuples of commuting self-adjoint n-by-n matrices. We prove a generalization to several variables of Nevanlinna's theorem describing analytic functions that map the upper half-plane to itself and satisfy a growth condition. We use this to characterize all rational functions of two variables that are operator monotone.

1 Introduction

In 1934, K. Löwner published a very influential paper [19] studying functions on an open interval $E \subseteq \mathbb{R}$ that are matrix monotone, *i.e.* functions f with the property that whenever S and T are self-adjoint matrices whose spectra are in E then

$$S \leq T \qquad \Rightarrow \qquad f(S) \leq f(T). \tag{1.1}$$

This property is equivalent (see Subsection 1.3) to being locally matrix monotone, *i.e.* if $S(t)$ is a C^1 arc of self-adjoint matrices with $\sigma(S(t)) \subset E$ then

$$S'(t) \geq 0 \qquad \Rightarrow \qquad \frac{d}{dt} f(S(t)) \geq 0. \tag{1.2}$$

*Partially supported by National Science Foundation Grant DMS 0801259
†Partially supported by National Science Foundation Grant DMS 0966845
‡Partially supported by London Mathematical Society Grant 4918

Roughly speaking, Löwner showed that if one fixes a dimension n and wants (1.1) or (1.2) to hold for n-by-n self-adjoint matrices, then certain matrices derived from the values of f must all be positive semi-definite. As n increases, the conditions become more restrictive. In the limit as $n \to \infty$ (equivalently, if one passes to self-adjoint operators on an infinite dimensional Hilbert space), then a necessary and sufficient condition is that the function f must have an analytic continuation to a function F that maps the upper half-plane Π to itself.

The goal of this paper is to extend the above notions to several variables. In particular, we want to study functions of d variables applied to d-tuples of commuting self-adjoint operators. Given two d-tuples $S = (S^1, \dots, S^d)$ and $T = (T^1, \dots, T^d)$, we shall say that $S \leq T$ if and only if $S^r \leq T^r$ for every $1 \leq r \leq d$. We want to study functions that satisfy (1.1) or (1.2) for d-tuples.

Before we can describe our results, we must first give a more detailed description of the one-dimensional case. We recommend the book [10] by W. Donoghue for a well-written modern account.

Note that there is another approach to extending Löwner's results to several variables where the operators $S^1, \dots, S^d$ act on different spaces $\mathcal{H}^1, \dots, \mathcal{H}^d$, and $f(S)$ is interpreted to act on $\mathcal{H}^1 \otimes \dots \otimes \mathcal{H}^d$. We refer the reader to the papers [12, 28, 17] and references therein.

Let us remark that we frequently make an *a priori* assumption that the function f be C^1. This is not always necessary; but it makes the statements of theorems cleaner. The class of matrix monotone functions is convex and closed under pointwise limits, so if f is in the class, one can convolve it with a smooth bump function to get an approximation that is matrix monotone on a slightly smaller set.

1.1 Dimension one

Let E be an open set in $\mathbb{R}$, and let $n \geq 2$ be a natural number. The Löwner class $\mathcal{L}^1_n(E)$ is the set of C^1 functions $f : E \to R$ with the property that, whenever $\{x_1, \dots, x_n\}$ is a set of n distinct points in E, then the matrix A, defined by

$$A_{ij} \;=\; \begin{cases} \dfrac{f(x_j) - f(x_i)}{x_j - x_i} & \text{if} \quad i \neq j \\ \left.\dfrac{\partial f}{\partial x}\right|_{x_i} & \text{if} \quad i = j, \end{cases}$$

is positive semi-definite.

We shall let M_n denote the n-by-n complex matrices, SAM_n the self-adjoint n-by-n matrices, and SA the bounded self-adjoint operators on an infinite dimensional separable Hilbert space.

Definition 1.3. *A function f is locally n-matrix monotone on the open set $E \subset \mathbb{R}$ if, whenever S is in SAM_n with $\sigma(S)$ consisting of n distinct points in E, and $S(t)$ is a C^1 curve in SAM_n with $S(0) = S$ and $\frac{d}{dt}S(t)|_{t=0} \geq 0$, then $\frac{d}{dt}f(S(t))|_{t=0} \geq 0$.*

Remark 1.4. This definition is slightly different from the one in the first paragraph, where the eigenvalues were not required to be distinct. We use this definition to be consistent with the multivariable Definition 1.9 below. However, using formula (6.6.31) in [13] for $\frac{d}{dt}f(S(t))$, it is easy to show that in the one variable case the two different definitions are equivalent.

We shall say that f is n-matrix monotone on E, or M_n-monotone, if, whenever S and T are in SAM_n and all their eigenvalues lie in E, then (1.1) holds. To emphasize the difference from locally monotone, we shall also call n-matrix monotone functions *globally M_n-monotone*. Replacing SAM_n by SA, we get the definitions of locally operator monotone and operator monotone.

Theorem 1.5 (Löwner)**.** *Let $E \subseteq \mathbb{R}$ be open, and let $f \in C^1(E)$. Then f is locally n-matrix monotone on E if and only if f is in $\mathcal{L}^1_n(E)$.*

Definition 1.6. *The Pick class on E, denoted $\mathcal{P}(E)$, is the set of real-valued functions f on E for which there exists an analytic function $F : \Pi \to \overline{\Pi}$ such that F extends analytically across E and*

$$\lim_{y \searrow 0} F(x+iy) = f(x) \quad \forall\, x \in E.$$

Theorem 1.7 (Löwner)**.** *Let $E \subseteq \mathbb{R}$ be open, and let $f \in C^1(E)$. The following are equivalent:*

(i) The function f is locally operator monotone on E.
(ii) The function f is in $\mathcal{L}^1_n(E)$ for all n.
(iii) The function f is in $\mathcal{P}(E)$.

1.2 Dimension $d \geq 2$: Local results

We shall let $CSAM_n^d$ denote the set of d-tuples of commuting self-adjoint n-by-n matrices, and CSA^d be the set of d-tuples of commuting self-adjoint bounded operators. If S is a commuting d-tuple of self-adjoint operators acting on the Hilbert space $\mathcal{H}$, and f is a real-valued continuous (indeed, measurable) function on the spectrum of S in $\mathbb{R}^d$, then $f(S)$ is a well-defined self-adjoint operator on $\mathcal{H}$.

Definition 1.8. *Let E be an open set in $\mathbb{R}^d$, and f be a real-valued C^1 function on E. Say f is locally operator monotone on E if, whenever S is in CSA^d with $\sigma(S) \subset E$, and $S(t)$ is a C^1 curve in CSA^d with $S(0) = S$ and $\frac{d}{dt}S(t)|_{t=0} \geq 0$, then $\frac{d}{dt}f(S(t))|_{t=0}$ exists and is ≥ 0.*

We shall not concern ourselves in this paper on what conditions on f guarantee that $f(S(t))$ is differentiable; for these see *e.g.* [22].

Definition 1.9. *Let E be an open set in $\mathbb{R}^d$, and f be a real-valued C^1 function on E. We say f is locally M_n-monotone on E if, whenever S is in $CSAM_n^d$ with $\sigma(S) = \{x_1, \dots, x_n\}$ consisting of n distinct points in E, and $S(t)$ is a C^1 curve in $CSAM_n^d$ with $S(0) = S$ and $\frac{d}{dt}S(t)|_{t=0} \geq 0$, then $\frac{d}{dt}f(S(t))|_{t=0}$ exists and is ≥ 0.*

We define the Löwner classes in d variables, $\mathcal{L}_n^d(E)$, by:

Definition 1.10. *Let E be an open subset of $\mathbb{R}^d$. The set $\mathcal{L}_n^d(E)$ consists of all real-valued C^1-functions on E that have the following property: whenever $\{x_1, \dots, x_n\}$ are n distinct points in E, there exist positive semi-definite n-by-n matrices $A^1, \dots, A^d$ so that*

$$
\begin{aligned}
A^r(i,i) &= \left.\frac{\partial f}{\partial x^r}\right|_{x_i} \\
\text{and} \quad f(x_j) - f(x_i) &= \sum_{r=1}^{d} (x_j^r - x_i^r) A^r(i,j) \qquad \forall\ 1 \leq i, j \leq n.
\end{aligned}
$$

Here is our d-variable version of Theorem 1.5.

Theorem 7.24 *Let E be an open set in $\mathbb{R}^d$, and f a real-valued C^1 function on E. Then f is locally M_n-monotone if and only if f is in $\mathcal{L}_n^d(E)$.*

In generalizations of Theorem 1.7, there turns out to be a difference between the case $d = 2$ and $d > 2$.

Definition 1.11. *The Löwner class, $\mathcal{L}^d$, is the set of functions $F : \Pi^d \to \overline{\Pi}$ with the property that there exist d positive semi-definite functions A^r, $1 \leq r \leq d$, on Π^d such that*

$$F(z) - \overline{F(w)} \ = \ (z^1 - \bar{w}^1)A^1(z,w) \ + \ \ldots \ + \ (z^d - \bar{w}^d)A^d(z,w).$$

When $d = 1$ or 2, the Löwner class coincides with the set of all analytic functions from Π^d to $\overline{\Pi}$, but for $d \geq 3$ it is a proper subset (see Section 5).

Definition 1.12. *Let $E \subseteq \mathbb{R}^d$ be open. The class $\mathcal{L}(E)$ is the set of real-valued functions f on E for which there exists an analytic function F in $\mathcal{L}^d$ such that F extends analytically across E and*

$$\lim_{y \searrow 0} F(x^1 + iy, \ldots, x^d + iy) = f(x^1, \ldots, x^d) \ \ \forall \ x \ \in \ E.$$

Theorem 8.1 *Let E be an open set in $\mathbb{R}^d$, and f a real-valued C^1 function on E. The following are equivalent:*
(i) The function f is locally operator monotone on E.
(ii) The function f is in $\mathcal{L}^d_n(E)$ for all n.
(iii) The function f is in $\mathcal{L}(E)$.

1.3 Local to Global

In one variable, provided E is an interval, local monotonicity implies global monotonicity immediately. Indeed, suppose $S \leq T$, and let $S(t) = (1-t)S + tT$. Then $S'(t) = T - S \geq 0$, so

$$f(T) - f(S) \ = \ \int_0^1 \frac{d}{dt} f(S(t)) \ dt \ \geq 0. \tag{1.13}$$

If E is not convex, this argument fails. Indeed, the function $-1/x$ is locally n-matrix monotone on $\mathbb{R} \setminus \{0\}$ for all n; but it is only globally monotone on sets that lie entirely on one side of 0.

For intervals, (1.13) shows that the word "locally" can be dropped in both Theorem 1.5 and 1.7. One problem in going to several variables is that this simple argument no longer works, because one may not be able to connect S and T by a path of commuting d-tuples. Indeed, the following example shows that there need not be any commuting tuples between two given ones.

Example 1.14. Let S and T be pairs in $CSAM_2^2$ given by

$$\begin{aligned} S &= \left(\begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right) \\ T &= \left(\begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix}, \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix} \right). \end{aligned}$$

If R is in $CSAM_2^2$ and $S \leq R \leq T$, then either $R = S$ or $R = T$.

We have been unable to resolve the question of whether the n-matrix monotone functions on a connected open set E are a proper subset of the locally n-matrix monotone functions on E. However, as n tends to infinity and we pass to locally operator monotone functions, analyticity enters the picture, and makes the problem more tractable — see Subsection 1.5.

1.4 The Nevanlinna Representation

To prove $(iii) \Rightarrow (i)$ in Theorem 1.7, one must understand analytic functions that map the upper half-plane to itself. A key fact is a characterization due to R. Nevanlinna [20] which says that, provided they have some regularity at infinity, they are all Cauchy transforms of measures on the line.

Theorem 1.15 (Nevanlinna). *If $F : \Pi \to \Pi$ is analytic and satisfies*

$$\limsup_{y \to \infty} y \, |F(iy) - C| < \infty,$$

for some $C \in \mathbb{R}$, then there exists a unique finite positive Borel measure ν on $\mathbb{R}$ so that

$$F(z) \ = \ C + \int \frac{d\nu(t)}{t - z}. \tag{1.16}$$

Nevanlinna's theorem was used by M. Stone to prove the spectral theorem [29], but one can adopt the reverse viewpoint, and write (1.16) in terms of the resolvent of a self-adjoint. Indeed, let X be the self-adjoint operator of multiplication by the independent variable on $L^2(\nu)$, and v the vector in $L^2(\nu)$ that is 1 a.e. Then (1.16) can be rewritten as

$$F(z) \ = \ C + \langle (X - z)^{-1} v, v \rangle. \tag{1.17}$$

This representation turns out to be useful in studying operator monotonicity, because then

$$F(S) \;=\; CI + R_v^*(I\otimes X - S\otimes I)^{-1}R_v, \qquad (1.18)$$

where $R_v : \mathcal{H} \to \mathcal{H}\otimes\mathcal{M}$ is given by $R_v : \xi \mapsto \xi\otimes v$

There is a several variable analogue of Theorem 1.15. It may require first perturbing F.

Definition 1.19. *For each real number t, define*

$$\rho_t(z) = \frac{z+t}{1-tz}.$$

For $F \in \mathcal{L}^d$, define

$$F_t \;:=\; \rho_t \circ F \circ \rho_t.$$

The following theorem follows from Theorem 6.33. We shall say that a function F on Π^d is analytic on a neighborhood of infinity if the function $F(1/z^1,\dots,1/z^d)$ extends to be analytic on a neighborhood of the origin. In Theorem 6.33, a weaker assumption is placed on F than being analytic in a neighborhood of infinity.

Theorem 1.20. *Let F be in $\mathcal{L}^d$, and assume that F is analytic in a neighborhood of infinity. Then for all sufficiently small t, except for at most countably many exceptions, the function F_t has the following representation. There is a Hilbert space $\mathcal{M}$, a densely defined self-adjoint operator X on $\mathcal{M}$, a vector v in $\mathcal{M}$, and d orthogonal projections $P^1,\dots,P^d$ with $\sum_{r=1}^d P^r = I_{\mathcal{M}}$ so that*

$$F_t(z) \;=\; C + \langle (X - \sum_{r=1}^d z^r P^r)^{-1} v, v\rangle. \qquad (1.21)$$

1.5 Dimension $d \geq 2$: Global operator monotonicity

Using the representation (1.21), we can prove results on (global) operator monotonicity. With notation as in Theorem 1.20, let us say that the μ-resolvent of X is the set of points

$$\{(z^1,\dots,z^d) \;\in\; \mathbb{C}^d \,:\, (X - \sum_{r=1}^d z^r P^r) \text{ has a bounded inverse}\}.$$

Theorem 9.2 *Let X be a densely-defined self-adjoint operator on a Hilbert space $\mathcal{M}$, let v be a vector in $\mathcal{M}$, and let $P^1, \ldots, P^d$ be projections with orthogonal ranges that sum to the identity. Let F be given by*

$$F(z) \;=\; C + \langle (X - \sum_{r=1}^{d} z^r P^r)^{-1} v, v\rangle.$$

Let E be an open rectangle in $\mathbb{R}^d$ that is in the μ-resolvent of X. Then F is globally operator monotone on E.

As an application, we can give a complete characterization of the rational functions of two variables that are operator monotone on rectangles.

Theorem 9.6 *Let F be a rational function of two variables. Let Γ be the zero-set of the denominator of F. Assume F is real-valued on $\mathbb{R}^2 \setminus \Gamma$. Let E be an open rectangle in $\mathbb{R}^2 \setminus \Gamma$. Then F is globally operator monotone on E if and only if F is in $\mathcal{L}(E)$.*

2 Some Notation

We shall let $\mathbb{D}$ denote the unit disk in the complex plane, Π the upper half-plane $\{z : \mathrm{Im}\,(z) > 0\}$, and $\mathbb{H}$ the right half-plane $\{z : \mathrm{Re}\,(z) > 0\}$. We shall let

$$\alpha(\lambda) \;=\; i\frac{1+\lambda}{1-\lambda}$$

be a linear fractional map that maps $\mathbb{D}$ to Π, and

$$\beta(z) \;=\; \frac{z-i}{z+i}$$

be its inverse.

We shall let d denote the number of variables. If z is a point in Π^d, we shall use $z^1, \ldots, z^d$ to denote its components; likewise $\lambda = (\lambda^1, \ldots, \lambda^d)$ will be a point in $\mathbb{D}^d$. We shall write $S = (S^1, \ldots, S^d)$ for a d-tuple of matrices or operators, and use $\|S\|$ for $\max_{1\leq r\leq d} \|S^r\|$. We shall also use α and β to denote the maps from $\mathbb{D}^d$ to Π^d and back again that are defined by applying α and β coordinate-wise.

A kernel on a set E is a map $K : E \times E \to \mathbb{C}$ with the property that for every finite set $\{\lambda_1, \ldots, \lambda_N\}$ of distinct points in E, the matrix $[K(\lambda_j, \lambda_i)]$ is positive semi-definite.

Definition 2.1. *The Pick class, $\mathcal{P}^d$, is the set of analytic functions $F : \Pi^d \to \overline{\Pi}$.*

Definition 2.2. *The Schur class, $\mathcal{S}^d$, is the set of analytic functions $\varphi : \mathbb{D}^d \to \overline{\mathbb{D}}$.*

Definition 2.3. *The Carathéodory class, $\mathcal{C}^d$, is the set of analytic functions $\psi : \mathbb{D}^d \to \overline{\mathbb{H}}$.*

Definition 2.4. *The Löwner class, $\mathcal{L}^d$, is the set of functions $F : \Pi^d \to \overline{\Pi}$ with the property that there exist d kernel functions A^r, $1 \leq r \leq d$ on Π^d such that*

$$F(z) - \overline{F(w)} \;=\; (z^1 - \bar{w}^1)A^1(z,w) \;+\; \ldots \;+\; (z^d - \bar{w}^d)A^d(z,w). \quad (2.5)$$

Definition 2.6. *The Schur-Agler class, $\mathcal{A}^d$, is the set of functions $\varphi : \mathbb{D}^d \to \overline{\mathbb{D}}$ with the property that there exist d kernel functions B^r, $1 \leq r \leq d$ on $\mathbb{D}^d$ such that*

$$1 - \varphi(\lambda)\overline{\varphi(\mu)} \;=\; (1 - \lambda^1\bar{\mu}^1)B^1(\lambda,\mu) + \ldots + (1 - \lambda^d\bar{\mu}^d)B^d(\lambda,\mu). \quad (2.7)$$

When the dimension is clear, we shall drop the superscript d.

Remark 2.8. If we exclude the constant function 1 from $\mathcal{S}$, we have the identification

$$F \;\in\; \mathcal{P} \quad\iff\quad \beta \circ F \circ \alpha \;\in\; \mathcal{S} \quad\iff\quad -iF \circ \alpha \;\in\; \mathcal{C}. \quad (2.9)$$

Moreover, we also have (again excluding the constant function 1)

$$F \;\in\; \mathcal{L} \quad\iff\quad \beta \circ F \circ \alpha \;\in\; \mathcal{A}, \quad (2.10)$$

(see Lemma 2.11). As all our results are trivial for constant functions, we shall use (2.9) and (2.10) without explicitly mentioning the exclusion of the constant function 1.

Lemma 2.11. *The function $F : \Pi^d \to \mathbb{C}$ is in the Löwner class if and only if $\varphi := \beta \circ F \circ \alpha$ is in the Schur-Agler class $\mathcal{A}^d$.*

Proof: Define $\varphi = \beta \circ F \circ \alpha$. Then φ is in $\mathcal{A}^d$ if and only if there are kernels B^r on $\mathbb{D}$ such that

$$1 - \varphi(\lambda)\overline{\varphi(\mu)} \;=\; \sum_{r=1}^{d}(1 - \lambda^r\bar{\mu}^r)B^r(\lambda,\mu). \quad (2.12)$$

When $z = \alpha(\lambda)$ and $w = \alpha(\mu)$, (2.12) becomes

$$1 - \beta \circ F(z)\overline{\beta \circ F(w)} \;=\; \sum_{r=1}^{d}\left(1 - \left[\frac{z^r - i}{z^r + i}\right]\overline{\left[\frac{w^r - i}{w^r + i}\right]}\right)B^r(\beta(z),\beta(w)). \quad (2.13)$$

Rearranging (2.13), we get

$$F(z) - \overline{F(w)} \;=\; \sum_{r=1}^{d} (z^r - \bar{w}^r)\, \frac{F(z) + i}{z^r + i} \frac{\overline{F(w)} - i}{\overline{w^r} - i} B^r(\beta(z), \beta(w)). \tag{2.14}$$

If A^r is defined for $r = 1, \dots, d$ by

$$A^r(z, w) \;=\; \frac{F(z) + i}{z^r + i} \frac{\overline{F(w)} - i}{\overline{w^r} - i} B^r(\beta(z), \beta(w))$$

(2.14) becomes

$$F(z) - \overline{F(w)} \;=\; \sum_{r=1}^{d} (z^r - \bar{w}^r) A^r(z, w), \tag{2.15}$$

which means F is in $\mathcal{L}^d$. Reversing the argument gives the converse. $\square$

Remark 2.16. It is known that $\mathcal{A}^d = \mathcal{S}^d$ for $d = 1$ or 2, and that for $d \geq 3$ $\mathcal{A}^d \subsetneq \mathcal{S}^d$ [9, 30, 2]. It follows similarly that the Löwner class equals the Pick class in dimensions 1 and 2, and is strictly contained in it for $d \geq 3$. Indeed, using Theorem 5.5.1 of [24], one can show that for each $d \geq 3$, there are rational functions that are real on $\mathbb{R}^d$ and that are in $\mathcal{P}^d \setminus \mathcal{L}^d$.

3 Models, B-points and C-points

For a function φ in $\mathcal{A}^d$, we can take the representation (2.7) and decompose the B^r's as Grammians to get a Hilbert space model for φ. That means we find a separable Hilbert space $\mathcal{M}$, an orthogonal decomposition of $\mathcal{M}$,

$$\mathcal{M} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^d, \tag{3.1}$$

and an analytic map $u : \mathbb{D}^d \to \mathcal{M}$ such that

$$1 - \overline{\varphi(\mu)}\varphi(\lambda) = \sum_{r=1}^{d} (1 - \overline{\mu^r}\lambda^r) \langle u_\lambda^r, u_\mu^r \rangle_{\mathcal{M}^r} \tag{3.2}$$

for all $\lambda, \mu \in \mathbb{D}^d$, where we write u_λ for $u(\lambda)$, P^r for the projection onto $\mathcal{M}^r$, and u_λ^r for $P^r[u_\lambda]$.

We shall view (3.1) interchangeably as a graded Hilbert space (*i.e.* one with a given orthogonal decomposition) or as a single Hilbert space with d given projections $P^1, \dots, P^d$ that are orthogonal and add up to the identity.

In general, if $\eta \in \mathcal{M}$, we set $\eta^r = P^r[\eta]$. If $\lambda \in \mathbb{C}^d$, we may regard λ as an operator on $\mathcal{M}$ by letting

$$\lambda\eta = \lambda^1\eta^1 + \cdots + \lambda^d\eta^d. \tag{3.3}$$

Equation (3.2) can then be rewritten as

$$1 - \overline{\varphi(\mu)}\varphi(\lambda) = \langle (1 - \mu^*\lambda)u_\lambda, u_\mu \rangle. \tag{3.4}$$

A lurking isometry argument yields the following result [2].

Theorem 3.5. *If $(\mathcal{M}, u)$ is a model of $\varphi \in \mathcal{A}^d$, then there exist $a \in \mathbb{C}$, vectors $\beta, \gamma \in \mathcal{M}$ and a linear operator $D : \mathcal{M} \to \mathcal{M}$ such that the operator*

$$\begin{bmatrix} a & 1 \otimes \beta \\ \gamma \otimes 1 & D \end{bmatrix}$$

is a contraction on $\mathbb{C} \oplus \mathcal{M}$ and, for all $\lambda \in \mathbb{D}^d$,

$$\begin{aligned} (1 - D\lambda)u_\lambda &= \gamma, && (3.5) \\ \varphi(\lambda) &= a + \langle \lambda u_\lambda, \beta \rangle. && (3.6) \end{aligned}$$

With notation as in Theorem 3.5, we shall call (a, β, γ, D) a realization of $(\mathcal{M}, u)$.

If we start instead with the representation (2.5) of a function F in $\mathcal{L}^d$, we can decompose the kernels A^r as the Grammians of some vectors v^r, in auxiliary separable Hilbert spaces $\mathcal{N}^r$. Then we get, in the analogous notation to above,

$$\begin{aligned} F(z) - \overline{F(w)} &= \sum_{r=1}^{d} (z^r - \bar{w}^r) A^r(z, w) \\ &= \sum_{r=1}^{d} (z^r - \bar{w}^r) \langle v_z^r, v_w^r \rangle_{\mathcal{N}^r} \\ &= \langle (z - w^*)v_z, v_w \rangle_{\mathcal{N}}. && (3.7) \end{aligned}$$

This decomposition leads to a lurking self-adjoint argument, which we shall discuss in Section 6.

In [4], we introduced the concept of a B-point for $\mathcal{S}$. Let us give a unified definition for each of the classes $\mathcal{S}, \mathcal{P}$ and $\mathcal{C}$; notice that it depends on the codomain of the function.

Definition 3.8. *Let U and V be fixed domains, and $f : U \to \overline{V}$ an analytic function. A point τ in ∂U is called a B-point of f if there is a sequence λ_n of points in U that converge to τ and such that*

$$\frac{\operatorname{dist}(f(\lambda_n), \partial V)}{\operatorname{dist}(\lambda_n, \partial U)} \tag{3.9}$$

is bounded.

So, for example, a point τ in $\partial \Pi^d$ is a B-point for a function F in $\mathcal{P}^d$ if there exists some sequence z_n in Π^d that tends to τ and such that the quantity

$$\frac{\operatorname{Im} F(z_n)}{\min_{r \in \{1,\ldots,d\}} (\operatorname{Im} z_n^r)}$$

is bounded.

For a function in $\mathcal{L}^d$ (respectively, $\mathcal{A}^d$) we shall call a point τ a B-point if it is a B-point for the function thought of as an element of $\mathcal{P}^d$ (resp. $\mathcal{S}^d$).

For each of the three classes $\mathcal{S}, \mathcal{P}$, and $\mathcal{C}$, it follows from results of F. Jafari [14] and M. Abate [1] that if τ is a B-point, then the ratio (3.9) remains bounded for every sequence λ_n that tends to τ non-tangentially. Moreover, the function f will then have a non-tangential limit at τ. (A sequence λ_n in U tends to the point τ non-tangentially if λ_n tends to τ and

$$\frac{\operatorname{dist}(\lambda_n, \tau)}{\operatorname{dist}(\lambda_n, \partial U)}$$

is bounded.)

The following result was proved in [4] for $d = 2$, but the proof generalizes to any d. We shall need it in the proof of Theorem 6.26.

Lemma 3.10. *Let $\varphi \in \mathcal{A}^d$ and $\tau \in \mathbb{T}^d$. Let $(\mathcal{M}, u)$ be a model for φ, and (a, β, γ, D) be a realization. The following are equivalent.*

(i) τ is a B-point for φ.

(ii) For some sequence λ_n converging to τ non-tangentially, the sequence $\|u_{\lambda_n}\|$ is bounded.

(iii) For any sequence λ_n converging to τ non-tangentially, the sequence $\|u_{\lambda_n}\|$ is bounded.

(iv) The vector γ is in the range of $(I - D\tau)$.

Moreover, if u_{λ_n} converges to a vector weakly as λ_n tends to τ non-tangentially, then u_{λ_n} converges in norm. The vector $u_\tau := \lim_{r \nearrow 1} u_{r\tau}$ exists for every B-point τ.

A stronger condition than being a B-point is being a C-point.

Definition 3.11. *A point $x \in \mathbb{R}^d$ is a C-point for $F \in \mathcal{L}$ if there are complex numbers $\eta^1, \ldots, \eta^d$ and a real number c so that*

$$F(z) - c - \sum_{r=1}^{d} \eta^r (z^r - x^r) \;=\; o(\|z - x\|)$$

as z tends to x non-tangentially.

In particular, if F is differentiable at x and $F(x)$ is real, then x is a C-point for F.

The following result was proved in [4].

Proposition 3.12. *Suppose $F \in \mathcal{L}$ has a model as in (3.7). If x is a C-point for F, then as z converges to x non-tangentially from Π^d, the vectors v_z converge in norm, to some vector v_x in $\mathcal{N}$.*

4 Analytically continuing Pick functions

Suppose F is analytic on Π^d, and E is an open set in $\mathbb{R}^d$. What conditions on F guarantee that it can be analytically continued across E? The edge-of-the-wedge theorem (see Theorem 4.11 below) is a common tool to give such extensions. Checking the hypotheses, however, requires knowledge of the values of F as one approaches points of E not just non-tangentially but also tangentially. If F is in the Pick class $\mathcal{P}^d$, Theorem 4.8 says that it suffices to know that every point of E is in a B-point (which can be checked by looking at the values of F on the inward-pointing normal).

As we are using bars to denote closure, we shall use stars for the complex conjugate of a set, and write Π^* for the lower half-plane.

4.1 One dimension

To understand the situtation, let us first consider the one dimensional case. Let $\psi : \mathbb{D} \to \mathbb{H}$ be non-constant. Then ψ has a Herglotz representation; if we assume $\psi(0)$ is positive, then

$$\psi(z) \; = \; \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} d\mu(\theta) \tag{4.1}$$

for some positive measure μ. There is an elegant analysis of when ψ has B-points in the paper [27] by D. Sarason, where the following two propositions are proved. Proposition 4.2 is originally due to M. Riesz [23], and Proposition 4.4 to R. Nevanlinna [21].

Proposition 4.2. *Let ψ be given by (4.1), and let τ be a point in $\mathbb{T}$. Then ψ has a B-point at τ if and only if*

$$\int \frac{1}{|e^{i\theta}-\tau|^2} d\mu(\theta) \; < \; \infty. \tag{4.3}$$

If $\varphi = \beta \circ (i\psi)$ is the Cayley transform of ψ, there is a distinction between B-points where $\varphi(\tau)$ equals 1, corresponding to $\psi(\tau) = \infty$, and all other cases.

Proposition 4.4. *Let $\varphi = \dfrac{\psi - 1}{\psi + 1}$, where ψ is given by (4.1), and let τ be a point in $\mathbb{T}$. Then φ has a B-point at τ with $\varphi(\tau) \neq 1$ if and only if (4.3) holds. The function φ has a B-point at τ with $\varphi(\tau) = 1$ if and only if τ is a mass point of μ.*

Suppose now that $\varphi : \mathbb{D} \to \mathbb{D}$ has an open arc I of B-points. Can φ be extended analytically across I? If we know that φ omits a value on I, then the answer is yes. Indeed, after a Möbius map, we can assume that φ is the Cayley transform of some ψ as (4.1). If condition (4.3) holds on an open arc I, then μ must vanish on I by Lemma 4.5 below. But then the formula (4.1) gives an analytic function on the extended plane less $\mathbb{T} \setminus I$.

However, without the assumption that φ omits a value, the answer may be no, as Example 4.6 shows.

Lemma 4.5. *Suppose μ is a measure on $[-\pi, \pi)$ and (4.3) holds for $\tau = e^{ix}$ for every x in an open arc $I \subset [-\pi, \pi)$. Then $\mu(I) = 0$.*

Proof: For μ a.e. point x in I, there is a constant $c > 0$ such that

$$\mu[x - \frac{1}{k}, x + \frac{1}{k}] \; \geq \; c\frac{2}{k},$$

by the Lebesgue differentiation theorem [26, Chap. 8]. For such an x, we have

$$\begin{aligned}\int_{-\pi}^{\pi} \frac{1}{|e^{i\theta}-\tau|^2} d\mu(\theta) &\geq \int_{x-1/k}^{x+1/k} \frac{1}{|e^{i\theta}-\tau|^2} d\mu(\theta) \\ &\geq \frac{1}{|1-e^{i/k}|^2}\, \mu[x-\frac{1}{k}, x+\frac{1}{k}] \\ &\geq ck.\end{aligned}$$

Letting k tend to infinity, the integral would be infinite; so μ must put no mass on I. $\square$

Example 4.6. Here is an example of a function in the Schur class of one variable that has B-points at every point of $\mathbb{T}$ but that cannot be analytically continued across every arc.

Let $\tau_n = e^{ix_n}$ be a sequence in $\mathbb{T}$ that converges to 1. Let $c_n = 2^{-n}|1-\tau_n|^2$. Then for every $\lambda \in \mathbb{T}$, the quantity $\left|\frac{1-\tau_n}{\lambda-\tau_n}\right|$ is less than or equal to 1 for all but finitely many values of n. Therefore

$$\sum 2^{-n} \left|\frac{1-\tau_n}{\lambda-\tau_n}\right|^2 < \infty \tag{4.7}$$

for every λ.

Let $\mu = \sum 2^{-n}\delta_{x_n}$, let F be the Herglotz transform of μ, and let $\varphi = \frac{F-1}{F+1}$. By Proposition 4.4, we have that φ has every point of $\mathbb{T}$ as a B-point, but φ cannot be analytically continued across any arc containing 1, as it takes the value 1 infinitely often on any such arc.

4.2 d dimensions

Our goal is to prove the following analytic continuation theorem:

Theorem 4.8. *Let E be an open subset of $\mathbb{R}^d$. Then there is an open set U in $\mathbb{C}^d$ that contains $\Pi^d \cup E \cup \Pi^{*d}$ with the following property: whenever F is in the Pick class, and every point of E is a B-point for F, then there is an analytic function G on U that agrees with F on Π^d.*

This theorem immediately implies the omit-a-value theorem. Let us say that a subset E' of $\mathbb{T}^d$ is a *B-set* for φ in $\mathcal{S}^d$ if every point of E' is a B-point for φ.

Theorem 4.9. *Let E' be an open subset of $\mathbb{T}^d$. Then there is an open set U in $\mathbb{C}^d$ containing $\mathbb{D}^d \cup E' \cup \{\mathbb{C} \setminus \bar{\mathbb{D}}\}^d$ such that the following two statements are equivalent for any φ in the Schur class:*

(1) *there is an analytic function ψ on U that agrees with φ on $\mathbb{D}^d$;*

(2) *the set E' is a B-set for φ and for every point τ in E' there exists a neighborhood V of τ in $\mathbb{T}^d$ and a point ω in $\mathbb{T}$ such that no nontangential limit of φ at any point of V is equal to ω.*

Condition (2) says that every point of E has a neighborhood where the non-tangential limits of φ omit some value in $\mathbb{T}$.

We start with the following proposition.

Proposition 4.10. *Let E be an open subset of $\mathbb{R}^d$. Then there is an open set U in $\mathbb{C}^d$ that contains $\Pi^d \cup E \cup \Pi^{*d}$ with the following property: if J is a non-empty interval in $\mathbb{R}$, F is in the Pick class, F has non-tangential limits at almost every point of E, and these limits are all in $\mathbb{R} \setminus J$, then there is an analytic function G on U that agrees with F on Π^d.*

Proof: Precomposing F with a Möbius transformation of Π if necessary, we can assume that J is an interval about infinity, so the non-tangential limits are in some compact set $[-M, M]$ a.e.

Let $H(z) = \log(1 + M + F(z))$. Then H maps Π^d into $\{z \in \mathbb{C} : 0 < \operatorname{Im} z < \pi\}$, and

$$\lim_{y \to 0} H(x + iy) \in [0, \log(2M + 1)] \quad \text{a.e. } x \in E.$$

We want to apply the edge-of-the-wedge theorem (Theorem 4.11 below) to H, and this will give the desired open set U to which H, and hence F, extends.

As H has bounded imaginary part, we can pass the limit inside the integral on the left-hand side of (4.12), and as H has real boundary values, we get that the limit is 0. Therefore by Theorem 4.11 we get an analytic extension of H to U. □

Here is the version of the edge-of-the-wedge theorem we want (Theorem C from [25]). We write $\mathbb{R}_+$ for the interval $(0, \infty)$.

Theorem 4.11. *[Edge-of-the-wedge] Let E be an open subset of $\mathbb{R}^d$. Then there is an open set U in $\mathbb{C}^d$ that contains $\Pi^d \cup E \cup \Pi^{*d}$ and is* *such that whenever H is an analytic function on Π^d with the property that for every g in $C_c^\infty(E)$,*

$$\lim_{\mathbb{R}^d_+ \ni y \to 0} \int_E g(x) \mathrm{Im}\, H(x+iy) dx \;=\; 0, \tag{4.12}$$

then there is an analytic function G on U that agrees with H on Π^d.

PROOF OF THEOREM 4.8: We can extend F to Π^{*d} by letting $F(z) = \overline{F(\bar{z})}$ on Π^{*d}. The difficulty is in showing that the definitions of F on the two disjoint domains Π^d and Π^{*d} are analytic continuations of each other across E. This is a local property. If we can show that every point of E has a neighborhood on which the boundary values of F take values in a bounded set, we can apply Proposition 4.10 to conclude that the reflection of F is an analytic continuation of F across this neighborhood in E. Since this is true at every point, the conclusion of the theorem will follow.

For convenience, we will change variables and consider the function $\varphi(\lambda) = (-i) F \circ \alpha$, which is in $\mathcal{C}^d$.

We can normalize to assume that $\varphi(0, \dots, 0) = 1$ and that the point of interest for φ is $\beta(0, \dots, 0) = (-1, \dots, -1)$. So for some $0 < c < \frac{\pi}{5}$, the set

$$\{(e^{i\theta_1}, \dots, e^{i\theta_d}) \;:\; \forall\, 1 \leq r \leq d,\ |\theta_r| \geq \pi - 5c\} \tag{4.13}$$

consists of B-points for φ. In what follows, we shall choose arg to take values in $[-\pi, \pi)$.

For each τ in the set

$$\{\tau \in \mathbb{T}^{d-1} \;:\; |\arg(\tau^j)| < 2c,\ \forall\, 1 \leq j \leq d-1\},$$

define g_τ in $\mathcal{C}^1$ by

$$g_\tau(z) \;=\; g(z, \tau^1 z, \tau^2 z, \dots, \tau^{d-1} z).$$

Then for each τ, the set

$$I_{3c} \;=\; \{\sigma \in \mathbb{T} : |\arg(\sigma)| > \pi - 3c\}$$

is a set of B-points for g_τ, and $g_\tau(0) = 1$. Each g_τ has a Herglotz representation, and by the results of subsection 4.1 the corresponding measure is supported off the set I_{3c}. So

$$g_\tau(z) \;=\; \int_{-\pi+3c}^{\pi-3c} \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu_\tau(\theta)$$

for some probability measure μ_τ. Therefore if σ is in the arc $I_c = \{\sigma \in \mathbb{T} : |\arg(\sigma)| > \pi - c\}$,

$$\begin{aligned} |g_\tau(\sigma)| &\leq \int_{-\pi+3c}^{\pi-3c} \left| \frac{e^{i\theta}+\sigma}{e^{i\theta}-\sigma} \right| d\mu_\tau(\theta) \\ &\leq \sec c \|\mu_\tau\| \\ &= \sec c. \end{aligned}$$

Therefore on the set $(I_c)^d$ we conclude that the non-tangential limits of φ take values in the bounded set $[-\sec c, \sec c]$.

Notice that c does not depend on F: we have shown that for any F, normalized to have $F(i, \ldots, i) = i$, if F has B-points on the set $\alpha((I_{5c})^d)$, then F is bounded on $\alpha((I_c)^d)$. By Proposition 4.10, this latter set now has a neighborhood to which F can be analytically extended, and this neighborhood can be chosen independently of F. So every point x in E has a neighborhood U_x to which all functions F in the Pick class with B-points on E can be extended; let U be the union of all the U_x as x ranges over E. □

5 The Löwner classes

We shall single out functions that have a representation on subsets of $\mathbb{R}^d$ as in (2.5).

Definition 5.1. *Let $E \subseteq \mathbb{R}^d$ be a non-empty open set, and let n be a positive integer. We define $\mathcal{L}_n(E)$ to be the set of real valued differentiable functions that have the following property: whenever $\{x_1, \ldots, x_n\}$ are n distinct points in E, there exist positive semi-definite n-by-n matrices $A^1, \ldots, A^d$ so that*

$$A^r(i,i) = \left. \frac{\partial f}{\partial x^r} \right|_{x_i} \tag{5.2}$$

$$\text{and} \quad f(x_j) - f(x_i) = \sum_{r=1}^{d} (x_j^r - x_i^r) A^r(i,j) \quad \forall\ 1 \leq i,j \leq n. \tag{5.3}$$

We shall give an alternative description of $\mathcal{L}(E)$. We shall temporarily call it $\mathcal{L}_\partial(E)$, but we shall show in Proposition 5.11 that it coincides with the set $\mathcal{L}(E)$ from Definition 1.12.

Definition 5.4. *Let $E \subseteq \mathbb{R}^d$ be a non-empty open set. We shall let $\mathcal{L}_\partial(E)$ denote the set of differentiable real valued functions f on E for* *which there exist positive semi-definite functions* $A^1, \ldots, A^d : E \times E \to \mathbb{C}$ *so that*

$$\begin{aligned} A^r(z,z) &= \left. \frac{\partial f}{\partial x^r} \right|_z && (5.5) \\ \text{and} \quad f(z) - f(w) &= \sum_{r=1}^{d} (z^r - w^r) A^r(z,w). && (5.6) \end{aligned}$$

If $E \subseteq \mathbb{R}^d$, a function f in $\mathcal{L}_\partial(E)$ can be extended to a function F in $\mathcal{L}$ that has f as its non-tangential boundary values on E.

Proposition 5.7. *Let* $E \subseteq \mathbb{R}^d$ *be open, and let* $f \in \mathcal{L}_\partial(E)$*. Then there exists* $F \in \mathcal{L}^d$ *such that every point of* E *is a* B*-point for* F*, and such that*

$$\lim_{z \xrightarrow{\mathrm{nt}} t} F(z) \;=\; f(t) \qquad \forall\, t \in E. \tag{5.8}$$

Proposition 5.7 follows immediately from the corresponding result on the polydisk, Theorem 5.9, which was proved by J.A. Ball and V. Bolotnikov [6] (we are changing their language slightly; they did not explicitly use the notion of B-point).

Theorem 5.9 (Ball-Bolotnikov)**.** *Let* $E' \subseteq \mathbb{T}^d$ *and let* $\psi : E' \to \mathbb{C}$*. Suppose there are positive semi-definite functions* $B^1, \ldots, B^d : E' \times E' \to \mathbb{C}$ *such that, for all* λ, μ *in* E',

$$1 - \psi(\lambda)\overline{\psi(\mu)} \;=\; (1 - \lambda^1 \bar{\mu}^1) B^1(\lambda,\mu) \;+\; \ldots \;+\; (1 - \lambda^d \bar{\mu}^d) B^d(\lambda,\mu). \tag{5.10}$$

Then there is a function φ *in* $\mathcal{A}$ *such that every point of* E' *is a* B*-point for* φ *and*

$$\lim_{\lambda \xrightarrow{\mathrm{nt}} \tau} \varphi(\lambda) \;=\; \psi(\tau) \qquad \forall\, \tau \in E'.$$

Moreover, if φ *is defined to equal* ψ *on* E'*, the kernels* B^r *can be extended to* $E' \cup \mathbb{D}^d$ *so that, for all* λ, μ *in* $E' \cup \mathbb{D}^d$,

$$1 - \varphi(\lambda)\overline{\varphi(\mu)} \;=\; (1 - \lambda^1 \bar{\mu}^1) B^1(\lambda,\mu) \;+\; \ldots \;+\; (1 - \lambda^d \bar{\mu}^d) B^d(\lambda,\mu).$$

We can pass back and forth between regarding functions in $\mathcal{L}(E)$ as functions in the Löwner class $\mathcal{L}^d$ that have B-points on E (and so can be analytically extended across E), and as functions that are characterized by their values on E by (5.4) and can then be analytically extended into Π^d.

Proposition 5.11. *Let $E \subseteq \mathbb{R}^d$ be a non-empty open set. The following four sets coincide.*

(i) $\cap_{n=1}^{\infty} \mathcal{L}_n(E)$.

(ii) The set $\mathcal{L}_\partial(E)$ *defined by Definition 5.4.*

(iii) The set $\mathcal{L}(E)$ *defined by Definition 1.12.*

(iv) The functions f on E for which there exists a function F in $\mathcal{L}^d$ such that every point x of E is a B-point of F and the non-tangential limit of F at x is $f(x)$.

Proof: It is immediate that $(ii) \subseteq (i)$. Theorem 4.8 asserts that $(iii) = (iv)$. Proposition 5.7 says that $(ii) \subseteq (iv)$.

To show $(iii) \subseteq (ii)$, choose a model for F so that (3.7) holds on $\Pi^d \times \Pi^d$:

$$F(Z) - \overline{F(W)} \;=\; \langle (Z - W^*) v_Z, v_W \rangle \qquad \forall\, Z, W \;\in\; \Pi^d.$$

As every point in E is a C-point for F, we can let Z and W tend to points in E non-tangentially, z and w respectively. By Proposition 3.12, the vectors v_Z and v_W converge to v_z and v_w. Let

$$A^r(z,w) \;=\; \langle v_z^r, v_w^r \rangle$$

and one gets (5.6). To get (5.5), let z be in E and let W in Π^d tend to z non-tangentially. As F is analytic at z, we have

$$F(W) - F(z) \;=\; \sum (W^r - z^r) \left. \frac{\partial f}{\partial x^r} \right|_z \;+\; o(\|z - W\|). \tag{5.12}$$

From the model,

$$\begin{aligned} F(W) - F(z) \quad &= \quad \langle (W - z) v_W, v_z \rangle \\ &= \quad \langle (W - z) v_z, v_z \rangle \;+\; \langle (W - z)(v_W - v_z), v_z \rangle \end{aligned} \tag{5.13}$$

The second term on the right of (5.13) is $o(\|z - W\|)$, so comparing with (5.12) we conclude that

$$\left. \frac{\partial f}{\partial x^r} \right|_z \;=\; \langle v_z^r, v_z^r \rangle,$$

and hence (5.5) holds.

To prove $(i) \subseteq (ii)$, we need to show that if (5.5) and (5.6) hold on every finite set, with perhaps a different choice of A^r's each time, then we can make one choice for the A^r's that works everywhere.

Let $f \in \cap_{n=1}^{\infty} \mathcal{L}_n(E)$. Consider any finite set $\{z_1, \ldots, z_n\}$ of distinct points in E. By Definition 5.1 there exist kernels $A^1, \ldots, A^d$ on E such that equations (5.2), (5.3) hold, and we have

$$A^r(i,i) \ \leq \ \frac{\partial f}{\partial x^r}(z_i), \qquad i = 1, \ldots, n, \ r = 1, \ldots, d.$$

Since the matrices A^r are positive we also obtain bounds on the off-diagonal entries of all the A^r. Hence the set K of all d-tuples $(A^1, \ldots, A^d)$ for which equations (5.2), (5.3) hold is a compact non-empty subset of M_n^d.

Moreover, if $(B^1, \ldots, B^d)$ is a d-tuple of kernels on any finite superset Z of $\{z_1, \ldots, z_n\}$ for which the analogs of equations (5.2), (5.3) hold, then the choice of A^r to be the principal submatrix of B^r corresponding to $\{z_1, \ldots, z_n\}$ gives a d-tuple that belongs to K. Therefore by Kurosh's theorem [5, p.74] or [3, p.30], there is a d-tuple $(A^1, \ldots, A^d)$ of kernels on E such that equations (5.2) and (5.3) hold for all points $z_i, \ z_j \in E$. $\square$

6 The μ-spectral theorem

A function in the Pick class of one variable, *i.e.* an analytic function from Π to Π, has an integral representation which can be obtained from the Herglotz representation (4.1) of functions from $\mathbb{D}$ to $\mathbb{H}$ by a change of variables.

Theorem 6.1. *[Herglotz] An analytic function $F : \Pi \to \Pi$ has a unique representation of the form*

$$F(z) \ = \ c + dz + \int \frac{1+zt}{t-z} d\mu(t) \tag{6.2}$$

where $\operatorname{Im} c \geq 0$ and $d \geq 0$, and μ is a finite positive Borel measure on $\mathbb{R}$. Conversely any function of this form is in the Pick class of one variable.

If in addition F decays up the imaginary axis, one gets that F is the Cauchy transform of a finite measure on $\mathbb{R}$. This is called the Nevanlinna representation, and was proved by R. Nevanlinna [20].

Theorem 6.3. *[Nevanlinna] If $F : \Pi \to \Pi$ is analytic and satisfies*

$$\limsup_{y \to \infty} |yF(iy)| < \infty,$$

then there exists a unique finite positive Borel measure ν on $\mathbb{R}$ so that

$$F(z) \;=\; \int \frac{d\nu(t)}{t-z}.$$

Remark 6.4. If one considers $\psi = -iF \circ \alpha : \mathbb{D} \to \mathbb{H}$, then d in Theorem 6.1 is the mass assigned to the point 1 in the Herglotz representation of ψ. Nevanlinna's condition in Theorem 6.3 is equivalent to saying that $\varphi = \beta \circ F \circ \alpha : \mathbb{D} \to \mathbb{D}$ has a B-point at 1 with $\varphi(1) = -1$.

One can prove the spectral theorem for a (possibly unbounded) self-adjoint operator by showing that that, if R_z is the resolvent, then for any vector u the function $\langle R_z u, u\rangle$ is in the one variable Pick class, and satisfies Nevanlinna's growth condition. Then Theorem 6.3 gives the scalar spectral measure. See [10] or [18]. Conversely, if X is the operator of multiplication by the independent variable on $L^2(\mu)$, and v is the constant function 1, then (6.2) becomes

$$F(z) \;=\; c + dz + \langle (1+zX)(X-z)^{-1}v, v\rangle. \tag{6.5}$$

In several variables, there is also a connection between Pick functions and self-adjoint operators, which could be called a μ-spectral theorem.

Definition 6.6. *Let $\mathcal{M}$ be a Hilbert space, with a fixed decomposition as $\mathcal{M} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^d$. Let T be a densely defined linear operator on $\mathcal{M}$. For $z = (z^1, \dots, z^d)$ in $\mathbb{C}^d$, define the μ-resolvent of T at z to be*

$$(T-z)^{-1} \;=\; (T - [z^1 I_{\mathcal{M}^1} \oplus \cdots \oplus z^d I_{\mathcal{M}^d}]\,)^{-1}.$$

The μ-spectrum of T is the complement of the set of points in $\mathbb{C}^d$ for which the μ-resolvent is bounded.

The expressions "μ-resolvent" and "μ-spectrum" are not standard, but they are suggested by usage in control engineering. The notion of μ-analysis provides an approach to robust stabilization in the presence of "structured uncertainty" [11]. Corresponding to the projections $P^1, \dots, P^d$ one defines the "cost function" $\mu(X)$ by

$$\frac{1}{\mu(X)} = \inf\{\|T\| : T \in B(\mathcal{M}), \quad \text{each } P^r\mathcal{M} \text{ reduces } T \text{ and } 1 - TX \text{ is singular }\}.$$

In what follows, we shall write z for $z^1 I_{\mathcal{M}^1} \oplus \cdots \oplus z^2 I_{\mathcal{M}^d}$ and z^* for $\bar{z}^1 I_{\mathcal{M}^1} \oplus \cdots \oplus \bar{z}^d I_{\mathcal{M}^d}$. Let us recall Definition 1.19.

Definition 6.7. *For each real number t, define*

$$\rho_t(z) = \frac{z+t}{1-tz}. \tag{6.8}$$

For $F \in \mathcal{L}^d$, define

$$F_t := \rho_t \circ F \circ \rho_t.$$

Note that, similarly to the maps α and β, we use ρ_t on $\mathbb{C}^d$ to mean the component-wise action.

Theorem 6.9. *Let F be in $\mathcal{L}^d$, and z_0 a point in Π^d. For all except at most a countable number of real numbers t, there is a Hilbert space $\mathcal{M} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^d$, a self-adjoint operator X on $\mathcal{M}$, a vector v in $\mathcal{M}$, and a real number c so that*

$$F_t(z) \ = \ c + \langle zv, v\rangle + \langle (z - z_0^*)(X-z)^{-1}(z-z_0)v, v\rangle. \tag{6.10}$$

Conversely, if z_0 is a point in Π^d, c is a real number, X is a densely defined self-adjoint operator on a Hilbert space $\mathcal{M} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^d$, and v is a vector in $\mathcal{M}$, then the function of z given by the right-hand side of (6.10) is in $\mathcal{L}^d$.

Proof: ($\Rightarrow$) Let $\varphi = \beta \circ F \circ \alpha$ in $\mathcal{A}^d$. Choose a model for φ so that (3.2) holds:

$$\begin{aligned} 1 - \overline{\varphi(\mu)}\varphi(\lambda) \quad &= \quad \sum_{r=1}^{d} (1 - \overline{\mu^r}\lambda^r)\langle u_\lambda^r, u_\mu^r\rangle_{\mathcal{M}^r} \\ &= \quad \langle (1-\mu^*\lambda)u_\lambda, u_\mu\rangle_{\mathcal{M}}. \end{aligned} \tag{6.11}$$

Define a linear operator V by

$$V \ : \begin{pmatrix} 1 \\ \lambda u_\lambda \end{pmatrix} \mapsto \begin{pmatrix} \varphi(\lambda) \\ u_\lambda \end{pmatrix},$$

and extend it by linearity to finite linear combinations of vectors of the form

$$\begin{pmatrix} 1 \\ \lambda_i u_{\lambda_i} \end{pmatrix}$$

where the points λ_i range over $\mathbb{D}^d$.

V is defined on a subspace of $\mathbb{C} \oplus \mathcal{M}$, and by (6.11) it is isometric on its domain. If the codimensions of the closures of the domain and range of V are the same, V can be extended to a unitary U. If they are different, after the addition of a separable infinite dimensional summand to one of the spaces $\mathcal{M}^r$, the codimensions become equal, and one can then extend V to a unitary U. So we can assume that we have a unitary $U : \mathbb{C} \oplus \mathcal{M} \to \mathbb{C} \oplus \mathcal{M}$ such that

$$U \ : \ \begin{pmatrix} 1 \\ \lambda u_\lambda \end{pmatrix} \ \mapsto \ \begin{pmatrix} \varphi(\lambda) \\ u_\lambda \end{pmatrix}. \tag{6.12}$$

Now, let τ be a point in the unit circle that is *not* in the point spectrum of U, and let

$$t \ = \ -i\frac{1-\tau}{1+\tau}.$$

As $\mathbb{C} \oplus \mathcal{M}$ is separable, the point spectrum of U is countable, so all but countably many real numbers t will arise in this way.

Let

$$Y \ = \ -i(U-\tau)^{-1}(U+\tau) \ : \ (U-\tau)\eta \mapsto \ -i(U+\tau)\eta.$$

Then Y is densely defined and self-adjoint. Its domain $\mathcal{D}$ is ran $(U-\tau)$. Moreover, by definition

$$Y \ : \ \begin{pmatrix} \varphi(\lambda)-\tau \\ (1-\tau\lambda)u_\lambda \end{pmatrix} \ \mapsto \ \begin{pmatrix} -\,i(\varphi(\lambda)+\tau) \\ -\,i(1+\tau\lambda)u_\lambda \end{pmatrix}.$$

Therefore

$$Y \ : \ \begin{pmatrix} 1 \\ \\ \dfrac{1-\tau\lambda}{\varphi(\lambda)-\tau}u_\lambda \end{pmatrix} \ \mapsto \ \begin{pmatrix} -\,i\ \dfrac{\varphi(\lambda)+\tau}{\varphi(\lambda)-\tau} \\ \\ -\,i\ \dfrac{1+\tau\lambda}{\varphi(\lambda)-\tau}u_\lambda \end{pmatrix}. \tag{6.13}$$

Let

$$\tilde{v}_z \ = \ \frac{1-\tau\beta(z)}{\beta\circ F(z)-\tau}u_{\beta(z)}.$$

Then one can rewrite (6.13) as

$$Y \ : \ \begin{pmatrix} 1 \\ \tilde{v}_z \end{pmatrix} \ \mapsto \ \begin{pmatrix} \rho_t\circ F(z) \\ -\rho_{-t}(z)\tilde{v}_z \end{pmatrix}. \tag{6.14}$$

Now let

$$v_z \ = \ \tilde{v}_{\rho_t(z)}.$$

Then (6.14) becomes

$$Y : \begin{pmatrix} 1 \\ v_z \end{pmatrix} \mapsto \begin{pmatrix} F_t(z) \\ -zv_z \end{pmatrix}. \tag{6.15}$$

As Y is self-adjoint, (6.15) implies that

$$F_t(z) - F_t(w)^* \; = \; \langle (z - w^*)v_z, v_w \rangle. \tag{6.16}$$

Let X be the compression of $-Y$ to $\mathcal{M}$. By Lemma 6.24, X is self-adjoint with dense domain equal to $\mathcal{D} \cap \mathcal{M}$.

If γ is in $\mathcal{D} \cap \mathcal{M}$, then

$$Y \begin{pmatrix} 0 \\ \gamma \end{pmatrix} \; = \; \begin{pmatrix} L(\gamma) \\ -X\gamma \end{pmatrix}.$$

for some linear functional L.

Define $v = v_{z_0}$, and let $a = F_t(z_0)$. Then

$$Y \begin{pmatrix} 1 \\ v \end{pmatrix} \; = \; \begin{pmatrix} a \\ -z_0 v \end{pmatrix}.$$

For $z \in \Pi^d$ let

$$\begin{aligned} Y \begin{pmatrix} 1 \\ v_z \end{pmatrix} \; &= \; Y \begin{pmatrix} 1 \\ v \end{pmatrix} + Y \begin{pmatrix} 0 \\ v_z - v \end{pmatrix} \\ &= \; \begin{pmatrix} a \\ -z_0 v \end{pmatrix} + \begin{pmatrix} L(v_z - v) \\ -X(v_z - v) \end{pmatrix}. \end{aligned}$$

By (6.15), we get the equations

$$F_t(z) \; = \; a + L(v_z - v) \tag{6.17}$$
$$zv_z \; = \; z_0 v + X(v_z - v). \tag{6.18}$$

For $\gamma \in \mathcal{D} \cap \mathcal{M}$,

$$\begin{aligned} \langle \begin{pmatrix} 0 \\ \gamma \end{pmatrix}, \begin{pmatrix} a \\ -z_0 v \end{pmatrix} \rangle \; &= \; \langle \begin{pmatrix} 0 \\ \gamma \end{pmatrix}, Y \begin{pmatrix} 1 \\ v \end{pmatrix} \rangle \\ &= \; \langle Y \begin{pmatrix} 0 \\ \gamma \end{pmatrix}, \begin{pmatrix} 1 \\ v \end{pmatrix} \rangle \\ &= \; \langle \begin{pmatrix} L(\gamma) \\ -X\gamma \end{pmatrix}, \begin{pmatrix} 1 \\ v \end{pmatrix} \rangle. \end{aligned}$$

Therefore

$$L(\gamma) \;=\; -\langle \gamma, z_0 v\rangle + \langle X\gamma, v\rangle. \tag{6.19}$$

If z is in Π^d, then by Lemma 6.25, $X - z$ is invertible, so (6.18) yields

$$v_z - v \;=\; (X - z)^{-1}(z - z_0)v. \tag{6.20}$$

Combining equations (6.17) to (6.20), we get

$$\begin{aligned} F_t(z) &= a - \langle v_z - v, z_0 v\rangle + \langle z v_z - z_0 v, v\rangle \\ &= a - \langle v_z - v, z_0 v - z^* v\rangle + \langle (z - z_0)v, v\rangle \\ &= a + \langle (X - z)^{-1}(z - z_0)v, (z^* - z_0)v\rangle + \langle (z - z_0)v, v\rangle \end{aligned} \tag{6.21}$$

By (6.16),

$$a - \bar{a} \;=\; F_t(z_0) - \overline{F_t(z_0)} \;=\; \langle (z_0 - z_0^*)v, v\rangle,$$

so $c := a - \langle z_0 v, v\rangle$ is real. Then (6.21) becomes (6.10), as desired.

($\Leftarrow$) To prove the converse, suppose X is a self-adjoint operator on $\mathcal{M}$ with dense domain $\mathcal{D}'$. Let $F(z)$ be given by the right-hand side of (6.10). Define v_z by (6.20), *i.e.*

$$v_z \;=\; v + (X - z)^{-1}(z - z_0)v. \tag{6.22}$$

Define a linear functional L on $\mathcal{D}'$ by

$$L(\gamma) \;=\; -\langle \gamma, z_0 v\rangle + \langle X\gamma, v\rangle.$$

Let $\mathcal{D}$ be the linear span in $\mathbb{C} \oplus \mathcal{M}$ of the vector $\begin{pmatrix} 1 \\ v \end{pmatrix}$ and the vector space $0 \oplus \mathcal{D}'$. Let $a = c + \langle z_0 v, v\rangle$. Finally, define Y on $\mathcal{D}$ by

$$Y \begin{pmatrix} t \\ \gamma + tv \end{pmatrix} \;=\; t \begin{pmatrix} a \\ -z_0 v \end{pmatrix} + \begin{pmatrix} L(\gamma) \\ -X(\gamma) \end{pmatrix}$$

for t in $\mathbb{C}$. It is routine to verify that Y is symmetric. Moreover, by (6.22), $(v_z - v)$ is in the domain of $X - z$, and therefore in $\mathcal{D}'$. So $\begin{pmatrix} 1 \\ v_z \end{pmatrix}$ is in $\mathcal{D}$ for every z in Π, and therefore

$$\langle Y \begin{pmatrix} 1 \\ v_z \end{pmatrix}, \begin{pmatrix} 1 \\ v_w \end{pmatrix}\rangle \;=\; \langle \begin{pmatrix} 1 \\ v_z \end{pmatrix}, Y \begin{pmatrix} 1 \\ v_w \end{pmatrix}\rangle. \tag{6.23}$$

Expanding (6.23) and rearranging, one gets

$$F(z) - \overline{F(w)} \;=\; \langle (z - w^*)u_z, u_w\rangle,$$

and hence F is in $\mathcal{L}^d$. $\square$

Lemma 6.24. *With notation as in the proof of the forward direction of Theorem 6.9, the domain of X is $\mathcal{D} \cap \mathcal{M}$. This domain is dense, and the operator X is self-adjoint.*

PROOF: Since $\mathcal{D}$ is dense in $\mathbb{C} \oplus \mathcal{M}$, there are vectors ξ_n in $\mathcal{M}$ that converge to zero and such that $\begin{pmatrix} 1 \\ \xi_n \end{pmatrix}$ are in $\mathcal{D}$. If γ is any vector in $\mathcal{M}$, there are vectors $\begin{pmatrix} a_n \\ \eta_n \end{pmatrix}$ in $\mathcal{D}$ that converge to $\begin{pmatrix} 0 \\ \gamma \end{pmatrix}$, hence so do the vectors

$$\begin{pmatrix} a_n \\ \eta_n \end{pmatrix} - a_n \begin{pmatrix} 1 \\ \xi_n \end{pmatrix} = \begin{pmatrix} 0 \\ \eta_n - a_n \xi_n \end{pmatrix}.$$

Therefore $\mathcal{D} \cap \mathcal{M}$ is dense in $\mathcal{M}$.

Let P be the projection from $\mathbb{C} \oplus \mathcal{M}$ onto $\mathcal{M}$. Let $X = -PY|_{\mathcal{M}}$ with domain $\mathcal{D}' = \mathcal{D} \cap \mathcal{M}$. Then for $\gamma,\ \eta$ in $\mathcal{D}'$, we have

$$\begin{aligned} \langle X\gamma, \eta \rangle &= -\langle PY\gamma, \eta \rangle \\ &= -\langle Y\gamma, \eta \rangle \\ &= -\langle \gamma, Y\eta \rangle \\ &= \langle \gamma, X\eta \rangle. \end{aligned}$$

So X is symmetric.

To prove X is self-adjoint, assume that there is some vector η in $\mathcal{M}$ such that

$$|\langle X\gamma, \eta \rangle| \ \leq \ C\|\gamma\|$$

for all $\gamma \in \mathcal{D}'$. Then for every vector $\begin{pmatrix} c \\ \delta \end{pmatrix}$ in $\mathcal{D}$ of norm at most one, we have

$$\begin{aligned} |\langle Y \begin{pmatrix} c \\ \delta \end{pmatrix}, \begin{pmatrix} 0 \\ \eta \end{pmatrix} \rangle| &= |\langle Y \left[\begin{pmatrix} 0 \\ \delta - c\xi_1 \end{pmatrix} + c \begin{pmatrix} 1 \\ \xi_1 \end{pmatrix} \right], \begin{pmatrix} 0 \\ \eta \end{pmatrix} \rangle| \\ &\leq \ C\|\delta - \xi_1\| + |c| \, \|Y \begin{pmatrix} 1 \\ \xi_1 \end{pmatrix} \| \, \|\eta\| \ \leq \ C'. \end{aligned}$$

So $\begin{pmatrix} 0 \\ \eta \end{pmatrix}$ is in $\mathcal{D}$, and therefore η is in $\mathcal{D}'$. $\Box$

Lemma 6.25. *Let X be a densely defined self-adjoint operator on $\mathcal{M}$. The μ-spectrum of X is disjoint from $\Pi^d \cup \Pi^{d*}$. Moreover,*

$$\|(X - z)^{-1}\| \ \leq \ 1/ \min_{1 \leq r \leq d} (|\mathrm{Im}\, z^r|) \qquad \forall\, z \in \Pi^d \cup \Pi^{d*}.$$

Proof: Let X be self-adjoint on $\mathcal{M} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^d$, and let $z = (x^1 + iy^1, \ldots, x^d + iy^d)$ be a point in $\Pi^d \cup \Pi^{d*}$. Then for any $v = v^1 \oplus \cdots \oplus v^d$ in $\mathcal{M}$,

$$\langle (X-z)v, v\rangle \;=\; \langle (X-x)v, v\rangle - i(y^1\|v^1\|^2 + \cdots + y^d\|v^d\|^2).$$

The first summand on the right is real, so $X - z$ is bounded below by $\min(|y^r|)$, and therefore has a left inverse. Applying the same argument to z^*, we get that $X - z^*$ has a left inverse, and taking adjoints we get that $X - z$ has a right inverse also. $\Box$

When F decays at infinity, we can sharpen Theorem 6.9 to get a theorem like Nevanlinna's in 6.3.

Let us write $\mathbb{1}$ for $(1, 1, \ldots, 1)$, and $s\mathbb{1}$ for $(s, s, \ldots, s)$, etc.

Theorem 6.26. *Let F be in $\mathcal{L}^d$, and assume F has a representation as in (6.10) with $t = 0$. Then the following are equivalent.*

(i)
$$\liminf_{y\to\infty} \; y|F(iy\mathbb{1})| \;<\; \infty.$$

(ii) *There exists a vector v_1 in $\mathcal{M}$ so that*
$$F(z) \;=\; \langle (X-z)^{-1}v_1, v_1\rangle \qquad z \;\in\; \Pi^d. \tag{6.27}$$

(iii) *The function $\varphi = \beta \circ F \circ \alpha$ in $\mathcal{S}$ has a B-point at $\mathbb{1}$ and $\varphi(\mathbb{1}) = -1$.*

(iv) $\lim_{y\to\infty} F(iy\mathbb{1}) = 0$ *and*
$$\liminf_{y\to\infty} \; y\|v_{(iy\mathbb{1})}\| \;<\; \infty.$$

(v) *The vector v is in the domain of X and* $\lim_{y\to\infty} F(iy\mathbb{1}) = 0$.

Proof: (i) $\Rightarrow$ (iii) Let $\varphi = \beta \circ F \circ \alpha$. Condition (i) becomes
$$\liminf_{s\to 1} \; \frac{1+s}{1-s}\left|\frac{1+\varphi(s\mathbb{1})}{1-\varphi(s\mathbb{1})}\right| \;<\; \infty. \tag{6.28}$$

The left-hand side of (6.28) dominates
$$\frac{1-|\varphi(s\mathbb{1})|}{1-s},$$
so $\mathbb{1}$ is a B-point. In order for (6.28) to hold, we must have $\varphi(\mathbb{1}) = -1$.

(iii) $\Leftrightarrow$ (iv) As the proof of Proposition 2.11 shows, one can pass between a model $(\mathcal{M}, u)$ for φ and a model $(\mathcal{M}, v)$ for F by letting

$$\begin{aligned} v_z^r &= \left(\frac{F(z)+i}{z^r+i}\right) u^r_{\beta(z)} \qquad\qquad (6.29)\\ u_\lambda^r &= \left(\frac{1-\varphi(\lambda)}{1-\lambda^r}\right) v^r_{\alpha(\lambda)} \qquad r = 1, \ldots, d. \end{aligned}$$

By Lemma 3.10, φ having a B-point at $\mathbb{1}$ is equivalent to $u_{(r\mathbb{1})}$ being bounded as $r \to 1^-$. Moreover, $\varphi(r\mathbb{1})$ tending to -1 is the same as $F(iy\mathbb{1})$ tending to 0 as $y \to \infty$. And as long as $F(iy\mathbb{1})$ has a finite limit, (6.29) says that $u_{(r\mathbb{1})}$ is bounded iff $[y\, v_{(iy\mathbb{1})}]$ is.

(iv) $\Rightarrow$ (v) As X is densely defined and self-adjoint, it is closed. By (6.18), the vectors $v_z - v$ all lie in $\mathcal{D}'$, the domain of X. Let $z = (iy\mathbb{1})$ and let $y \to \infty$. Then $v - v_{(iy\mathbb{1})}$ tends to v. Moreover, $X(v - v_{(iy\mathbb{1})}) = z_0 v - iyv_{(iy\mathbb{1})}$ contains a bounded sequence as $y \to \infty$, and therefore a subsequence that converges weakly to some vector, w say. So (v, w) is in the weak closure of the graph of X, therefore in the graph of X, and hence v is in $\mathcal{D}'$.

Therefore $y_n\, v_{z_n}$ also converges in norm, and hence

(v) $\Rightarrow$ (ii) If $v \in \mathcal{D}'$, then (6.19) becomes

$$L(\gamma) = \langle \gamma, Xv - z_0 v\rangle. \qquad (6.30)$$

Let $v_1 = (X - z_0)v$. Then (6.18) says

$$(X - z)v_z = v_1. \qquad (6.31)$$

Combining (6.17), (6.30) and (6.31), we get

$$F_t(z) \;=\; a - \langle v, v_1\rangle + \langle (X-z)^{-1} v_1, v_1\rangle. \qquad (6.32)$$

Now let $z = (iy\mathbb{1})$ in (6.32) and let $y \to \infty$. By Lemma 6.25, the last term on the right tends to zero, so we must have $a - \langle v, v_1\rangle = 0$.

(ii) $\Rightarrow$ (i) Lemma 6.25 implies that

$$\|F(z)\| \;\leq\; \|v_1\|^2 / \min(|\mathrm{Im}\, z^r|),$$

and so (i) follows. $\square$

For later use, let us record a slight variant of Theorem 6.26; it is proved in the same way.

Theorem 6.33. *Let F be in $\mathcal{L}^d$, and assume F has a representation as in (6.10) with $t = 0$. Then the following are equivalent.*

(i) *There exists a constant $C \in \mathbb{R}$ so that*

$$\liminf_{y\to\infty} \ y|F(iy\mathbb{1}) \ - \ C| \ < \ \infty.$$

(ii) *There exists a vector v_1 in $\mathcal{M}$ and a constant C in $\mathbb{R}$ so that*

$$F(z) \ = \ C \, + \, \langle (X-z)^{-1} v_1, v_1 \rangle \qquad z \ \in \ \Pi^d.$$

(iii) *The function $\varphi = \beta \circ F \circ \alpha$ in $\mathcal{S}$ has a B-point at $\mathbb{1}$ and $\varphi(\mathbb{1}) \neq 1$.*

(iv) $\lim_{y\to\infty} F(iy\mathbb{1}) = C \ \in \ \mathbb{R}$ *and*

$$\liminf_{y\to\infty} \ y\|v_{(iy\mathbb{1})}\| \ < \ \infty.$$

(v) *The vector v is in the domain of X.*

7 Locally matrix monotone functions

Recall the definition of locally n-matrix monotone.

Definition 1.9 *Let E be an open set in $\mathbb{R}^d$, and f be a real-valued C^1 function on E. Say f is locally M_n-monotone on E if, whenever S is in $CSAM_n^d$ with $\sigma(S) = \{x_1, \dots, x_n\}$ consisting of n distinct points in E, and $S(t)$ is a C^1 curve in $CSAM_n^d$ with $S(0) = S$ and $\frac{d}{dt} S(t)|_{t=0} \geq 0$, then $\frac{d}{dt} f(S(t))|_{t=0}$ exists and is ≥ 0.*

If S is in $CSAM_n^d$, we can choose an orthonormal basis of eigenvectors that diagonalize all the S^r's simultaneously, so

$$S^r \ = \ \begin{pmatrix} x_1^r & & \\ & \ddots & \\ & & x_n^r \end{pmatrix} \qquad \forall\ 1 \leq r \leq d. \tag{7.1}$$

If $S(t)$ is a C^1 curve of commuting self-adjoints, then $S(0) + t\,\frac{d}{dt} f(S(t))|_{t=0}$ commutes to first order.

For any $X \in M_n$ we define diag X to be the diagonal matrix in M_n with diagonal entries X_{ii}, and for any $\Delta \in SAM_n^d$ we define diag Δ to be (diag $\Delta^1, \dots,$ diag Δ^d).

Definition 7.2. *We shall say that S in $CSAM_n^d$ is generic if its spectrum consists of n distinct points.*

Lemma 7.3. *Let S be in $CSAM_n^d$ and Δ be in SAM_n^d, with S generic. Then there exists a C^1 curve $S(t)$ of commuting self-adjoints with $S(0) = S$ and $S'(0) = \Delta$ if and only if*

$$[S^r, \Delta^s] \ = \ [S^s, \Delta^r] \qquad \forall\ 1 \leq r \neq s \leq d. \tag{7.4}$$

Proof: ($\Rightarrow$): If $S(t) = S + t\Delta + o(t)$ is commutative, calculate

$$[S^r(t), S^s(t)] \ = \ t\,([S^r, \Delta^s] - [S^s, \Delta^r]) + o(t).$$

The coefficient of t must vanish, giving (7.4).

($\Leftarrow$): Suppose S is as in (7.1), and (7.4) holds. This means

$$\Delta_{ij}^s(x_j^r - x_i^r) \ = \ \Delta_{ij}^r(x_j^s - x_i^s) \qquad \forall\, r \neq s, \tag{7.5}$$

so

$$\Delta_{ij}^r \frac{1}{x_j^r - x_i^r} \ = \ \Delta_{ij}^s \frac{1}{x_j^s - x_i^s} \qquad \text{if } x_j^r - x_i^r \neq 0 \neq x_j^s - x_i^s. \tag{7.6}$$

Define a skew-symmetric matrix Y by

$$Y_{ij} \ = \ \Delta_{ij}^r \frac{1}{x_j^r - x_i^r} \qquad \text{for any } r \text{ such that } x_j^r - x_i^r \neq 0. \tag{7.7}$$

For any $i \neq j$, there is some r with $x_j^r - x_i^r \neq 0$, so (7.7) defines Y_{ij}; and (7.6) says it doesn't matter which r we choose. Let all the diagonal terms of Y be 0.

Define

$$S^r(t) \ = \ e^{tY}(S^r + t \text{ diag } \Delta^r)e^{-tY}. \tag{7.8}$$

Since e^{tY} is a unitary matrix and $S^r + t$ diag Δ^r is diagonal, $S(t) \in CSAM_n^d$ and

$$\frac{d}{dt}S^r(t)|_{t=0} = [Y, S^r] + \text{diag } \Delta^r = \Delta^r.$$

□

If S and Δ satisfy (7.4) and S is generic then for any function f that is C^1 on a neighborhood of $\sigma(S)$ we define the directional derivative of f at S in direction Δ by

$$D_\Delta f(S) = \frac{d}{dt}f(S(t))|_{t=0} \tag{7.9}$$

where $S(t)$ is the curve given by equations (7.8) and (7.7). We shall show in Proposition 7.18 that (7.9) is actually unchanged if $S(t)$ is replaced by any other curve that agrees with it to first order. First, let us show that the right-hand side of (7.9) exists. Indeed,

$$f(S(t)) = e^{tY} f(S + t \text{ diag } \Delta) e^{-tY}. \tag{7.10}$$

Since $S + t$ diag Δ is diagonal, $f(S + t \text{ diag } \Delta)$ is diagonal, with ith entry

$$f(x_i + t\Delta_{ii}) = f(x_i) + t \sum_{r=1}^{d} \Delta_{ii}^r \frac{\partial f}{\partial x^r}(x_i) + o(t).$$

In other words,

$$f(S + t \text{ diag } \Delta) = f(S) + t \sum_{r=1}^{d} (\text{diag } \Delta^r) \frac{\partial f}{\partial x^r}(S) + o(t).$$

Hence, on differentiating equation (7.10) at 0 we obtain

$$\frac{d}{dt} f(S(t))|_{t=0} = [Y, f(S)] + \sum_{r=1}^{d} (\text{diag } \Delta^r) \frac{\partial f}{\partial x^r}(S).$$

We have shown the following.

Proposition 7.11. *Let S be a generic d-tuple of commuting self-adjoint matrices in M_n. Fix an orthonormal basis of eigenvectors, so every S^r is diagonal:*

$$S^r = \begin{pmatrix} x_1^r & & \\ & \ddots & \\ & & x_n^r \end{pmatrix}.$$

Let Δ be a d-tuple of self-adjoints satisfying (7.4). Let f be C^1 on a neighborhood of $\{x_1, \ldots, x_n\}$ in $\mathbb{R}^d$, where each x_j is the d-tuple $(x_j^1, \ldots, x_j^d)$. Then

$$[D_\Delta f(S)]_{ij} = \begin{cases} \Delta_{ij}^r \frac{f(x_j)-f(x_i)}{x_j^r - x_i^r} & \text{if } i \neq j, \text{ where } x_j^r \neq x_i^r \\ \\ \sum_{r=1}^{d} \Delta_{ii}^r \frac{\partial f}{\partial x^r}|_{x_i} & \text{if } i = j. \end{cases} \tag{7.12}$$

Corollary 7.13. *For S, Δ as in Proposition 7.11, if f, g are C^1 functions that agree to first order on $\sigma(S)$, then $D_\Delta f(S) = D_\Delta g(S)$.*

Lemma 7.14. *Let R and S be in $CSAM_n^d$. For every point μ in the joint spectrum of R there is an x_p in the joint spectrum of S with*

$$\|\mu - x_p\| \ \leq \ \sqrt{dn}\|R - S\|. \tag{7.15}$$

Proof. Choose an orthonormal basis that diagonalizes S, so that S is as in (7.1). Let $\Delta = R - S$. Let μ be a joint eigenvalue of R with corresponding eigenvector $\xi = (\xi_1, \ldots, \xi_n)^t$. Choose p so that $|\xi_p| \geq |\xi_j|$ for all $1 \leq j \leq n$.

Then for each $1 \leq r \leq d$, we have

$$R^r \xi \ = \ \mu^r \xi \ = \ (S^r + \Delta^r)\xi.$$

So in particular,

$$\sum_{j=1}^n R^r_{pj} \xi_j \ = \ \mu^r \xi_p.$$

Therefore

$$(\mu^r - x_p^r)\xi_p \ = \ \sum_{j=1}^n \Delta^r_{pj} \xi_j.$$

So

$$\begin{aligned} |\mu^r - x_p^r| \quad &\leq \quad \sum_{j=1}^n |\Delta^r_{pj}| \\ &\leq \quad \sqrt{n} \sqrt{\sum_j |\Delta^r_{pj}|^2} \\ &\leq \quad \sqrt{n} \|\Delta^r\|, \end{aligned}$$

and hence

$$\sum_{r=1}^d |\mu^r - x_p^r|^2 \ \leq \ dn\|\Delta\|^2.$$

$\square$

Lemma 7.16. *If $R(t)$ is a Lipschitz path in $CSAM_n^d$, $0 \leq t < 1$, with $R(0) = S$ generic then there exists $\varepsilon > 0$ and Lipschitz maps $X_1, \ldots, X_n : [0, \varepsilon) \to \mathbb{R}^d$ such that $\sigma(R(t)) = \{X_j(t) : j = 1, \ldots, n\}$.*

Proof. Choose an orthonormal basis that diagonalizes S, so S is as in (7.1). The joint eigenvalues of S are the points $x_i = (x_i^1, \ldots, x_i^d)$, and genericity means $\|x_i - x_j\| > 0$ if $i \neq j$. Choose ε so that for all $0 \leq t \leq \varepsilon$,

$$\sqrt{dn}\|R(t) - S\| \;\leq\; \frac{1}{3} \min_{i \neq j} \|x_i - x_j\|. \tag{7.17}$$

By Lemma 7.14, for every joint eigenvalue x of S there is a joint eigenvalue μ of $R(t)$ within $\sqrt{dn}\|R(t) - S\|$ of it. By (7.17), this means that $R(t)$ is also generic, and it makes sense to talk of the joint eigenvalue of $R(t)$ that is closest to x_j. Let us call these joint eigenvalues $X_j(t)$. We have proved that

$$\|X_j(t) - x_j\| \;\leq\; \sqrt{dn}\|R(t) - S\| \qquad \forall\ 0 \leq t \leq \varepsilon.$$

Repeating the argument with $R(t_1)$ in place of S, we get

$$\|X_j(t_2) - X_j(t_1)\| \;\leq\; \sqrt{dn}\|R(t_2) - R(t_1)\| \qquad \forall\ 0 \leq t_1, t_2 \leq \varepsilon.$$

As R is assumed to be Lipschitz, we get that each X_j is Lipschitz also. □

Proposition 7.18. *If S is generic in $CSAM_n^d$, Δ is in SAM_d^n, and they satisfy the commutation relations (7.4), then for any C^1 path $R(t) \in CSAM_n^d$ such that $R(0) = S$, $R'(0) = \Delta$ and any $f \in C^1$,*

$$\frac{d}{dt} f(R(t))|_{t=0} = D_\Delta f(S). \tag{7.19}$$

Proof. If g is a monomial then a simple calculation shows that

$$\frac{d}{dt} g(R(t))|_{t=0}$$

exists and depends only on g, S and Δ. It follows that, for any polynomial g,

$$\frac{d}{dt} g(R(t))|_{t=0} = \frac{d}{dt} g(S(t))|_{t=0} = D_\Delta g(S). \tag{7.20}$$

Consider any $f \in C^1$ and pick a polynomial g that agrees with f to first order on $\sigma(S)$. By Corollary 7.13,

$$D_\Delta f(S) = D_\Delta g(S). \tag{7.21}$$

We claim that

$$\frac{d}{dt} g(R(t))|_{t=0} = \frac{d}{dt} f(R(t))|_{t=0}. \tag{7.22}$$

For by Lemma 7.16 there exist Lipschitz functions $X_1, \dots, X_n : [0, \varepsilon) \to \mathbb{R}^d$ such that $\sigma(R(t)) = \{X_1(t), \dots, X_n(t)\}$ for all t. Then $f(S) = g(S)$ and

$$\begin{aligned} \|(f-g)(R(t))\| &= \max_i |(f-g)(X_i(t))| \\ &= o(\max_i |X_i(t) - X_i(0)|) \\ &= o(t). \end{aligned}$$

Hence

$$\left\| \frac{f(R(t)) - f(S)}{t} - \frac{g(R(t)) - g(S)}{t} \right\| \to 0 \text{ as } t \to 0.$$

In view of equation (7.20),

$$\frac{f(R(t)) - f(S)}{t} \to \frac{d}{dt} g(R(t))|_{t=0} = D_\Delta g(S) \text{ as } t \to 0.$$

On combining this relation with equation (7.21) we obtain equation (7.19). ☐

Corollary 7.23. *A real-valued C^1 function f on an open set $E \subseteq \mathbb{R}^d$ is locally M_n-monotone if and only if*

$$D_\Delta f(S) \ \geq \ 0$$

for every generic S in $CSAM_n^d$ with spectrum in E and every Δ in SAM_n^d such that $\Delta \geq 0$ and

$$[S^r, \Delta^s] \ = \ [S^s, \Delta^r] \qquad \forall\ 1 \leq r \neq s \leq d.$$

The statement follows immediately from Definition 1.9 and Proposition 7.18.

We can now characterize locally matrix monotone functions.

Theorem 7.24. *Let E be an open set in $\mathbb{R}^d$, and f a real-valued C^1 function on E. Then f is locally M_n-monotone if and only if f is in $\mathcal{L}_n(E)$.*

Proof: ($\Leftarrow$) We must show: if S is generic with $\sigma(S) \subset E$, if Δ is a positive d-tuple and $[S^r, \Delta^s] = [S^s, \Delta^r]$ for all r, s, then $D_\Delta f(S) \geq 0$.

Let $\sigma(S) = \{x_1, \ldots, x_n\}$. Choose A^r as in Definition 5.1. For $i \neq j$, assume without loss of generality that $x_j^1 \neq x_i^1$. Then

$$\begin{aligned} [D_\Delta f(S)]_{ij} &= \Delta^1_{ij} \frac{f(x_j) - f(x_i)}{x_j^1 - x_i^1} \\ &= \frac{\Delta^1_{ij}}{x_j^1 - x_i^1} \left(\sum_{r=1}^d (x_j^r - x_i^r) A^r(i,j) \right) \\ &= \sum_{r=1}^d \Delta^r_{ij} \, A^r(i,j). \end{aligned}$$

(We get the last line by using (7.6)). By (7.12) the same formula holds for $[D_\Delta f(S)]_{ij}$ when $i = j$, so $D_\Delta f(S)$ is the sum of the Schur products of Δ^r with A^r, so is positive.

($\Rightarrow$) Let f be locally M_n-monotone, and fix $\{x_1, \ldots, x_n\}$ distinct points in E. Let S be given by (7.1). We wish to find positive matrices A^r such that (5.3) and (5.2) hold.

Let $\mathcal{G}$ be the set of all skew-symmetric real n-by-n matrices Γ with the property that there exists a d-tuple A of real positive semi-definite matrices satisfying

$$A^r(i,i) = \left.\frac{\partial f}{\partial x^r}\right|_{x_i} \qquad 1 \leq i \leq n, \ 1 \leq r \leq d \tag{7.25}$$

$$\sum_{r=1}^d (x_j^r - x_i^r) \, A^r(i,j) = \Gamma_{ij} \qquad 1 \leq i \neq j \leq n. \tag{7.26}$$

Let Λ be the matrix $\Lambda_{ij} = f(x_i) - f(x_j)$. We wish to show Λ is in $\mathcal{G}$.

Notice that $\mathcal{G}$ is a closed convex set. Moreover, it is non-empty, because $\left.\frac{\partial f}{\partial x^r}\right|_{x_i}$ is always greater than or equal to 0. (This last assertion can be seen by letting Δ be 0 except in the r^{th} slot, where it is I, and calculating $D_\Delta f(S)$).

So if Λ is not in $\mathcal{G}$, by the Hahn-Banach theorem there is a real linear functional $L : M_n \to \mathbb{R}$ that is non-negative on $\mathcal{G}$ and negative on Λ. Any such linear functional is of the form $L(T) = tr(TK)$ for some matrix K. Replacing K by $K - K^t$ will not change the value of L on skew-symmetric real matrices, so we can assume that there is a real skew-symmetric matrix K such that $tr(\Gamma K) \geq 0$ for all Γ in $\mathcal{G}$, and $tr(\Lambda K) < 0$.

Define Δ by

$$\Delta^r_{ij} \;=\; (x^r_j - x^r_i)K_{ji}.$$

Then Δ is in SAM^d_n, and

$$[\Delta^s, S^r]_{ij} \;=\; (x^s_j - x^s_i)K_{ji}(x^r_j - x^r_i) \;=\; [\Delta^r, S^s],$$

so Δ satisfies (7.4).

Moreover $\Delta \geq 0$. Indeed, fix s between 1 and d, and let $c_1, \dots, c_n$ be complex numbers. We want to show that

$$\sum_{i,j=1}^{n} c_i \bar{c}_j \Delta^s_{ij} \;\geq\; 0. \tag{7.27}$$

For $r \neq s$, let A^r be the diagonal matrix with entries given by (7.25). Define A^s to be the sum of the diagonal matrix from (7.25) with the rank one matrix $[c_i \bar{c}_j]$. Define Γ by (7.26). Then Γ is in $\mathcal{G}$, and since K and Δ^s both vanish on the diagonal,

$$\begin{aligned} tr(\Gamma K) \;&=\; \sum_{i,j=1}^{n} (x^s_j - x^s_i) A^s(i,j) K_{ji} \\ &=\; \sum_{i,j} \Delta^s_{ij} c_i \bar{c}_j \\ &\geq\; 0, \end{aligned}$$

yielding (7.27).

As f is locally M_n monotone, we must have then that $D_\Delta f(S) \geq 0$. But

$$\begin{aligned} 0 \;>\; tr(\Lambda K) \;&=\; \sum_{1 \leq i \neq j \leq n} [f(x_j) - f(x_i)] K_{ji} \\ &=\; \sum_{1 \leq i,j \leq n} \frac{f(x_j) - f(x_i)}{x^r_j - x^r_i} \Delta^r_{ij} \\ &=\; \sum_{1 \leq i,j \leq n} [D_\Delta f(s)]_{ij}, \end{aligned}$$

which is a contradiction. $\square$

As the dimension of the matrices increases, the condition that a function f be locally monotone becomes more stringent. On an infinite dimensional Hilbert space, the requirement becomes that f be in the Loewner class, as we shall see in the next section.

8 Locally operator monotone functions

We defined locally operator monotone functions in Definition 1.8. We shall show that being locally operator monotone is the same as being locally M_n-monotone for all n, which in turn is the same as being in the Löwner class $\mathcal{L}(E)$.

Theorem 8.1. *Let E be an open set in $\mathbb{R}^d$, and f a real-valued C^1 function on E. The following are equivalent.*
(i) The function f is locally M_n-monotone on E for all $n \geq 1$.
(ii) The function f is in $\mathcal{L}(E)$.
(iii) The function f is locally operator monotone on E.

The equivalence of (i) and (ii) follows from Theorem 7.24 and Proposition 5.11. The implication $(iii) \Rightarrow (i)$ is obvious. We need to prove that $(ii) \Rightarrow (iii)$. First we need some preliminary results.

Proposition 8.2. *Let E be an open set in $\mathbb{R}^d$, and let $f \in \mathcal{L}(E)$. Then there is a model $(\mathcal{M}, v)$ for f such that v_z is locally Lipschitz on E.*

Proof: By Proposition 5.11 we can extend f to a function F in $\mathcal{L}$ that extends analytically across E and agrees with f on E. For this F we have a model $(\mathcal{M}, v)$ so that

$$F(z) - \overline{F(w)} \; = \; \langle (z - w^*) v_z, v_w \rangle_{\mathcal{M}} \qquad \forall \; z, w \in E \cup \Pi^d, \tag{8.3}$$

and by Proposition 3.12, if w is in E, then v_w is the limit of v_z as z tends to w non-tangentially from inside Π^d.

Fix w in E (so $F(w)$ is real). Then, by analyticity, we have for z close to w :

$$F(z) - F(w) \; = \; \sum_{r=1}^{d} \left. \frac{\partial f}{\partial x^r} \right|_w (z^r - w^r) \; + \sum_{|\alpha| \geq 2} \left. \frac{\partial^\alpha f}{\partial x^\alpha} \right|_w \frac{(z-w)^\alpha}{\alpha!}. \tag{8.4}$$

From (8.3), we get

$$\begin{aligned} F(z) - F(w) \quad &= \quad \langle (z-w) v_z, v_w \rangle \\ &= \quad \sum_{r=1}^{d} (z^r - w^r) \langle v_w^r, v_w^r \rangle \; + \; \sum_{r=1}^{d} (z^r - w^r) \langle v_z^r - v_w^r, v_w^r \rangle. \end{aligned}$$

As z tends to w non-tangentially, the second term is $o(\|z-w\|)$, so comparing with (8.4) we see that

$$\|v_w^r\|^2 \;=\; \left.\frac{\partial f}{\partial x^r}\right|_w \qquad \forall\, 1 \leq r \leq d. \tag{8.5}$$

Now let z and w both be in E. Comparing (8.3) and (8.4), we get

$$\langle (z-w)v_z, v_w\rangle \;-\; \langle (z-w)v_w, v_w\rangle \;=\; \sum_{|\alpha|\geq 2} \left.\frac{\partial^\alpha f}{\partial x^\alpha}\right|_w \frac{(z-w)^\alpha}{\alpha!}. \tag{8.6}$$

Swapping z and w, we get

$$\langle (w-z)(v_w-v_z), v_z\rangle \;=\; \sum_{|\alpha|\geq 2} \left.\frac{\partial^\alpha f}{\partial x^\alpha}\right|_z \frac{(w-z)^\alpha}{\alpha!}. \tag{8.7}$$

Subtracting (8.6) from (8.7), we get

$$\langle (z-w)(v_z-v_w), (v_z-v_w)\rangle \;=\; \sum_{|\alpha|\geq 2} \left((-1)^{|\alpha|} \left.\frac{\partial^\alpha f}{\partial x^\alpha}\right|_z \;-\; \left.\frac{\partial^\alpha f}{\partial x^\alpha}\right|_w \right) \frac{(z-w)^\alpha}{\alpha!}. \tag{8.8}$$

But since f is analytic,

$$\left(\left.\frac{\partial^\alpha f}{\partial x^\alpha}\right|_z \;-\; \left.\frac{\partial^\alpha f}{\partial x^\alpha}\right|_w \right) \;=\; O(\|z-w\|),$$

and so the right-hand side of (8.8) is $O(\|z-w\|^3)$. Therefore

$$\langle (z-w)(v_z-v_w), (v_z-v_w)\rangle \;=\; O(\|z-w\|^3). \tag{8.9}$$

If all the differences $|z^r-w^r|$ were comparable, we could conclude immediately that $\|v_z-v_w\| = O(\|z-w\|)$. If they are not, we can get round this difficulty by connecting z to w by two line segments.

Indeed, suppose $\max_{1\leq r\leq d}|z^r-w^r| = \varepsilon$. Choose numbers a^r and b^r with modulus between $1/2$ and 2 so that

$$z^r - w^r \;=\; (a^r - b^r)\varepsilon \qquad \forall\; 1 \leq r \leq d.$$

Let

$$x^r \;=\; w^r + a^r\varepsilon \;=\; z^r + b^r\varepsilon.$$

Then applying (8.9) to the pairs (z,x) and (x,z), we get

$$\begin{aligned} \|v_z - v_w\| \quad &\leq \quad \|v_z - v_x\| + \|v_x - v_w\| \\ &= \quad O(\|z-x\| + \|x-w\|) \\ &= \quad O(\|z-w\|), \end{aligned}$$

as desired. $\square$

Suppose now E, f and $(\mathcal{M}, v)$ are as in Proposition 8.2. So $v : z \mapsto v_z$ is a map from E to $\mathcal{M}$. Let S be a d-tuple of bounded commuting self-adjoint operators on a Hilbert space $\mathcal{H}$, with $\sigma(S) \subset E$. We want to define an operator $\tilde{v}(S) \in B(\mathcal{H}, \mathcal{H} \otimes \mathcal{M})$.

We do this by choosing an orthonormal basis for $\mathcal{M}$, and writing

$$v(z) \ := \ v_z \ = \ \begin{pmatrix} v_1(z) \\ v_2(z) \\ \vdots \end{pmatrix}.$$

Then

$$\tilde{v}(S) \ := \ \begin{pmatrix} v_1(S) \\ v_2(S) \\ \vdots \end{pmatrix} \ : \mathcal{H} \to \mathcal{H} \otimes \mathcal{M}. \tag{8.10}$$

The operator $\tilde{v}(S)$ is bounded, because if S has spectral measure Λ and ξ is a unit vector in $\mathcal{H}$, then

$$\begin{aligned} \|\tilde{v}(S)\xi\|^2 \quad &= \quad \sum_j \int_{\sigma(S)} |v_j|^2 d\langle \Lambda\xi, \xi\rangle & (8.11) \\ &= \quad \int_{\sigma(S)} \sum_{r=1}^{d} \frac{\partial f}{\partial x^r} d\langle \Lambda\xi, \xi\rangle \\ &\leq \quad \sup_{z \in \sigma(S)} \ \sum_{r=1}^{d} \left.\frac{\partial f}{\partial x^r}\right|_z , & (8.12) \end{aligned}$$

and the last sum is finite because $\sigma(S)$ is compact and f is C^1.

The operator $\tilde{v}(S)$ does not depend on the choice of orthonormal basis in $\mathcal{M}$. A simple calculation shows that, for any $h \in \mathcal{H}$ and $m \in \mathcal{M}$,

$$\tilde{v}(S)^*(h \otimes m) = \langle m, v(.)\rangle(S)h.$$

This gives a coordinate-free expression for $\tilde{v}(S)^*$, hence also for $\tilde{v}(S)$.

Lemma 8.13. *Let E, f and $(\mathcal{M}, v)$ be as in Proposition 8.2, and S be a d-tuple of bounded commuting self-adjoint operators on a Hilbert space $\mathcal{H}$, with $\sigma(S) \subset E$. Then*

$$\|\tilde{v}(S)\| \ \leq \ \left(\sup_{\sigma(S)} \operatorname{div}\, f \right)^{\frac{1}{2}}. \tag{8.14}$$

Moreover, $\tilde{v}$ is continuous.

PROOF: Inequality (8.14) has been proved in (8.12). To prove continuity of $\tilde{v}$, let K be a compact subset of E with $\sigma(S) \subset int(K) \subset E$. Let $\varepsilon > 0$.

As v is continuous on K and K is compact, there exists N such that $\sum_{j=N+1}^{\infty} |v_j(z)|^2 \leq \varepsilon^2/9$ on K. For $1 \leq j \leq N$, there is a polynomial p_j such that $\|p_j - v_j\|_\infty \leq \varepsilon/9N$ on K. There exists $\delta > 0$ so that if $\|T^r - S^r\| \leq \delta$, then $\sigma(T) \subseteq K$ and $\|p_j(T) - p_j(S)\| \leq \varepsilon/9N$ for each $1 \leq j \leq N$.

Let $\tilde{v}_N(S)$ be the operator

$$\begin{pmatrix} v_1(S) \\ \vdots \\ v_N(S) \\ 0 \\ \vdots \end{pmatrix}.$$

Then $\|\tilde{v}_N(S) - \tilde{v}(S)\| \leq \varepsilon/3$ by (8.11), and similarly $\|\tilde{v}_N(T) - \tilde{v}(T)\| \leq \varepsilon/3$. As

$$\|v_j(T) - v_j(S)\| \ \leq \ \|v_j(T) - p_j(T)\| + \|p_j(T) - p_j(S)\| + \|p_j(S) - v_j(S)\|,$$

and each entry is at most $\varepsilon/9N$,

$$\|\tilde{v}_N(S) - \tilde{v}_N(T)\| \leq N\left(\frac{\varepsilon}{9N} + \frac{\varepsilon}{9N} + \frac{\varepsilon}{9N}\right) = \frac{\varepsilon}{3},$$

and hence

$$\|\tilde{v}(T) - \tilde{v}(S)\| \ \leq \ \varepsilon.$$

$\square$

We assume that $\mathcal{M}$ is decomposed as $\mathcal{M} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^d$, and P^r is the orthogonal projection from $\mathcal{M}$ onto $\mathcal{M}^r$. If $S = (S^1, \ldots, S^d)$ is a d-tuple of operators on $\mathcal{H}$, we shall write

$$S \odot I \ := \ S^1 \otimes P^1 \oplus \cdots \oplus S^d \otimes P^d \tag{8.15}$$

for the operator on $\mathcal{H} \otimes \mathcal{M}$.

Proposition 8.16. *Let E be open in $\mathbb{R}^d$, let $f \in \mathcal{L}(E)$, and assume $(\mathcal{M}, v)$ is a model of f for which v is continuous. Let S and T be d-tuples of commuting self-adjoint operators on a Hilbert space $\mathcal{H}$ with spectrum in E. Then*

$$f(T) - f(S) \ = \ \tilde{v}(S)^* \left[T \odot I - S \odot I\right] \tilde{v}(T). \tag{8.17}$$

Proof: First assume that S and T are (separately) diagonalizable. Let ξ be an eigenvector of S, and η an eigenvector of T, so for some numbers z^r, w^r we have

$$\begin{array}{rcll} S^r\xi &=& w^r\xi & \\ T^r\eta &=& z^r\eta & \qquad \forall\ 1 \leq r \leq d. \end{array}$$

Then

$$\langle [f(T) - f(S)]\, \eta, \xi\rangle_{\mathcal{H}} \;=\; \langle [f(z) - f(w)^*]\, \eta, \xi\rangle_{\mathcal{H}}.$$

Also,

$$\begin{array}{l} \langle \tilde{v}(S)^*\,[T \odot I - S \odot I]\, \tilde{v}(T)\eta, \xi\rangle_{\mathcal{H}} \\ \qquad = \quad \langle [T \odot I - S \odot I]\, \eta \otimes v(z), \xi \otimes v(w)\rangle_{\mathcal{H}\otimes\mathcal{M}} \\ \qquad = \quad \displaystyle\sum_{r=1}^{d} \langle \eta, \xi\rangle_{\mathcal{H}}\ \langle (z^r - \bar{w^r})v^r(z), v^r(w)\rangle_{\mathcal{M}^r} \\ \qquad = \quad (f(z) - f(w))\langle \eta, \xi\rangle_{\mathcal{H}}. \end{array}$$

So both sides of (8.17) agree if you apply them to an eigenvector of T and then take the inner product with an eigenvector of S. By linearity, this is true also for linear combinations of eigenvectors, and as these are assumed dense, we get that (8.17) holds.

If S and T are not diagonalizable, by the spectral theorem we can approximate them in norm by operators that are, and as $\tilde{v}$ and f are both continuous, one gets (8.17) in the limit. $\Box$

Proof of Theorem 8.1: Assume f is in $\mathcal{L}(E)$, and $S(t)$ is a curve of commuting self-adjoint d-tuples with $S(0) = S$ and $S'(0) = \Delta \geq 0$. Choose a model $(\mathcal{M}, v)$ with v locally Lipschitz. Then by Proposition 8.16,

$$f(S(t)) - f(S) \;=\; \tilde{v}(S)^*\,[(S(t) - S) \odot I]\, \tilde{v}(S(t)).$$

As

$$S(t) = S + t\Delta + o(t),$$

we get

$$\begin{array}{rcl} \left.\dfrac{d}{dt} f(S(t))\right|_0 &=& \displaystyle\lim_{t\to 0} \tilde{v}(S)^*\,[\Delta \odot I]\, \tilde{v}(S(t)) \;+\; \lim_{t\to 0} \tilde{v}(S)^*\,[o(1)]\, \tilde{v}(S(t)) \\ &=& \tilde{v}(S)^*\,[\Delta \odot I]\, \tilde{v}(S). \end{array}$$

Hence $f(S(t))$ is differentiable at 0, and its derivative is a positive operator. $\Box$

9 Globally Operator Monotone Functions

Definition 9.1. *Let E be an open set in $\mathbb{R}^d$, and f be a real-valued C^1 function on E. Say f is globally operator monotone on E if, whenever S and T are d-tuples of commuting bounded self-adjoint operators on a Hilbert space with $\sigma(S) \cup \sigma(T) \subset E$, and $S \leq T$, then $f(S) \leq f(T)$.*

If F has the form in Theorem 6.26, then it is globally monotone on rectangles in the μ-resolvent of X.

Theorem 9.2. *Let X be a densely-defined self-adjoint operator on a graded Hilbert space $\mathcal{M} = \mathcal{M}^1 \oplus \cdots \oplus \mathcal{M}^d$, let $v \in \mathcal{M}$, and let F be given by*

$$F(z) \;=\; \langle (X - z)^{-1} v, v \rangle.$$

Let E be an open rectangle in $\mathbb{R}^d$ that is in the μ-resolvent of X. Then F is globally operator monotone on E.

Proof: First observe that if S is a commuting d-tuple of self-adjoint operators on $\mathcal{H}$ and $\sigma(S) \subset E$, then

$$F(S) = R_v^* (I_{\mathcal{H}} - S \odot I)^{-1} R_v \tag{9.3}$$

where $\odot$ is as in equation (8.15) and

$$\begin{aligned} R_v \;:\; \mathcal{H} &\;\to\; \mathcal{H} \otimes \mathcal{M} \\ h &\;\mapsto\; h \otimes v. \end{aligned}$$

Thus equation (9.3) means that for any vectors ξ and η in $\mathcal{H}$,

$$\langle F(S)\eta, \xi \rangle_{\mathcal{H}} \;=\; \langle (I_{\mathcal{H}} \otimes X - \sum_{r=1}^{d} S^r \otimes P^r)^{-1}\; \eta \otimes v, \xi \otimes v \rangle_{\mathcal{H}\otimes\mathcal{M}}. \tag{9.4}$$

Indeed, if η is an eigenvector of S with eigenvalues a^r, then $F(S)\eta = F(a)\eta$, so the left-hand side of (9.4) is $F(a)\langle \eta, \xi \rangle$. But we have

$$(I_{\mathcal{H}} \otimes X - \sum_{r=1}^{d} S^r \otimes P^r)^{-1}\; \eta \otimes v \;=\; \eta \otimes (X - a)^{-1} v,$$

as one can verify by applying $(I_{\mathcal{H}} \otimes X - \sum_{r=1}^{d} S^r \otimes P^r)$ to both sides. So the right-hand side of (9.4) is

$$\langle \eta, \xi \rangle \langle (X - a)^{-1} v, v \rangle,$$

which is the same. If S has a spanning set of eigenvectors, our claim is proved. If it does not, one can approximate it in norm by a d-tuple that does, and the claim follows by continuity.

Now let S and T be d-tuples of commuting self-adjoint operators with $\sigma(S) \cup \sigma(T) \subset E$ and $\Delta := T - S \geq 0$. Let

$$R^r(t) \;=\; (1-t)S^r + tT^r, \qquad 1 \leq r \leq d.$$

Then for t in the range $(0,1)$, the d-tuple $R(t)$ will consist of self-adjoint operators that need not commute with each other. Nonetheless, letting

$$Y(t) \;=\; (I_{\mathcal{H}} \otimes X - R(t) \odot I)$$

then $R_v^* Y(t)^{-1} R_v$ still makes sense by Lemma 9.5. Moreover,

$$\frac{d}{dt} Y(t)^{-1} \;=\; Y(t)^{-1}(\Delta \odot I)\; Y(t)^{-1},$$

and so is positive. Therefore

$$\begin{aligned} F(T) - F(S) \;&=\; R_v^* Y(1)^{-1} R_v - R_v^* Y(0)^{-1} R_v \\ &=\; R_v^* \int_0^1 \frac{d}{dt} Y(t)^{-1} dt R_v \\ &=\; \int_0^1 R_v^* Y(t)^{-1}(\Delta \odot I) Y(t)^{-1} R_v dt \\ &\geq\; 0. \end{aligned}$$

$\square$

Lemma 9.5. *Let $a^r < b^r$, $1 \leq r \leq d$, and let X be a densely defined self-adjoint operator on a graded Hilbert space $\mathcal{M} = \oplus_{r=1}^d \mathcal{M}^r$. Suppose that for every t in $(0,1)$, the point $\lambda_t = (1-t)a + tb$ is not in the μ-spectrum of X. Let $S = (S^1, \dots, S^d)$ be a d-tuple of bounded self-adjoint operators on a Hilbert space $\mathcal{H}$, with $\sigma(S^r) \subset (a^r, b^r)$ for each r. Then $I \otimes X - \sum_{r=1}^d S^r \otimes P^r$ has a bounded inverse.*

Proof: First, suppose $a^r = -1$ and $b^r = 1$ for each r. Then $(-1,1) \cap \sigma(X)$ is empty, and so is $(-1,1) \cap \sigma(I \otimes X)$. So $\|I_{\mathcal{H}} \otimes X \xi\| \geq \|\xi\|$ for every ξ in $\mathcal{H} \otimes \mathcal{M}$. But if $\sigma(S^r) \subset (-1,1)$ for each r, the operator $\sum S^r \otimes P^r$ has norm less than one. Therefore $I \otimes X - \sum_{r=1}^d S^r \otimes P^r$ is invertible.

In the general case, let m^r be the midpoint and c^r half the length of the interval (a^r, b^r), so $a^r = m^r - c^r,\ b^r = m^r + c^r$. Let

$$Y \;=\; \left(\sum_{r=1}^{d} \frac{1}{\sqrt{c^r}} I \otimes P^r\right)\left(I \otimes X - \sum_{r=1}^{d} m^r I_{\mathcal{H}} \otimes P^r\right)\left(\sum_{r=1}^{d} \frac{1}{\sqrt{c^r}} I \otimes P^r\right).$$

If σ_μ denotes the μ-spectrum,

$$\sigma_\mu(Y) = c^{-1}(\sigma_\mu(X) - m),$$

and hence the point $(1-t)(-\mathbb{1}) + t\mathbb{1}$ lies in the μ-resolvent set of Y for $0 < t < 1$. Let $T^r = (1/c^r)(S^r - m^r I_{\mathcal{H}})$. Then T^r is a strict contraction, and so, by the previous case, $Y - \sum T^r \otimes P^r$ is invertible. As

$$\begin{aligned} Y &- \sum_{r=1}^{d} T^r \otimes P^r \\ &= \left(\sum \tfrac{1}{\sqrt{c^r}} I \otimes P^r\right)\left(I \otimes X - \sum S^r \otimes P^r\right)\left(\sum \tfrac{1}{\sqrt{c^r}} I \otimes P^r\right), \end{aligned}$$

we get the desired result. □

We can now prove a global result for rational functions of two variables.

Theorem 9.6. *Let F be a rational function of two variables. Let Γ be the zero-set of the denominator of F. Assume F is real-valued on $\mathbb{R}^2 \setminus \Gamma$. Let E be an open rectangle in $\mathbb{R}^2 \setminus \Gamma$. Then F is globally operator monotone on E if and only if F is in $\mathcal{L}(E)$.*

Proof: Necessity follows from Theorem 8.1.

For sufficiency, by Lemma 9.7 it is sufficient to prove the theorem for F_{t_n} with $t_n \searrow 0$. Suppose the degree of F is n^1 in z^1 and n^2 in z^2. Let $\varphi = \beta \circ F \circ \alpha$. By a result of G. Knese [16], there is a model for φ in a Hilbert space $\mathcal{M} = \mathcal{M}^1 \oplus \mathcal{M}^2$ with $\dim(\mathcal{M}^r) = n^r$ for $r = 1$ and 2; see also the paper [7] by J.A. Ball, C. Sadosky and V. Vinnikov.

Accordingly, in Theorem 6.9, we obtain a realization of F_t on $\mathcal{M}^1 \oplus \mathcal{M}^2$ of the form (6.10); since $\mathcal{M}$ is finite-dimensional, the vector v is in the domain of X. By Theorem 6.33 (v)⇒(ii), for some $v_1 \in \mathcal{M}$

$$F_t(z) \;=\; C + \langle (X - z)^{-1} v_1, v_1 \rangle_{\mathcal{M}},$$

where $\dim(\mathcal{M}^r) = n^r$ for $r = 1$ and 2, and $F_t(\infty, \infty) = C < \infty$. Then the pole-set Γ_t of F_t is contained in the zero-set of $\det(X - z)$ which is a rational function of degree $(\dim(\mathcal{M}^1), \dim(\mathcal{M}^2))$. As these two algebraic sets have the same degree, they must be equal. So the μ-resolvent of X is $\mathbb{R}^2 \setminus \Gamma_t$, and now the result follows from Theorem 9.2. $\square$

Let ρ_t be as in (6.8). The following lemma is elementary.

Lemma 9.7. *Let $t > 0$. Let U be an open set in $\mathbb{R}^d$. Then:*

(i) The function F is globally operator monotone on $U \cap (-1/t, \infty)^d$ if and only if $F \circ \rho_t$ is globally operator monotone on $\rho_t^{-1}(U) \cap (-\infty, 1/t)^d$.

(ii) The function F is globally operator monotone on $U \cap F^{-1}(-1/t, 1/t)$ if and only if $\rho_t \circ F$ is globally operator monotone on the same set.

What happens to Theorem 9.6 in $d \geq 3$ variables? It is still true that rational Löwner functions have finite-dimensional models [8, 7]. However, a recent example of Knese [15] shows that the minimal dimension n^r needed may be strictly greater than the degree of F in z^r. So we cannot rule out the possibility that the μ-spectrum of X contains some other algebraic sets in $\mathbb{R}^d$ than just the zero set of the denominator of F.

Example 9.8. For $0 \leq s \leq 1/2$, the function $(z^1 z^2)^s$ is operator monotone on $(0, \infty) \times (0, \infty)$.

Indeed, if $(0,0) < (A^1, A^2) \leq (B^1, B^2)$ and s is between 0 and 1/2, then

$$\|(A^r)^s (B^r)^{-s}\| \;\leq\; 1 \qquad \text{for } r = 1, 2.$$

Therefore the norm of

$$(B^1)^{-s}(A^1)^s(A^2)^s(B^2)^{-s}$$

is less than or equal to 1, so the largest eigenvalue is less than or equal to 1, and therefore the largest eigenvalue of

$$(B^2)^{-s/2}(B^1)^{-s/2}(A^1)^s(A^2)^s(B^1)^{-s/2}(B^2)^{-s/2} \tag{9.9}$$

is also less than or equal to 1. But (9.9) is self-adjoint, so less than or equal to the identity. Therefore

$$(A^1 A^2)^s \;\leq\; (B^1 B^2)^s.$$

We do not know if $(z^1 z^2)^s$ can be approximated by rational functions in the Löwner class.

Let us close with some questions.

• Is Theorem 9.6 true for rational functions of more than 2 variables?

• Can E be an arbitrary open set in Theorem 9.2?

• Is every function in $\mathcal{L}(E)$ globally operator monotone on E?

• Is every function in $\mathcal{L}_n(E)$ M_n-monotone on E?

References

[1] M. Abate. The Julia-Wolff-Carathéodory theorem in polydisks. *J. Anal. Math.*, 74:275–306, 1998.

[2] J. Agler. On the representation of certain holomorphic functions defined on a polydisc. In *Operator Theory: Advances and Applications, Vol. 48*, pages 47–66. Birkhäuser, Basel, 1990.

[3] J. Agler and J.E. McCarthy. *Pick Interpolation and Hilbert Function Spaces.* American Mathematical Society, Providence, 2002.

[4] J. Agler, J.E. McCarthy, and N.J. Young. A Carathéodory theorem for the bidisk via Hilbert space methods. To appear.

[5] A.V. Arkhangel'skiĭ and L.S. Pontryagin (eds.). *General Topology I.* Springer, Berlin, 1990.

[6] J.A. Ball and V. Bolotnikov. A tangential interpolation problem on the distinguished boundary of the polydisk for the Schur-Agler class. *J. Math. Anal. Appl.*, 273:328–348, 2002.

[7] J.A. Ball, C. Sadosky, and V. Vinnikov. Scattering systems with several evolutions and multidimensional input/state/output systems. *Integral Equations and Operator Theory*, 52:323–393, 2005.

[8] B.J. Cole and J. Wermer. Andô's theorem and sums of squares. *Indiana Math. J.*, 48:767–791, 1999.

[9] M.J. Crabb and A.M. Davie. Von Neumann's inequality for Hilbert space operators. *Bull. London Math. Soc.*, 7:49–50, 1975.

[10] W.F. Donoghue. *Monotone matrix functions and analytic continuation.* Springer, Berlin, 1974.

[11] G. E. Dullerud and F. Paganini. *A course in robust control theory.* Springer Verlag, New York, 2000.

[12] F. Hansen. Operator monotone functions of several variables. *Math. Inequal. Appl.*, 6:1–17, 2003.

[13] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[14] F. Jafari. Angular derivatives in polydisks. *Indian J. Math.*, 35:197–212, 1993.

[15] G. Knese. Rational inner functions in the Schur-Agler class of the polydisk. To appear.

[16] G. Knese. Polynomials with no zeros on the bidisk. *Analysis and PDE*, 3(2):109–149, 2010.

[17] A. Korányi. On some classes of analytic functions of several variables. *Trans. Amer. Math. Soc.*, 101:520–559, 1961.

[18] P. Lax. *Functional Analysis*. Wiley, 2002.

[19] K. Löwner. Über monotone Matrixfunktionen. *Math. Z.*, 38:177–216, 1934.

[20] R. Nevanlinna. Asymptotisch Entwicklungen beschränkter Funktionen und das Stieltjessche Momentproblem. *Ann. Acad. Sci. Fenn. Ser. A*, 18, 1922.

[21] R. Nevanlinna. Remarques sur le lemme de Schwarz. *C.R. Acad. Sci. Paris*, 188:1027–1029, 1929.

[22] V.V. Peller. Hankel operators in the theory of perturbations of unitary and self-adjoint operators. *Funktsional. Anal. i Prilozhen*, 19:37–51, 1985. English translation: Funct. Anal. Appl. 19 (1985) 111-123.

[23] M. Riesz. Sur certains inégalités dans la théorie des fonctions. *Fysiogr. Sällsk. Lund Forh.*, 1(4):18–38, 1931.

[24] W. Rudin. *Function Theory in Polydiscs*. Benjamin, New York, 1969.

[25] W. Rudin. *Lectures on the Edge-of-the-Wedge Theorem*. AMS, Providence, 1971.

[26] W. Rudin. *Real and Comnplex Analysis*. McGraw-Hill, New York, 1986.

[27] D. Sarason. Nevanlinna-Pick interpolation with boundary data. *Integral Equations and Operator Theory*, 30:231–250, 1998.

[28] M. Singh and H.L. Vasudeva. Monotone matrix functions of two variables. *Linear Algebra Appl.*, 328(1-3):131–152, 2001.

[29] M. Stone. *Linear transformations in Hilbert space.* New York, 1932.

[30] N.Th. Varopoulos. On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory. *J. Funct. Anal.*, 16:83–100, 1974.