MŰEGYETEM 1782

Operator monotone and operator convex functions and applications

Bachelor Thesis

Ábel Komálovics

Supervisor: Dr. József Pitrik

Budapest, Hungary

2021

Contents

Chapter 1

Intruduction

A partially ordered set is one of the simplest relational structures, so it is quite natural to investigate the homomorphisms of such structures. As we will see, one can define a partial order on the set of self-adjoint operators of a Hilbert space, to make it such a structure, the monotone functions of which we will examine mainly one the line of the proof of Loewner's Theorem by a recently published book of Barry Simon [2]. The original paper of Loewner [1] is from 1934, and several remarkable results have been achived since then. In quantum mechanics, the observable algebra is that of the self adjoint operators of the associated Hilbert space of the quantum mechanical system. Quasi-entropies are induced by real functions, and that of monotone decreasing functions have some desired properties. Even John von Neumann and Eugene Paul Wigner wrote about Loewner's Theorem in [6]. By the II. Gelfand-Naimark Theorem (Proposition 15.2 of [9]), every C*-algebra can be isometrically embedded into the algebra of bounded operators on some Hilbert space, thus an operator monotone function composed with the inverse of such embedding is a monotone function on the said C*-algebra.

Loewner's student, Fritz Kraus [5] initiated the study of matrix convex functions. Otto Heinävaara [4] also proved great results in this topic. The Hansen–Jensen–Pedersen Theorem (Theorem 11.1 of [2]) is a surprising generalisation of the well-known Jensen's Theorem, but due to length affecting reasons, we will not prove this theorem, but Lemma 4.0.1 is related to it.

In this paper we will first induce a partial order on the bounded operators of a Hilbert space and introduce the notion of operator monotone functions. After the basics, we state and prove Loewner's Theorem through which we will encounter some surprising regularity properties of operator monotone functions. Then we will show that an operator monotone function is monotone even on the operators of a infinite dimensional space. After that, we prove a pleasent remark about operator concave and operator monotone functions, which will help us in showing a surprising theorem. Then we explore the properties of some means when used on operators, and generalise the notion of means on the ground of those. Finally, we shall show the remarkable connection of operator means with operator monotone functions by the Kubo-Ando Theorem, framing the whole thesis.

The thesis may be lengthy, but two of the main aims of it is to be precise and as self contained as possible.

1.1 Basics

Definition 1.1.1 *In a normed space $(V, \|\cdot\|)$, we denote the open ball with radius $r > 0$ and center $x \in V$, $\{y \in V \mid \|x - y\| < r\}$, with $B_r(x)$. The* *clousure of a set H is denoted as $\overline{H}$, and the boundary by ∂H.*

If $(\mathfrak{H}, \langle \cdot | \cdot \rangle)$ is a Hilbert space, we say that a linear operator $A : \mathfrak{H} \to \mathfrak{H}$ is bounded, if $\sup\limits_{x \in \overline{B_1(0)}} \|Ax\| \in \mathbb{R}$. In this case, this supremum is the operator norm of A. We denote the bounded linear operators of $\mathfrak{H}$ with $\mathscr{L}(\mathfrak{H})$.

If $A \in \mathscr{L}(\mathfrak{H})$, we say that A is positive ($A \geq 0$), if $\langle v | Av \rangle \geq 0$ holds for all $v \in \mathfrak{H}$.

We denote the set of positive operators with $\mathscr{L}(\mathfrak{H})^+$, which is a positive cone, that is $\alpha A + \beta B \in \mathscr{L}(\mathfrak{H})^+$ for all $\alpha, \beta \geq 0$ and $A, B \in \mathscr{L}(\mathfrak{H})^+$

A is referred to as strictly positive, if $\langle v | Av \rangle > 0$ holds for all $v \in \mathfrak{H} \setminus \{0\}$.

One can see, that even the $\forall v \in \mathfrak{H} \quad \langle v | Av \rangle \in \mathbb{R}$ condition is equivalent to A being self-adjoint by $\langle v | A^* v \rangle = \overline{\langle v | Av \rangle}$, and the fact that $A = 0 \iff \langle v | Av \rangle = 0 \; \forall v \in \mathfrak{H}$.

Definition 1.1.2 *We define a partial order relation $\leq \subseteq \mathscr{L}(\mathfrak{H}) \times \mathscr{L}(\mathfrak{H})$ by $A \leq B \Leftrightarrow B - A \in \mathscr{L}(\mathfrak{H})^+$*

Since any $A \in \mathscr{L}(\mathfrak{H})$ can be written in the form of $A = \mathfrak{Re}\, A + i\, \mathfrak{Im}\, A = \dfrac{A + A^*}{2} + i\dfrac{iA - iA^*}{2}$, where both $\mathfrak{Re}\, A$ and $\mathfrak{Im}\, A$ are self-adjoint, we see that $A \leq B \iff \mathfrak{Re}\, A \leq \mathfrak{Re}\, B \wedge \mathfrak{Im}\, A = \mathfrak{Im}\, B$. If $\mathfrak{Im}\, A \neq \mathfrak{Im}\, B$, nor $A \leq B$ neither $B \leq A$ hold, thus it is reasonable to continue just with a set of operators, such that $\mathfrak{Im}\, A = A_0$, but to facilitate our progress, we will examine the $A_0 = 0$ equivalency class that is, the self-adjoint operators. We will denote this subspace $\mathscr{L}(\mathfrak{H})_{sa}$.

Proposition 1.1.1 *If A is self-adjoint, $\|A\| = \sup\limits_{x \in \partial B_1(0)} \langle x | Ax \rangle$*

Proof: $\forall x \in \partial\,\mathrm{B}_1(0)$ $\langle x\,|Ax\rangle \leq \|x\|\|Ax\| \leq \|A\|\|x\|^2 = \|A\|$ on the other hand, let $K = \sup\limits_{x\in\partial\,\mathrm{B}_1(0)} \langle x\,|Ax\rangle$ and $y, x \in \partial\,\mathrm{B}_1(0)$. By multiplying with a norm one constant, we can assume that $\langle y\,|Ax\rangle \in \mathbb{R}$. Then

$$
\begin{aligned}
&|\,\langle y\,|Ax\rangle\,| = |\,\langle x+y\,|A(x+y)\rangle - \langle x-y\,|A(x-y)\rangle\,|/4\\
&\leq (|\,\langle x+y\,|A(x+y)\rangle\,| + |\,\langle x-y\,|A(x-y)\rangle\,|)/4\\
&\leq K(\|x+y\|^2 + \|x-y\|^2)/4 = K(2\|x^2\| + 2\|y\|^2)/4 = K
\end{aligned}
$$

where the second to last equation used the paralellogram law. One can choose $y = \dfrac{Tx}{\|Tx\|}$ to see that this implies $\|A\| \leq K$. ■

This is a helpful result, because this implies $0 \leq A \leq B \Rightarrow \|A\| \leq \|B\|$. Now that we have established a partially ordered structure, we can start specifying our question, but to get to functions on operators, we need one more definition

Definition 1.1.3 *If $A \in \mathscr{L}(\mathfrak{H})$, let spec $A = \{\lambda \in \mathbb{C} \mid \nexists (A - \lambda\mathbf{1})^{-1}\}$ denote the spectrum of A, where the non-existence means that $A - \lambda\mathbf{1}$ can not be inverted as a continuous linear operator.*

It is clear, that we can take a polynomial of any $\mathscr{L}(\mathfrak{H})_{sa}$, since the addition and the multiplication with a bounded operator or scalar is continuous. Since we can we can uniformly approximate any continuous function with polynomials, we can take the continuous function of any self-adjoint operator, even in the infinite dimensional case, by using some functional calculus. One can read more about the continuous functional calculus in Chapter VII of [3]. If $\mathbf{Ran} f \subseteq \mathbb{R}$, the image of a self-adjoint operator is self-adjoint, so we search for such functions, since we want to compare the images via the order relation.

We want a function f, that is an order endomorphism on $\mathscr{L}(\mathfrak{H})_{sa}$, that is $A \leq B \Rightarrow f(A) \leq f(B)$. This problem however can be raised in a handier fashion. Let us limit the problem to $\mathfrak{H} \equiv \mathbb{C}^n$ first.

Definition 1.1.4 *If $f : (a,b) \to \mathbb{R}$ function, such that $\forall A, B \in M_n[\mathbb{C}]_{sa}$ $A \leq B \Rightarrow f(A) \leq f(B)$, we refer to f as n-monotone. We will denote the set such functions with $\mathfrak{M}_n(a,b)$.*

Notice, that if $f \in \mathfrak{M}_n(a,b)$, then for any $A \in M_{n-1}[\mathbb{C}]_{sa}$ we can define

$$A_1 = \begin{pmatrix} A & 0 \\ 0 & \dfrac{a+b}{2} \end{pmatrix} \in M_n[\mathbb{C}]_{sa} \quad f(A_1) = \begin{pmatrix} f(A) & 0 \\ 0 & f\left(\dfrac{a+b}{2}\right) \end{pmatrix}$$

so $A \leq B \Rightarrow A_1 \leq B_1 \Rightarrow f(A_1) \leq f(B_1) \Rightarrow f(A) \leq f(B)$ which yields $\mathfrak{M}_{n+1}(a,b) \subseteq \mathfrak{M}_n(a,b)$, thus it is natural to define

Definition 1.1.5 *If $f \in \bigcap\limits_{n \in \mathbb{N}^+} \mathfrak{M}_n(a,b)$, we say that f is operator monotone. Let $\mathfrak{M}_\infty(a,b)$ denote the set of operator monotone functions on (a,b).*

We will see in the applications, that the name is justified, that is any $f \in \mathfrak{M}_\infty(a,b)$ is monotone even on operators of infinite dimensional separable Hilbert spaces. It is easy to see that $\mathfrak{M}_n(a,b)$ is a positive cone that is also closed under composition, thus so is $\mathfrak{M}_\infty(a,b)$.

One will encounter several function spaces throughout the thesis, thus we introduce the notation of the ones, which may deviate from common notations. $C^b(a,b)$ is the set of bounded continuous functions on (a,b). $C_0(a,b)$ Is the space of continuous functions on (a,b) with compact support. $\overline{C_0(\mathbb{R})}$ is the set of continuous functions vanishing at inftinity. All of these spaces use the $\|f\| = \sup\limits_{\operatorname{supp} f} |f|$ norm. We will only consider real valued functions.

The paper uses limits several times in multiple spaces, but I would rather not indicate the convergence as it is consistent in each space, except for the space of bounded linear operators, in which some notation will be used. To avoid confusion, I would like to introduce the topologies which are not indicated.

Definition 1.1.6 $(\mathfrak{I}, <)$ *is a directed set, if* $\mathfrak{I}$ *is a nonempty set, and* $<\subseteq \mathfrak{I} \times \mathfrak{I}$ *is a reflexive and transitive relation, wich obeys* $\forall i, j \in \mathfrak{I}\ \exists c \in \mathfrak{I}:\ c > a$ *and* $c > b$.

If $a : \mathfrak{I} \to (X, \tau)$ *is a function from a directed set to a topological space, then* a *is called a net.* a *converges to a point* $b \in X$*, if* $\forall U \in \tau \quad b \in U \Rightarrow \exists i_0 \in \mathfrak{I}:\ a_i \in U\ \forall i > i_0$*. In this case, we write* $\lim\limits_{i,\mathfrak{I}} a_i = b$.

As the convergence of nets determines a topology, we will use them to define the topologies as opposed to defining the open sets.

Normed spaces will be endowed with the norm topology, where

$\lim\limits_{n,\mathbb{N}} a_n = a \Leftrightarrow \lim\limits_{n,\mathbb{N}} \|a_n - a\|$

The space of measures on a compact topological space $\mathfrak{X}$ will be interpretted as the subspace of the dual space $(C(\mathfrak{X}), \|\cdot\|_\infty)$ which inherits the *weak** topology, that is

$\lim\limits_{i,\mathfrak{I}} \mu_i = \mu \Leftrightarrow\ \forall f \in C(\mathfrak{X})\ \lim\limits_{i,\mathfrak{I}} \int f \mathrm{d}\mu_i = \int f \mathrm{d}\mu.$

Measures on a locally compact space X will use the vague convergence, where

$\lim\limits_{i,\mathfrak{I}} \mu_i = \mu \Leftrightarrow\ \forall f \in C_0(X)\ \lim\limits_{i,\mathfrak{I}} \int f \mathrm{d}\mu_i = \int f \mathrm{d}\mu.$

Distributions use a similar convergence, but use smooth functions:

$\lim\limits_{i,\mathfrak{I}} T_i = T \Leftrightarrow\ \forall f \in C_0^\infty(X)\ \lim\limits_{i,\mathfrak{I}} T_i(f) = T(f)$

Chapter 2

Loewner's Theorem

Notation 2.0.1 $\mathbb{C}_+ = \{z \in \mathbb{C} \mid \mathfrak{Im}\, z > 0\}$

Theorem 2.0.1 (Loewner's Theorem) *If $a < b$ ($a = -\infty$ and/or $b = \infty$ is allowed), and $f : (a, b) \to \mathbb{R}$ is a function, then the following are equivalent:*

($\mathfrak{a}$) $f \in \mathfrak{M}_\infty(a, b)$

($\mathfrak{b}$) *$\exists\mu$ finite measure with* $\operatorname{supp}\mu = \mathbb{R} \setminus (a, b) \equiv J$ *and $\exists A, C \in \mathbb{R}$ with $A \geq 0$, such that*

$$f(x) = C + Ax + \int_J \frac{1 + xy}{y - x} \mathrm{d}\mu(y) \tag{2.1}$$

($\mathfrak{c}$) *f is the restriction of a $g : (\mathbb{C} \setminus \mathbb{R}) \cup (a, b) \to \mathbb{C}$ analytic function which obeys*

$$g(\mathbb{C}_+) \subseteq \mathbb{C}_+$$

Loewner called functions with $g(\mathbb{C}_+) \subseteq \mathbb{C}_+$ *"positive" functions*, but they are also called *Herglotz functions, Pick functions* and *Nevanlinna functions* as well.

We see right away that any $f \in \mathfrak{M}_\infty(a,b)$ is necessarily smooth. What is more, one can notice that every function in $f \in \mathfrak{M}_\infty(\mathbb{R})$ is affine, moreover, even $f \in \mathfrak{M}_2(\mathbb{R}) \Rightarrow\ f\ is\ affine$ holds, as we shall see in Theorem 4.0.2.

Corollary 2.0.1 *If $s > 0$ and $f_s(x) = x^s$, then $f_s \in \mathfrak{M}_\infty(0,\infty) \Leftrightarrow s \leq 1$. In particular, the square root function is operator monotone. Furthermore, $log \in \mathfrak{M}_\infty(0,\infty)$ holds as well*

Proof: $f_s(re^{i\varphi}) = r^s e^{is\varphi}$, and for $r > 0$, $re^{i\varphi} \in \mathbb{C}_+ \Leftrightarrow \varphi \in (0,\pi)$. Thus, if $r > 0$ and $re^{i\varphi} \in \mathbb{C}_+$, then $f_s(re^{i\varphi}) \in \mathbb{C}_+ \Leftrightarrow \varphi \in (0,\pi) \Leftrightarrow s \in (0,1]$.
If we take the principal branch of logarithm on $\mathbb{C} \setminus (-\infty, 0]$, then $log(re^{i\varphi}) = log(r) + i\varphi$, thus $re^{i\varphi} \in \mathbb{C}_+ \Rightarrow \varphi \in (0,\pi) \Rightarrow \varphi > 0$. ∎

It is clear that $(\mathfrak{b}) \Leftrightarrow (\mathfrak{c})$ is the most subtle logical jump since integral representations of analytic functions are not rare to say the least. We shall prove this part first.

2.1 $(\mathfrak{b}) \Leftrightarrow (\mathfrak{c})$

Notation 2.1.1 *$\mathfrak{K} : \mathbb{C}^2 \setminus \Delta \to \mathbb{C}$, $\mathfrak{K}(w,z) = \dfrac{w+z}{w-z}$, where $\Delta = \{(x,x) \mid x \in \mathbb{C}\}$. The open unit ball of $\mathbb{C}$ is traditionally referred to as $\mathbb{D}$, thus we shall not deviate from this notation. For a $z \in \mathbb{C}$, let $\overline{z}$ denote the complex conjugate of z. The linear span of a set H in a vectorspace is denoted by $< H >$.*

Theorem 2.1.1 (Poisson representation) *If f is analytic in a neighbor-*

hood of $\overline{\mathbb{D}}$*, then*

$$f(z) = i\,\mathfrak{Im}\, f(0) + \int_0^{2\pi} \mathfrak{K}(e^{i\theta}, z)\,\mathfrak{Re}\, f(e^{i\theta})\frac{\mathrm{d}\theta}{2\pi} \tag{2.2}$$

Proof: First we notice that

$$\mathfrak{K}(e^{i\theta}, z) = \frac{1 + ze^{-i\theta}}{1 - ze^{-i\theta}} = (1+ze^{-i\theta})\left(\sum_{n\in\mathbb{N}} z^n e^{-in\theta}\right) = 1+2\sum_{n\in\mathbb{N}^+} z^n e^{-in\theta} \tag{2.3}$$

where the sum converges uniformly on $[0, 2\pi]$ for each $z \in \mathbb{D}$, since such sets are compact subsets of $\mathbb{D}$. Since f is analytic in the neighborhood of $\overline{\mathbb{D}}$, the sum $f = \sum_{n\in\mathbb{N}} a_n z^n$ converges uniformly on $\overline{\mathbb{D}}$ Now we can rewrite

$\mathfrak{Re}\, f(e^{i\theta}) = \mathfrak{Re}\, a_0 + \frac{1}{2}\sum_{n\in\mathbb{N}^+} a_n e^{in\theta} + \overline{a}_n e^{-in\theta}$ also converging uniformly on $[0, 2\pi]$.

The last thing we need is that $\int e^{-in\theta} e^{im\theta}\frac{\mathrm{d}\theta}{2\pi} = \delta_{nm}$

$$\begin{aligned}
&\int \mathfrak{K}(e^{i\theta}, z)\,\mathfrak{Re}\, f(e^{i\theta})\frac{\mathrm{d}\theta}{2\pi} \\
&= \int \left(1 + 2\sum_{n\in\mathbb{N}^+} z^n e^{-in\theta}\right)\left(\mathfrak{Re}\, a_0 + \frac{1}{2}\sum_{k\in\mathbb{N}^+} a_k e^{ik\theta} + \overline{a_k} e^{-ik\theta}\right)\frac{\mathrm{d}\theta}{2\pi} \\
&= \int \mathfrak{Re}\, a_0 + 2\frac{1}{2}\sum_{n\in\mathbb{N}^+} a_n z^n \frac{\mathrm{d}\theta}{2\pi} = \mathfrak{Re}\, a_0 + f(z)
\end{aligned}$$

■

Lemma 2.1.1 $\mathscr{H} =< \{\mathfrak{K}(\cdot, z), \overline{\mathfrak{K}(\cdot, z)} | z \in \mathbb{D}\} > \Rightarrow \overline{\mathscr{H}} = C(\partial\mathbb{D})$

Proof: As $\mathfrak{K}(e^{i\theta}, z) = 1 + 2\sum_{n\in\mathbb{N}^+} z^n e^{-in\theta}$ and

$\forall u, v \in \mathbb{D}$ $\frac{\mathfrak{K}(e^{i\theta}, u) - \mathfrak{K}(e^{i\theta}, v)}{u - v} \in \mathscr{H}$, we see that the derivative in the second variable is in the closure, since it is a limit. Now we have that

$\left.\frac{d}{dz}\mathfrak{K}(e^{i\theta}, z)\right|_{z=0} = e^{-i\theta} \in \overline{\mathscr{H}}$. The same argument can be made with the nth derivatives, to see that $e^{-i(n+1)\theta} \in \overline{\mathscr{H}}$, since $\overline{\mathscr{H}}$ is closed under multiplication by scalars. Using conjugates to get the positive powers, we arrive at $\{e^{ik\theta} \mid k \in \mathbb{Z}\} \subseteq \overline{\mathscr{H}}$, which is a dense set according to Weierstrass Approximation Theorem, so $C(\partial\mathbb{D}) \subseteq \overline{\{e^{ik\theta} \mid k \in \mathbb{Z}\}} \subseteq \overline{\mathscr{H}}$ ■

Theorem 2.1.2 (Herglotz representation on $\mathbb{D}$) *If f is an analytic function on $\mathbb{D}$ with $\mathfrak{Re}\, f(z) > 0\ \forall z \in \mathbb{D}$, then $\exists \mu!$ finite measure on $\partial\mathbb{D}$ such that*

$$f(z) = i\,\mathfrak{Im}\, f(0) + \int_0^{2\pi} \mathfrak{K}(e^{i\theta}, z)\mathrm{d}\mu(\theta) \tag{2.4}$$

Proof: If f satisfies the conditions, then so does $g = \frac{f - i\,\mathfrak{Im}\, f(0)}{\mathfrak{Re}\, f(0)}$, so let us consider this function, since it has $g(0) = 1$. For all $r \in (0, 1)$ g is analytic in the neighborhood of $r\overline{\mathbb{D}}$, so by Theorem 2.1.1, $\forall z \in \mathbb{D}$

$$g(rz) = \int_0^{2\pi} \mathfrak{K}(e^{i\theta}, z)\mathrm{d}\mu_r(\theta), \quad \mathrm{d}\mu_r(\theta) = \mathfrak{Re}\, g(re^{i\theta})\frac{\mathrm{d}\theta}{2\pi}$$

Notice, that this gives us $1 = g(0) = \int_0^{2\pi} \mathrm{d}\mu_r(\theta)$. By continuity of g, we have that $\lim_{r \nearrow 1} \int_0^{2\pi} \mathfrak{K}(e^{i\theta}, z)\mathrm{d}\mu_r(\theta) = g(z)$ exists for each $z \in \mathbb{D}$. The predecessing lemma shows us, that $\mu = \lim_{r \nearrow 1} \mu_r$ exists, since it is defined on a dense subset of $C(\partial\mathbb{D})$. Since the convergence is that of the weak* topology, we can see that $\int_{\partial\mathbb{D}} \mathbf{1}\mathrm{d}\mu = \int_{\partial\mathbb{D}} \mathbf{1}\mathrm{d}\mu_r = 1$, thus $\mathrm{d}\mu$ is a probability measure. Since the set of probability measures on $\partial\mathbb{D}$ is metrizable, thus the limit μ is unique. One can use the inverse of the transform in the beginning of the proof to arrive at (2.4)

■

The proof shows us that the measure

$$\mathrm{d}\mu = \lim_{r \nearrow 1} \mathfrak{Re}\, f(re^{i\theta}) \frac{d\theta}{2\pi} \tag{2.5}$$

and that $\mu(\partial\mathbb{D}) = \mathfrak{Re}\, f(0)$ thus $\mathrm{d}\mu$ is finite.

Theorem 2.1.3 (Herglotz representation on $\mathbb{C}_+$) *If $f : \mathbb{C}_+ \to \mathbb{C}_+$ is an analytic function, then $\exists! \mu$ finite measure on $\mathbb{R}$ and $\exists A \geq 0$ constant, such that*

$$f(z) = \mathfrak{Re}\, f(i) + Az + \int_{\mathbb{R}} \frac{1 + xz}{x - z} \mathrm{d}\mu(x) \tag{2.6}$$

Conversely, any f in such form with $A \geq 0$ obeys $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$

Proof: Starting with the converse, we can write

$$\frac{1 + xz}{x - z} = -x + \frac{1 + x^2}{x - z}$$

hence

$$\mathfrak{Im}\, \frac{1 + xz}{x - z} = \mathfrak{Im}\, \frac{(1 + x^2)(x - \overline{z})}{|x - z|^2} = \mathfrak{Im}\, \frac{(1 + x^2)z}{|x - z|^2} = \frac{(1 + x^2)}{|x - z|^2} \mathfrak{Im}\, z$$

which means that any function in the form of (2.6) obeys $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$ Let us consider the fractional linear function $T : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$

$$T(z) = \frac{z - i}{z + i} \qquad T^{-1}(w) = i\frac{1 + w}{1 - w}$$

One can see, that $T(\overline{\mathbb{R}}) = \partial\mathbb{D}$ and $T(i) = 0$ which implies $T(\mathbb{C}_+) = \mathbb{D}$ and $T^{-1}(\mathbb{D}) = \mathbb{C}_+$ by ground properties of fractional linear maps. Thus $f : \mathbb{C}_+ \to \mathbb{C}_+ \Rightarrow -if \circ T^{-1} : \mathbb{D} \to \{z \in \mathbb{C} \mid \mathfrak{Re}\, z > 0\}$, which justifies the use of the previous theorem, which hands us a unique $\tilde{\mu}$ measure on $\partial\mathbb{D}$, such that

$$-if(T^{-1}w) = i\,\mathfrak{Im}(-if(i)) + \int_0^{2\pi} \mathfrak{K}(e^{i\theta}, w)\mathrm{d}\tilde{\mu}(\theta) \tag{2.7}$$

By multiplying the equation with i, setting $z = T^{-1}(w)$ and $\mu = \tilde{\mu} \circ T$ we arrive at

$$f(z) = \mathfrak{Re} f(i) + az + \int_{\mathbb{R}} \mathfrak{K}(T(x), T(z))\mathrm{d}\mu(x)$$

where $\mathfrak{K}(T(x), T(z)) = \dfrac{1+xz}{x-z}$ follows from elementary computation. One can rewrite (2.6) as

$$f(z) = \mathfrak{Re} f(i) + az + \int_{\mathbb{R}} \frac{1}{x-z} - \frac{x}{1+x^2}\mathrm{d}\nu(x) \tag{2.8}$$

where

$$\mathrm{d}\nu = (1+x^2)\mathrm{d}\mu \tag{2.9}$$

ν is no longer a finite measure, but still is a σ-finite one. This form is used frequently as well. ■

Now we see the connection between the representation and the $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$ property. What is left is extending this funtion to the lower half plane $\mathbb{C}_-$ through an (a, b) interval. After we get to know the measure of (2.6), this problem will be more menageable.

Proposition 2.1.1 *The measure of (2.6) is given by*

$$d\mu(x) = \lim_{\varepsilon \searrow 0} \frac{1}{1+x^2}\frac{1}{\pi}\,\mathfrak{Im}\, f(x+i\varepsilon)\,dx \tag{2.10}$$

where lim *resebles the vague limit.*

Proof: Since (2.8) holds, we have that

$$\frac{1}{\pi}\,\mathfrak{Im}\,f(x+i\varepsilon) = \frac{a\varepsilon}{\pi}+\int\limits_{\mathbb{R}} \frac{1}{\pi}\,\mathfrak{Im}\,\frac{1}{y-(x+i\varepsilon)}\mathrm{d}\nu(x) = \frac{a\varepsilon}{\pi}+\int\limits_{\mathbb{R}} \frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\mathrm{d}\nu(x)$$

If we make a $u = \dfrac{x-y}{\varepsilon}$ substitution, we get $\dfrac{arctan(u)+C}{\pi}$ as an antiderivative, so the integral is 1 for every $\varepsilon > 0$. One can also see, that the integrand goes to 0 in y uniformly on every $\{(x-y)^2 > \delta\}$ as $\varepsilon \searrow 0$, so it is an approximate delta function for every fixed x. Let $g \in C_0(\mathbb{R})$

$$\begin{aligned}
&\lim_{\varepsilon\searrow 0}\int\limits_{\mathbb{R}} g(x)\frac{1}{1+x^2}\frac{1}{\pi}\,\mathfrak{Im}\,f(x+i\varepsilon)dx \\
&= \lim_{\varepsilon\searrow 0}\int\limits_{\mathbb{R}} g(x)\frac{1}{1+x^2}\left(\frac{a\varepsilon}{\pi}+\int\limits_{\mathbb{R}}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\mathrm{d}\nu(y)\right)dx \\
&= \lim_{\varepsilon\searrow 0}\int\limits_{\mathbb{R}} g(x)\frac{1}{1+x^2}\frac{a\varepsilon}{\pi}dx + \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} g(x)\frac{1}{1+x^2}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\mathrm{d}\nu(y)dx
\end{aligned}$$

Since the first integrand converges to 0 uniformly, it is enough to continue with the second one, where we have a product of two σ-finite measure spaces and the iterated absolute integral

$$\begin{aligned}
&\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}}\left|g(x)\frac{1}{1+x^2}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\right|\mathrm{d}\nu(y)\mathrm{d}x\\
&\leq\left\|\frac{g(x)}{1+x^2}\right\|\int\limits_{\operatorname{supp}g}\int\limits_{\mathbb{R}}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\mathrm{d}\nu(y)\mathrm{d}x\\
&=\left\|\frac{g(x)}{1+x^2}\right\|\int\limits_{\operatorname{supp}g}\frac{1}{\pi}\mathfrak{Im}(f(x+i\varepsilon)-a\varepsilon)\mathrm{d}x\\
&\leq\left\|\frac{g(x)}{1+x^2}\right\|\int\limits_{\operatorname{supp}g}\frac{1}{\pi}\sup_{\operatorname{supp}g}\mathfrak{Im}\,f(y+i\varepsilon)\mathrm{d}x\\
&=\lambda(\operatorname{supp}g)\left\|\frac{g(x)}{1+x^2}\right\|\frac{1}{\pi}\sup_{\operatorname{supp}g}\mathfrak{Im}\,f(y+i\varepsilon)<\infty
\end{aligned}$$

is finite, since $\operatorname{supp} g$ is compact. In the last row, λ denotes the Lebesgue measure. Hence we can use the Fubini-Tonelli Theorem to interchange the interals . Notice, that $\int_{\mathbb{R}} g(x)\frac{1}{1+x^2}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}dx$ has a pointwise limit as $\varepsilon \searrow 0$, $\frac{g(y)}{1+y^2}$. It can be dominated by $\|g\|\mathbf{1}_{\operatorname{supp} g}$, thus by the Dominated Convergence Theorem we can interchange the limit and the outer integral, to arrive at

$$\begin{aligned}
&\lim_{\varepsilon\searrow 0}\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}}g(x)\frac{1}{1+x^2}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\mathrm{d}\nu(y)\mathrm{d}x\\
&=\lim_{\varepsilon\searrow 0}\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}}g(x)\frac{1}{1+x^2}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\mathrm{d}x\mathrm{d}\nu(y)\\
&=\int\limits_{\mathbb{R}}\lim_{\varepsilon\searrow 0}\int g(x)\frac{1}{1+x^2}\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}\mathrm{d}x\mathrm{d}\nu(y)\\
&=\int\limits_{\mathbb{R}}g(y)\frac{1}{1+y^2}\mathrm{d}\nu(y)=\int g(y)\mathrm{d}\mu(y)
\end{aligned}$$

■

By combining (2.5), (2.7) and $\mathrm{d}\mu \circ T^{-1} = d\tilde{\mu}$, one can also see that $\mathrm{d}\mu(T^{-1}(e^{i\theta})) = d\tilde{\mu}(\theta) := \lim_{r \nearrow 1} \mathfrak{Re}(-if(T^{-1}(re^{i\theta})))\frac{d\theta}{2\pi} = \lim_{r \nearrow 1} \mathfrak{Im}\, f(T^{-1}(re^{i\theta}))\frac{d\theta}{2\pi}$ and we have a radially directed limit to the unit circle inside T^{-1}, $\lim_{r \nearrow 1} T^{-1}(re^{i\theta})$, which is equivalent to a directed limit from $T^{-1}(0) = i$ to $T^{-1}(e^{i\theta}) = cot(\theta/2)$ in the form of $\lim_{t \nearrow 1} ti + (1-t)cot(\theta/2)$. By further transformations, we can arrive at the same form of $\mathrm{d}\mu$ this way as well.

Corollary 2.1.1 *A $f : \mathbb{C}_+ \to \mathbb{C}_+$ function is a restriction of a $g : (\mathbb{C} \setminus \mathbb{R}) \cup (a, b) \to \mathbb{C}$ analytic function obeying $g((a, b)) \subseteq \mathbb{R}$ if and only if in the representation (2.6) we have $\mu(a, b) = 0$.*

Proof: If $\mathfrak{Im}\, g((a, b)) = \{0\}$, then by (2.10) we know that $\mu(a, b) = 0$, since g is continuous and $f(z) = g(z)\ \forall z \in \mathbb{C}^+$.

If $\mu(a, b) = 0$, then f is continuously defined on (a, b) by the representation with $f((a, b)) \subseteq (c, d)$ for some $c < d$. We can define

$$g(z) = \begin{cases} f(z) & if\ z \in \mathbb{C}_+ \cup (a, b) \\ \overline{f(\overline{z})} & if\ z \in \mathbb{C}_- \cup (a, b) \end{cases} \tag{2.11}$$

g is defined on the (a, b) interval twice, to make it even more clear that g is continuous on $\mathbb{C}_+ \cup (a, b)$ and on $\mathbb{C}_- \cup (a, b)$ as well, since f is. g is continuous, for given a $Z \subseteq \mathbf{Ran}g$ closed set, write $Z_+ = Z \cap \overline{\mathbb{C}}_+ \subseteq g(\mathbb{C}_+ \cup (a, b))$ and $Z_- = Z \cap \overline{\mathbb{C}}_- \subseteq g(\mathbb{C}_- \cup (a, b))$ which are closed sets, the preimages of which are closed. Since $g^{-1}(Z) = g^{-1}(Z_+) \cup g^{-1}(Z_-)$, it is closed.

As a function is analytic if and only if it is holomorphic, we only need to check that $\int_\gamma g = 0$ for any simple closed curve γ. For any $\gamma \subseteq \mathbb{C}_+$ or $\gamma \subseteq \mathbb{C}_-$ it is trivial. If $\gamma \cap \mathbb{C}_+ \neq \varnothing \neq \gamma \cap \mathbb{C}_-$, we can divide γ into γ_1 and

γ_2 where γ_1 is the portion of γ which lies in $\mathbb{C}_+$ with the connection of the two ends, and γ_2 is the same, just on $\mathbb{C}_-$ so that $\gamma_1 + \gamma_2 = \gamma$. These are simple closed curves, thus compact sets. Hence $\gamma_1 + [0, i]$ and $\gamma_2 + [0, -i]$ are also compact resulting in f being uniformly continuous on them, therefore $\int_{\gamma_1} g(x)dz = \lim_{\varepsilon \searrow 0} \int_{\gamma_1} f(z + i\varepsilon)dz = \lim_{\varepsilon \searrow 0} 0$. The same argument can be made to show, that $\int_{\gamma_2} g(x)dz = 0$ which tells us

$$\int_\gamma g = \int_{\gamma_1} g + \int_{\gamma_2} g = 0$$

For any $\gamma \cap \mathbb{C}_+ = \varnothing$ or $\gamma \cap \mathbb{C}_- = \varnothing$, the same argument works as on γ_1 or γ_2

■

With this corollary, we completed the proof of the $(\mathfrak{b}) \Leftrightarrow (\mathfrak{c})$ part of Loewner's Theorem, that is

Theorem 2.1.4 *For any $f : \mathbb{C} \to \mathbb{C}$ function the following are equivalent*

$(\mathfrak{b})$ $f : (\mathbb{C}\backslash\mathbb{R}) \cup (a, b) \to \mathbb{C}$ *is an analytic function which obeys* $f((a, b)) \subseteq \mathbb{R}$ *and* $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$

$(\mathfrak{c})$ $\exists \mu$ *finite measure with* $supp(\mu) = \mathbb{R} \setminus (a, b) \equiv J$ *and* $\exists A, B \in \mathbb{R}$ *with* $A \geq 0$, *such that*

$$f(x) = C + Ax + \int_J \frac{1 + xy}{y - x} \mathrm{d}\mu(y) \tag{2.12}$$

Proof: It follows from Corollary 2.1.1 and Theorem 2.1.3. ■

2.2 $(\mathfrak{b}) \Rightarrow (\mathfrak{a})$

We could prove $(\mathfrak{b}) \Leftrightarrow (\mathfrak{c})$ in the original form, but for the $(\mathfrak{a}) \Leftrightarrow ((\mathfrak{b}) \vee (\mathfrak{c}))$ part we will reduce the problem from any interval to $(-1,1)$. Firstly, we state this special case of Loewner's Theorem.

Theorem 2.2.1 *For any $f : (-1,1) \to \mathbb{R}$ fuction, the following are equivalent:*

($\mathfrak{a}$) $f \in \mathfrak{M}_\infty(-1,1)$

($\mathfrak{b}$) $\exists\mu$ *finite measure on* $[-1,1]$, *such that* $\exists A, C \in \mathbb{R}$, $A \geq 0$ *obeying*

$$f(x) = C + Ax + \int_{-1}^{1} \frac{x}{1+\lambda x} \mathrm{d}\mu(\lambda) \tag{2.13}$$

($\mathfrak{c}$) f *is a restriction of a* $g : (\mathbb{C} \setminus \mathbb{R}) \cup (-1,1) \to \mathbb{C}$ *analytic function, which satisfies*

$$g(\mathbb{C}_+) \subseteq \mathbb{C}_+$$

We can prove *($\mathfrak{b}$)* $\Leftrightarrow$ *($\mathfrak{c}$)* right away.

Proof: Since $(\mathfrak{b}) \Leftrightarrow (\mathfrak{c})$ has already been proven for the general case with $(\mathfrak{b})$ taking a different form, we only need to show that *($\mathfrak{b}$)* is equivalent to the $(a,b) \equiv (-1,1)$ case of $(\mathfrak{b})$, which is

$$f(z) = \mathfrak{Re}\, f(i) + az + \int_{\mathbb{R}\setminus(-1,1)} \frac{1+xz}{x-z} \mathrm{d}\mu(x)$$

One can notice, that $\dfrac{1+xz}{x-z} = x^{-1} + \dfrac{xz}{x-z}\dfrac{1+x^2}{x^2}$.

Let us define $c = \mathfrak{Re}\, f(i) + \int_{\mathbb{R}\setminus(-1,1)} x^{-1} \mathrm{d}\mu(x)$ and change the measure to

$\mathrm{d}\tilde{\nu}(-x^{-1}) = \dfrac{1+x^2}{x^2}\mathrm{d}\mu(x)$ and also the variable to $\lambda = -x^{-1}$. Notice that $\mathbf{Dom}\,\tilde{\nu} = [-1,1] \setminus \{0\}$. Now we have

$$f(z) = c + az + \int\limits_{\mathbf{Dom}\,\tilde{\nu}} \frac{z}{1+\lambda z}\mathrm{d}\tilde{\nu}(\lambda)$$

Since the integrand is z at $\lambda = 0$, we can absorb the linear part part with setting $\nu = \tilde{\nu} + a\delta$, so $\mathbf{Dom}\,\nu = [-1,1]$ and we arrive at the desired form

$$f(z) = f(0) + \int\limits_{-1}^{1} \frac{z}{1+\lambda z}\mathrm{d}\nu(\lambda)$$

Notice that we only used equivalent transformations, hence the two forms are equivalent.

■

Theorem 2.2.2 *If* $(\mathfrak{a}) \Leftrightarrow (\mathfrak{c})$ *holds in Loewner's Theorem for the* $(a,b) \equiv (-1,1)$ *case, then it holds for all* $a < b$.

Proof: Suppose first that $-\infty < a < b < \infty$. Let us define

$$T : (-1,1) \to (a,b) \qquad T(x) = \frac{b-a}{2}x + \frac{a+b}{2}$$

which is a monotone bijection. As an affine map with positive coefficient, clearly T and T^{-1} are operator monotone, thus for every $f : (a,b) \to \mathbb{R}$

$$f \in \mathfrak{M}_\infty(a,b) \Leftrightarrow f \circ T \in \mathfrak{M}_\infty(-1,1)$$

T extends analytically to $\mathbb{C}$ with $T(\mathbb{C}_+) \subseteq \mathbb{C}_+$, thus $\exists\, g : (\mathbb{C}\setminus\mathbb{R}) \cup (a,b) \to \mathbb{C}$ analytic with $g(\mathbb{C}_+) \subseteq \mathbb{C}_+$ and $f \subseteq g$ if and only if $\exists\, g : (\mathbb{C}\setminus\mathbb{R}) \cup (-1,1) \to \mathbb{C}$

analytic with $g(\mathbb{C}_+) \subseteq \mathbb{C}_+$ and $f \circ T \subseteq g$. This proves our proposition for bounded intervals.

For $(a,b) \equiv (0,\infty)$, if $f \in \mathfrak{M}_\infty(0,\infty)$, then $f \in \bigcap_{n\in\mathbb{N}^+} \mathfrak{M}_\infty(0,n)$, so for all $n \in \mathbb{N}^+$, f has a suitable analytic continuation to $(\mathbb{C}\setminus\mathbb{R}) \cup (0,n)$ of which f is a restriction onto $(0,n)$. By the uniqueness of analytic continuation, f has a suitable continuation, which is defined on $(\mathbb{C}\setminus\mathbb{R}) \cup (0,\infty)$. Conversely, if f has a suitable continuation to $(\mathbb{C}\setminus\mathbb{R}) \cup (0,\infty)$, then it has one to $(\mathbb{C}\setminus\mathbb{R}) \cup (0,n)$ for every $n \in \mathbb{N}^+$, hence $f \in \mathfrak{M}_\infty(0,n)$ for every $n \in \mathbb{N}^+$, which means that $f \in \mathfrak{M}_\infty(0,\infty)$, since $\forall k \in \mathbb{N}^+\ \forall A, B \in M_k[\mathbb{C}]$ $\big(A, B$ is self-adjoint $\Rightarrow \exists\, n \in \mathbb{N}^+ :\ A, B < n\mathbf{1}\big)$. The same line of reasoning works for any infinite interval. ■

Theorem 2.2.3 (Schur decomposition) *If $A, B, C \in M_n[\mathbb{C}]$ and C is invertible, then*

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} = \begin{pmatrix} \mathbf{1} & BC^{-1} \\ 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} A - BC^{-1}B^* & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 \\ C^{-1}B^* & \mathbf{1} \end{pmatrix} \tag{2.14}$$

holds, and

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \geq 0 \quad \Leftrightarrow \quad A \geq BC^{-1}B^* \text{ and } C \geq 0 \tag{2.15}$$

$A - BC^{-1}B^*$ is often reffered to as the Schur complement of A.

Proof: (2.14) is straightforward computation, which after setting $M = \begin{pmatrix} \mathbf{1} & 0 \\ C^{-1}B^* & \mathbf{1} \end{pmatrix}$, takes the form of

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} = M^* \begin{pmatrix} A - BC^{-1}B^* & 0 \\ 0 & C \end{pmatrix} M$$

Since M is invertible, it is a bijection, so by the definition of positivity, we see that

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \geq 0 \quad \Leftrightarrow \quad \begin{pmatrix} A - BC^{-1}B^* & 0 \\ 0 & C \end{pmatrix} \geq 0$$

The R.H.S. is by definition equivalent to the R.H.S. of (2.15).

■

Theorem 2.2.4 *If* $n \in \mathbb{N}^+$ *and* $A, B \in M_n[\mathbb{C}]$ *are self-adjoint matrices with eigenvalues in* $(-\infty, 0)$*, then*

$$A \leq B \Leftrightarrow -A^{-1} \leq -B^{-1} \tag{2.16}$$

Proof: $\mathbb{R}^{2n} = \mathbb{R}^n \oplus \mathbb{R}^n$, hence it is easy to check by the definition of positivity, that for any $C, D, E \in M_n[\mathbb{C}]$ where E is invertible

$$\begin{pmatrix} C & D \\ D & E^{-1} \end{pmatrix} \geq 0 \Leftrightarrow \begin{pmatrix} E^{-1} & D \\ D & C \end{pmatrix} \geq 0 \tag{2.17}$$

If we take $D = \mathbf{1}$ and consider the fact, that the Schur decomposition gives us

$$\forall C, E \geq 0 \qquad \begin{pmatrix} C & \mathbf{1} \\ \mathbf{1} & E^{-1} \end{pmatrix} \geq 0 \Leftrightarrow 0 < E \leq C$$

we can set $C = -A$ and $E = -B$ for (2.17) to give us

$$0 < -B \leq -A \Leftrightarrow 0 < -A^{-1} \leq -B^{-1}$$

by the special case of the decomposition. ■

This shows us that if $f(x) = -x^{-1}$, then $f \in \mathfrak{M}_\infty(-\infty, 0)$. However, one can also consider the fact that $0 < A \leq B \Rightarrow -B \leq -A < 0$
$\Rightarrow B^{-1} = f(-B) \leq f(-A) = A^{-1} \Rightarrow f(A) = -A^{-1} \leq -B^{-1} = f(B)$, which yields $f \in \mathfrak{M}_\infty(0, \infty)$.

Theorem 2.2.5 $(\mathfrak{b}) \Rightarrow (\mathfrak{a})$ *in Loewner's Theorem.*

Proof: Since $(\mathfrak{b}) \Leftrightarrow (\mathfrak{c})$ is already proven, and $(\mathfrak{c}) \Rightarrow (\mathfrak{a})$ for $(a,b) \equiv (-1,1)$ implies $(\mathfrak{c}) \Rightarrow (\mathfrak{a})$ for all (a,b), it is enough to show that $(\mathfrak{b}) \Rightarrow (\mathfrak{a})$ for $(a,b) \equiv (-1,1)$.

Due to $\mathfrak{M}_\infty(a,b)$ being closed under addition, and $\forall A \geq 0\ \forall B \in \mathbb{R}$ $(x \mapsto Ax + B) \in \mathfrak{M}_\infty(a,b)$, we only need to check, that the integrand of (2.13) is in $\mathfrak{M}_\infty(a,b)$, that is,

$$\forall \lambda \in [-1,1] \quad -\mathbf{1} < A \leq B < \mathbf{1} \;\Rightarrow\; \frac{A}{1+\lambda A} \leq \frac{B}{1+\lambda B}$$

Since for $\lambda = 0$, it is trivial, we can set $\mu = |\lambda|^{-1} \quad \mu \geq 1$, to get

$$\forall \mu \geq 1 \quad -\mathbf{1} < A \leq B < \mathbf{1} \;\Rightarrow\; \frac{A}{\mu \pm A} \leq \frac{B}{\mu \pm B}$$

Since $\dfrac{x}{\mu \pm x} = 1 \mp \dfrac{\mu}{\mu \pm x}$,

$$\forall \mu \geq 1 \quad -\mathbf{1} < A \leq B < \mathbf{1} \;\Rightarrow\; \mp(\mu \pm A)^{-1} \leq \mp(\mu \pm B)^{-1}$$

Notice that $0 < (\mu \pm B), (\mu \pm A)$, so we can use the fact that $(x \mapsto -x^{-1}) \in \mathfrak{M}_\infty(0,\infty)$. The implication follows in both cases, that is

$$\forall \mu \geq 1 \quad -\mathbf{1} < A \leq B < \mathbf{1} \;\Rightarrow\; -(\mu + A)^{-1} \leq -(\mu + B)^{-1}$$

$$\forall \mu \geq 1 \quad -\mathbf{1} < -B \leq -A < \mathbf{1} \;\Rightarrow\; -(\mu - B)^{-1} \leq -(\mu - A)^{-1}$$

∎

2.3 $(\mathfrak{a}) \Rightarrow (\mathfrak{c})$

For this section, we prove some results in the finite dimensional case, some of which are remarkable on their own. One will encounter propositions that only have a citation for the proof, because the they are not closely related to the topic or the machinery of it.

Definition 2.3.1 *If* $n \in \mathbb{N}^+$, $A = (a_{ij})_{i,j \leq n} \in M_n[\mathbb{C}]$ *and* $I \subseteq \{1, \dots, n\}$, *let* $d_I(A)$ *denote* $\det A_I$, *where* $A_I = (a_{ij})_{(i,j) \in I^2}$. *If* $I = \{1, \dots, k\}$, $d_I(A)$ *is called main principle determinant and principle determinant otherwise.*

Lemma 2.3.1 *If* $A \in M_n[\mathbb{C}]$ *is self-adjoint, then* A *is strictly positive, if and only if each main principle determinant is positive.* A *is positive, if and only if each principle determinant is non-negative.*

Proof: Proposition 5.8 of [2] ■

Definition 2.3.2 *Let* $f \in C(a, b)$ *be a function and* $x_1, \dots, x_n \in (a, b)$ *distinct points. We define the nth divided difference of* f *recursively by*
$[x_1; f] = f(x_1)$, $[x_1, \dots, x_n; f] = \dfrac{[x_2, \dots, x_n; f] - [x_1, \dots, x_{n-1}; f]}{x_n - x_1}$.

Proposition 2.3.1 *If* $f \in C(a, b)$, *then the nth divided difference is a symmetric function on* $(a, b)^n$. *If* $f \in C^{n-1}(a, b)$, *then the nth divided difference has a continuous extension to* $(a, b)^n$ *by*

$$[x_1, \dots, x_n; f] = \sum_{j=1}^{l} \frac{1}{(m_j - 1)!} D^{m_j - 1} \left[f(x) \prod_{k \neq j} (x - x_k)^{-m_k} \right] \Bigg|_{x=0}$$

where $(x_1, \dots, x_n) = \prod_{j=1}^{l} \left(\prod_{i=1}^{m_j} y_j \right)$

Proof: Theorem 5.13 of [2] ■

This proposition is useful, because it asserts that if we have a function regular enough, we can approximate the nth derivative with the divided difference, since $\forall (x, \dots, x) \in (a, b)^{n+1} \quad [x, \dots, x; f] = \dfrac{\mathrm{D}^n f(x)}{n!}$.

Definition 2.3.3 (Loewner matrix) *For any $n \in \mathbb{N}^+$, $f \in C^1(a,b)$ and $a < x_1 < \ldots < x_n < b$ we define the $n \times n$ Loewner matrix of f by*

$$L_n(x_1, \ldots, x_n; f)_{ij} = \begin{cases} [x_i, x_j; f] & if\ i \neq j \\ \mathrm{D}f(x_i) & if\ i = j \end{cases}$$

Proposition 2.3.2 *Let $k \in \mathbb{N}$ and $f \in \mathfrak{M}_n(a,b) \cap C^k(a,b)$. Then there exist $f_m \in C^\infty(a+\frac{1}{m}, b-\frac{1}{m}) \cap \mathfrak{M}_n(a+\frac{1}{m}, b-\frac{1}{m})$ such that $\mathrm{D}^j f_m$ converges uniformly to $\mathrm{D}^j f$ on all compact subintervals of (a,b) for all $j \in \{0, \ldots, k\}$.*

Proof: Let $g_1 \in C^\infty(\mathbb{R})$ such that $\int_{\mathbb{R}} g_1 = 1$, $g_1 \geq 0$ and $\operatorname{supp} g_1 \subseteq [-1,1]$, a Gaussian distribution function for example, and let $g_m(x) = m g_1(mx)$ be an approximate delta function with $\operatorname{supp} g_m \subseteq [-\frac{1}{m}, \frac{1}{m}]$. With this we can define

$$f_m(x) = \int\limits_{\operatorname{supp} g_m} f(x-y) g_m(y) \mathrm{d}y$$

For all $y \in \operatorname{supp} g_m$, $f(\cdot - y) \in \mathfrak{M}_n(a + \frac{1}{m}, b - \frac{1}{m})$, hence $f_m \in \mathfrak{M}_n(a + \frac{1}{m}, b - \frac{1}{m})$. With a $t = x - y$ substitution, we can see that

$$\mathrm{D}f_m(x) = \mathrm{D} \int\limits_{\operatorname{supp} g_m} f(t) g_m(x-t) \mathrm{d}t = \lim_{h \to 0} \int\limits_{\operatorname{supp} g_m} f(t) \frac{g_m(x-t+h) - g_m(x-t)}{h} \mathrm{d}t$$
$$= \int\limits_{\operatorname{supp} g_m} f(t) \mathrm{D}g_m(x-t) \mathrm{d}t$$

The last equation can be justified with the Dominated convergence theorem. Combining this with $g_m \in C^\infty(\mathbb{R})$, we see that $f_m \in C^\infty(a + \frac{1}{m}, b - \frac{1}{m})$.

For every $[c,d] \subseteq (a,b)$ compact subinterval there exists $m \in \mathbb{N}^+$ such that $[c - \frac{1}{m}, d + \frac{1}{m}] \subseteq (a,b)$, which is a compact set, thus f is uniformly continuous on it. This implies that for every $\varepsilon > 0$ there exists $m \in \mathbb{N}^+$ such that $\forall x \in [c,d]\ \forall y \in [-\frac{1}{m}, \frac{1}{m}] \quad |f(x-y) - f(x)| < \varepsilon$. Hence $\forall x \in [c,d]\ \forall k \geq m$

$$|f_k(x) - f(x)| = \left| \int\limits_{\operatorname{supp} g_k} f(x-y)g_k(y) - f(x)g_k(y)\mathrm{d}y \right|$$
$$\leq \int\limits_{\operatorname{supp} g_k} |f(x-y) - f(x)| g_k(y)\mathrm{d}y < \varepsilon \int\limits_{\operatorname{supp} g_k} g_k(y)\mathrm{d}y = \varepsilon$$

hence $f_m \rightrightarrows f$ on $[c,d]$.

Since $\mathrm{D}f_k = (\mathrm{D}f)_k$, $f \in C^k(a,b) \Rightarrow \forall j \in \{0,\ldots,k\}\ \mathrm{D}^j f \in C(a,b) \Rightarrow \mathrm{D}^j f_m \rightrightarrows \mathrm{D}^j f$ on every $[c,d] \subseteq (a,b)$. ■

Definition 2.3.4 (Schur product) *For any $A, B \in M_{n\times m}[\mathbb{C}]$, the Schur product $(A \odot B)_{ij} = A_{ij}B_{ij}$.*

Lemma 2.3.2 $\forall A, B \in M_n[\mathbb{C}] \quad A, B \geq 0 \Rightarrow A \odot B \geq 0$

Proof: Since $\odot$ is clearly bilinear, and given the spectral theorem, any positive matrix is the linear combination of orthogonal rank one projections with non-negative coefficients, it is enough to check for any two rank one projections, which take the form

$$P^\varphi = \varphi\overline{\varphi}^T, \quad P^\varphi_{ij} = \varphi_i\overline{\varphi}_j$$

But in this case $P^\varphi \odot P^\vartheta = P^{\varphi\odot\vartheta}$, which is clearly positive since it is a projection.

■

Theorem 2.3.1 (Daleckiĭ-Kreĭn formula) *Let $a < x_1 < \ldots < x_n < b$, $f \in C^1(a,b)$, let $C \in M_n[\mathbb{C}]$ be a self-adjoint matrix and define*

$$A = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix} \tag{2.18}$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f(A + \lambda C)\bigg|_{\lambda=0} = L_n(x_1, \dots, x_n; f) \odot C \tag{2.19}$$

The presented proof is due to Loewner.

Proof: From the definition of A, one can notice that e_k is an eigenvector of A with eigenvalue x_k for all $k \leq n$. Since the eigenvalues $x_k(\lambda)$ and eigenvectors $v_k(\lambda)$ are continuous in λ in finite dimensions, for sufficiently small λ, we have

$$(A + \lambda C) v_k(\lambda) = x_k(\lambda) \quad x_k(0) = x_k \quad v_k(0) = e_k \quad x_k(\lambda) \neq_{k \neq l} x_l$$

Since $\mathbf{Ran} f \subseteq \mathbb{R}$, $f(A + \lambda C)$ is self-adjoint, thus

$$\begin{aligned} \langle v_k(\lambda) \,|[f(A + \lambda C) - f(A)] e_l \rangle &= \langle f(A + \lambda C) v_k(\lambda) \,| e_l \rangle - \langle v_k(\lambda) \,| f(A) e_l \rangle \\ &= [f(x_k(\lambda)) - f(x_l)] \, \langle v_k(\lambda) \,| e_l \rangle \end{aligned}$$

In the special case of $f = id$, $\langle v_k(\lambda) \,| \lambda C e_l \rangle = (x_k(\lambda) - x_l) \, \langle v_k(\lambda) \,| e_l \rangle$. Combining these two, we see that

$$\left\langle v_k(\lambda) \middle| \frac{f(A + \lambda C) - f(A)}{\lambda} e_l \right\rangle = \frac{f(x_k(\lambda)) - f(x_l)}{x_k(\lambda) - x_l} \, \langle v_k(\lambda) \,| C e_l \rangle$$

Considering that the R.H.S. is continuous in λ, we can take $\lambda \to 0$ to arrive at

$$\left[\frac{\mathrm{d}}{\mathrm{d}\lambda} f(A + \lambda C)\bigg|_{\lambda=0} \right]_{kl} = [x_k, x_l, f] C_{kl} = [L_n(x_1, \dots, x_n; f) \odot C]_{kl}$$

■

One can notice, that $\frac{\mathrm{d}}{\mathrm{d}\lambda} f(A+\lambda C)\Big|_{\lambda=0}$ is the Gâteaux derivative of f at A in the C direction, $\mathrm{D}f(A)(C)$. In partivular, if we restrict a $C^1(a,b)$ function onto $M_n[\mathbb{C}]_{sa}$, then f is Gâteaux differentiable in all directions at all points (matrices), that have distinct eigenvalues in (a,b). The following theorem comes from Loewner as well.

Theorem 2.3.2 *For any $f:(a,b)\to\mathbb{R}$ function, the following are equivalent:*

(a) $f \in \mathfrak{M}_n(a,b)$

(b) For each $a < x_1 < x_2 < \ldots < x_n < b$,

$$L_n(x_1,\ldots,x_n;f) \geq 0 \tag{2.20}$$

Proof: $(a) \Rightarrow (b)$: Pick any set of x_k in strictly ascending order as in (b), define A as (2.18), and take a $C \geq 0$. By first handling the $\lambda > 0$ case, we can see, that $A+\lambda C \geq A \Rightarrow f(A+\lambda C) \geq f(A) \Rightarrow \frac{f(A+\lambda C)-f(A)}{\lambda} \geq 0$. We arrive at the same conclusion in the $\lambda < 0$ case, so

$$0 \leq \frac{\mathrm{d}}{\mathrm{d}\lambda} f(A+\lambda C)\Big|_{\lambda=0} = L_n(x_1,\ldots,x_n;f)\odot C$$

If we define $P^v = v\overline{v}^T \geq 0$ for $v\in\mathbb{C}^n$ and notice that every positive operator has a non-negative trace by definition, the predecessing equation yields

$$\forall v\in\mathbb{C}^n \quad 0 \leq \mathrm{Tr}(L_n(x_1,\ldots,x_n;f)\odot P^v) = \langle v\,|L_n(x_1,\ldots,x_n;f)v\rangle$$

Thus $L_n(x_1,\ldots,x_n;f)\geq 0$ holds by definition.

$(b) \Rightarrow (a)$: By lemma 2.3.2, we have

$$\forall C\geq 0 \quad \mathrm{D}f(A)(C) = \frac{\mathrm{d}}{\mathrm{d}\lambda} f(A+\lambda C)\Big|_{\lambda=0} \geq 0$$

Holds for any A in the form of (2.18). This implies that for any $B \geq A$, if $A+\lambda(B-A)$ has distinct eigenvalues for every $\lambda \in [0,1]$, then $f(A) \leq f(B)$, since $f(B) = f(A) + \int_0^1 \mathrm{D}f(A+\lambda(B-A))(B-A)\mathrm{d}\lambda$ by the Fundamental Theorem of Calculus for Gâteaux derivatives.

If A has distinct eigenvalues, by the finality of the dimensions and the continuity of the eigenvalues in λ, we have that $C(\lambda) = A + \lambda(B-A)$ has distinct eigenvalues outside a finite $0 < \lambda_1 < \cdots < \lambda_l < 1$ set, and maybe $\lambda_{l+1} = 1$. Since they are distinct, for any j we can choose $\varepsilon > 0$ such that $\lambda_j + \varepsilon < \lambda_{j+1} - \varepsilon$, so $f(C(\lambda_j+\varepsilon)) \leq f(C(\lambda_{j+1}-\varepsilon))$. By taking $\varepsilon \searrow 0$, by the continuity of f, we get $f(C(\lambda_j)) \leq f(C(\lambda_{j+1}))$, thus $f(A) \leq f(B)$ even if only A has distinct eigenvalues.

Since we can approximate any A with eigenvalues in (a,b) with such matrices A_m with distinct eigenvalues, and $f(A_m) \leq f(A_m + (B-A))$ for m large enough by the previous case, we still arrive at $f(A) \leq f(B)$ in the general case. ∎

Definition 2.3.5 *For any $n \in \mathbb{N}^+$, $f \in C^{2n-1}(a,b)$, $x \in (a,b)$, we define the $n \times n$ Dobsch matrix of f at x,*

$$B_n(x;f)_{ij} = \frac{D^{i+j-1}f(x)}{(i+j-1)!}$$

For any $n \in \mathbb{N}^+$, $f \in C(a,b)$, $x_1,\ldots,x_n \in (a,b)$ we define the $n \times n$ multipoint Loewner matrix

$$A_n(x_1,\ldots,x_n;f)_{ij} = [x_1,\ldots,x_i,x_1,\ldots,x_j,f]$$

The mutipoint Loewner matrix is often used to approximate the Dobsch matrix, since $A_n(x,\ldots,x;f) = B_n(x;f)$, and the kth divided difference is continuous by (2.3.1). The convergence here is the operator norm convergence, but in finite dimension, it is equivalent to the entrywise convergence which is easier to notice.

We call a measure trivial, if it has finite points of support.

Lemma 2.3.3 *If μ is a nontrivial finite measure $\mathbb{R} \setminus (a,b)$ and $f : (a,b) \to \mathbb{R}$ is defined by*

$$f(x) = \int\limits_{\mathit{supp}\,\mu} (y-x)^{-1} \mathrm{d}\mu(y) \tag{2.21}$$

then $A_n(x_1, \ldots, x_n; f)$ is strictly positive for each $x_1, \ldots, x_n \in (a,b)$.

Proof: One can see by induction, that for any $y \notin (a,b)$, $[x_1, \ldots, x_k; (y-x)^{-1}] = \prod_{j=1}^{k} (y-x_j)^{-1}$. With this in consideration, we can see that

$$A_n(x_1, \ldots, x_n; f)_{ij} = \int\limits_{\mathrm{supp}\,\mu} \prod_{k=1}^{i} (y-x_k)^{-1} \prod_{l=1}^{j} (y-x_l)^{-1} \mathrm{d}\mu(y)$$

Let $v \in \mathbb{C}^n$

$$\begin{aligned} \langle v \,| A_n(x_1, \ldots, x_n; f) v \rangle &= \sum_{i,j=1}^{n} \overline{v}_i A_n(x_1, \ldots, x_n; f)_{ij} v_j \\ &= \int\limits_{\mathrm{supp}\,\mu} |\sum_{i=1}^{n} v_i \prod_{j=1}^{i} (y-x_j)^{-1}|^2 \mathrm{d}\mu(y) \end{aligned}$$

which is strictly positive for any non-zero vector, since $\mathrm{d}\mu$ is nontrivial. ∎

Theorem 2.3.3 *If $f \in \mathfrak{M}_n(a,b) \cap C^{2n-1}(a,b)$, then $B_n(x,f) \geq 0$ for all $x \in (a,b)$*

Proof: Let us take $x_1, \dots, x_n \in (a,b)$. For space efficiency, we shorthand L for $L_n(x_1, \dots, x_n; f)$. Since the determinant is invariant under additive transformations of the row space and column space, and we can bring scalar multiplicators of rows and columns in front of it, we substract the first row from the remaining ones to get

$$\det L = \prod_{j\geq 2}(x_j - x_1) \det A^{(1)} \quad A^{(1)}_{ij} = \begin{cases} [x_1, x_j; f] & if \ i = 1 \\ [x_1, x_i, x_j; f] & if \ i \geq 2 \end{cases}$$

Now substract the second row from the rows of greater index to get

$$\det L = \prod_{j\geq 2}(x_j - x_1) \prod_{j\geq 3}(x_j - x_2) \det A^{(2)} \quad A^{(2)}_{ij} = \begin{cases} [x_1, x_j; f] & if \ i = 1 \\ [x_1, x_2, x_j; f] & if \ i = 2 \\ [x_1, x_2, x_i, x_j; f] & if \ i \geq 3 \end{cases}$$

Repeating this algorithm, we arrive at

$$\det L = \prod_{i<j}(x_j - x_i) \det A^{(n)} \quad A^{(n)}_{ij} = [x_1, \dots, x_i, x_j; f]$$

If we use the same argument on this equation with the columns, we get

$$\det L = \prod_{i<j}(x_j - x_i)^2 \det A_n(x_1, \dots, x_n; f)$$

By the predecessing theorem, this implies $\det A_n(x_1, \dots, x_n; f) \geq 0$. One can choose a g in a form of (2.21) with μ being the Lebesgue measure on some $[c,d]$, and define

$$d_k(t) = d_k(A_n(x_1, \dots, x_n; tf + (1-t)g))$$

By the previous lemma and (2.3.1), $d_k(0) > 0$ for all $k \leq n$. Since $\mathfrak{M}_n(a,b)$ is convex, $d_k(t) \geq 0$ for all $k \leq n$, but $d_k(t)$ is a polynomial of t, so it has finite roots in $[0, 1)$. If we gather all the roots for each $k \le n$, which is still a finite set, and examine its complement, $\mathfrak{T}$ in $[0, 1)$, we see that $A_n(x_1, \ldots, x_n; tf + (1-t)g) > 0$ for all $t \in \mathfrak{T}$. Taking $t \to 1$ in $\mathfrak{T}$, we see that $tf + (1-t)g \rightrightarrows f$, so $A_n(x_1, \ldots, x_n; tf + (1-t)g) \to A_n(x_1, \ldots, x_n; f)$, thus it is positive.

By taking $x_1, \ldots, x_n \to x$, we have that $A_n(x_1, \ldots, x_n; f) \to B_n(x; f)$, hence it is positive as well. The last convergence used the diffenetiability of f. ■

Definition 2.3.6 (Distributions) *A distribution on an interval (a, b) is a linear functional of the space $C_0^\infty((a, b))$*

Any $f : (a, b) \to \mathbb{R}$ locally integrable, measurable function induces a ditribution on (a, b) by

$$T_f(g) = \int_a^b g(x)f(x)dx$$

Definition 2.3.7 *A distribution T is positive, if $f \ge 0 \Rightarrow T(f) \ge 0$.*
For any distribution T, the distributional derivative of T is defined by $\mathrm{D}T(f) = T(-\mathrm{D}f)$.

One can notice, that for any $f \in C^1(a, b)$, $\mathrm{D}T_f = T_{\mathrm{D}f}$, which comes from integrating by parts.

Proposition 2.3.3 *Every $f \in \mathfrak{M}_1(a, b)$ is a measurable function.*
If $f \in \mathfrak{M}_n(a, b)$, then $\mathrm{D}^{2n-1}f$, the $(2n-1)st$ distributional derivative is positive.

Proof: Let $f \in \mathfrak{M}_1(a,b)$, so a monotone function. If we define $q(c) = \sup\{x \in (a,b) \mid f(x) < c\}$ and $p(d) = \sup\{x \in (a,b) \mid f(x) > d\}$, we see that

$$f^{-1}((-\infty, c)) = (a, q(c)) \textit{ or } (a, q(c)], \quad f^{-1}((d, \infty)) = (p(d), b) \textit{ or } [p(d), b)$$

which are all Borel sets, so the preimage of any interval, hence of any open set (, since any open set is a countable union of intervals), thus of any Borel set is a Borel set.

If $f \in \mathfrak{M}_n(a,b)$ then by lemma 2.3.2, we have f_m smooth functions converging pointwise on (a,b), so the distributions T_m of f_m converge to T of f. By the former theorem, $B_n(x; f_m) \geq 0$ for all m and $x \in (a,b)$, so $B_n(x; f_m)_{nn} = \dfrac{\mathrm{D}^{2n-1} f(x)}{(2n-1)!} \geq 0$, so $\mathrm{D}^{2n-1} f \geq 0$, thus $\mathrm{D}^{2n-1} T_m$. By the definition of the distributional derivative and the convergence, $T_m \to T \quad \Leftrightarrow \quad \mathrm{D}^k T_m \to \mathrm{D}^k T \ \forall k \in \mathbb{N}$, hence $\mathrm{D}^{2n-1} T_m \to \mathrm{D}^{2n-1} T$, thus $\mathrm{D}^{2n-1} T \geq 0$. ■

Lemma 2.3.4 *If T is a distribution on (a,b) with $\mathrm{D}^2 T \geq 0$, then T is induced by a continuous function.*

Proof: Since $\mathrm{D}^2 T$ is a positive linear functional of $((a,b), \tau_{(a,b)})$, which is a locally compact Hausdorff space, there exists a μ regular Borel measure due to the Riesz-Markov Theorem, such that $\forall f \in C_0^\infty(a,b) \quad \mathrm{D}^2 T(f) = \int_a^b f \mathrm{d}\mu$. Set

$$g(x) = \int_a^x \mu([a,x]) dx$$

so $D^2 g = \dfrac{\mathrm{d}\mu}{\mathrm{d}\lambda}$, the Radon-Nykodim derivative of μ. The function g is clearly continuous, and we have that $T(f) = \int_{\mathbb{R}} f(x) g(x) dx$ ■

This lemma combined with the preceding proposition tells us that any $f \in \mathfrak{M}_n(a,b)$ agrees with a C^{2n-3} function almost everywhere with non-negative odd order derivatives. Considering the fact, that if a continuous and a monotone function agrees almost everywhere, then they are identical, we see that

Corollary 2.3.1 $\mathfrak{M}_n(a,b) \subseteq C^{2n-3}(a,b)$*, in particular* $\mathfrak{M}_\infty(a,b) \subseteq C^\infty(a,b)$*. Furthermore,* $\mathrm{D}^{2k-3}f(x) \geq 0$ *for all* $k \in \{2,\dots,n\}$ *and* $x \in (a,b)$*.*

Theorem 2.3.4 (Bernstein's Theorem) *If* $f \in C^\infty(-1,1)$ *obeys* $D^k f(x) \geq 0 \quad \forall x \in (-1,1)\ \forall k \in \mathbb{N}^+$*, then* f *is a restriction of a function analytic on* $\mathbb{D}$*.*

Proof: Let $T_n f$ be the Taylor approximation of order n about 0, that is

$$T_n f(x) = \sum_{k=1}^{n} \frac{\mathrm{D}^k(0)}{k!} x^k$$

By writing $f(x) = f(y) + \int_0^x \mathrm{D}f(t)\mathrm{d}t$, and using it repeatedly to expand the integrand, we get

$$R_n(x) \equiv f(x) - T_n f(x) = \int_0^x \mathrm{d}x_1 \int_0^{x_1} \mathrm{d}x_2 \dots \int_0^{x_n} \mathrm{d}x_{n+1}\ \mathrm{D}^{n+1} f(x_{n+1}) \quad (2.22)$$

$$= \int_0^x \mathrm{d}x_{n+1} \int_{x_{n+1}}^x \mathrm{d}x_n \dots \int_{x_2}^x \mathrm{d}x_{n+1}\ \mathrm{D}^{n+1} f(x_{n+1}) \quad (2.23)$$

$$= \int_0^x \mathrm{D}^{n+1} f(x_{n+1}) \frac{(x - x_{n+1})^n}{n!} \mathrm{d}x_{n+1} \quad (2.24)$$

Since there is only one variable of integration is left, we can write $t = x_{n+1}$.

At first, let us limit ourselves to the $x \geq 0$ case. By the criterion of the theorem, $T_n f(x) \geq 0$, thus $R_n(x) \leq f(x)$, and $R_n(x) \geq 0$ as well. If we fix an $\alpha \in (0,1)$, then $\dfrac{x-t}{\alpha-t} \leq x$ holds for all $0 \leq t \leq x \leq \alpha$. Thus we can write

$$\begin{aligned}
0 \leq R_n(x) &= \int_0^x \frac{\mathrm{D}^{n+1} f(t)}{n!}(x-t)^n \mathrm{d}t \\
&= \int_0^x \frac{\mathrm{D}^{n+1} f(t)}{n!}(\alpha-t)^n \frac{(x-t)^n}{(\alpha-t)^n} \mathrm{d}t \\
&\leq x^n \int_0^x \frac{\mathrm{D}^{n+1} f(t)}{n!}(\alpha-t)^n \mathrm{d}t \leq x^n \int_0^\alpha \frac{\mathrm{D}^{n+1} f(t)}{n!}(\alpha-t)^n \mathrm{d}t \\
&= x^n R_n(\alpha) \leq x^n f(\alpha)
\end{aligned}$$

Hence we see that $R_n \rightrightarrows 0$ on $[0,\alpha]$.

If $-\alpha \leq x \leq t \leq 0$, then $T_n f(x) \leq 0$, thus $R_n(x) \leq 0$. In this case $\dfrac{x-t}{\alpha-t} \geq x$ holds, thus $R_n(x)$ can be upper estimated by $x^n(f\alpha)$, Therefore $R_n \rightrightarrows 0$ on $[-\alpha, 0]$ as well. Since $|f(x) - T_n f(x)| = |R_n(x)| \rightrightarrows 0$ on $[-\alpha,\alpha]$, $T_n f \rightrightarrows f$ on $[-\alpha,\alpha]$, but since $\alpha \in (0,1)$ was arbitrary, we conclude that $g(x) = \sum_{k\in\mathbb{N}} \dfrac{\mathrm{D}^k f(0)}{k!} x^k = f(x)$ on $(-1,1)$, which also mean that $\sum_{k\in\mathbb{N}} \dfrac{\mathrm{D}^k f(0)}{k!} z^k$ is absolute convergent on the entire $\mathbb{D}$, thus it is convergent, defining an analytic function.

■

The integral (2.24) is called the integral form of the Taylor's remainder. One can also see that (2.22) is an integral of a continuous function $\mathrm{D}^{n+1} f$ on $\mathrm{conv}\{0,x\}^{n+1}$, thus by the Mean value theorem, there exists a $t_{n,x} \in \mathrm{conv}\{0,x\}$, such that $\mathrm{D}^{n+1} f(t_{n,x})$ multiplied by the volume of $\mathrm{conv}\{0,x\}^{n+1}$

is the $(n+1)$-fold integral. $R_n(x) = \dfrac{D^{n+1}f(t_{n,x})}{(n+1)!}x^{n+1}$. This is known as the Lagrange form of the Taylor's remainder, which will be used in the proof of the following theorem.

Lemma 2.3.5 *If f is a C^2 function on an I interval such that there is a $\delta > 0$ with $[-\delta, \delta] \subseteq I$, then*

$$|Df(0)| \leq \frac{2}{\delta} \sup_{[-\delta,\delta]} |f| + \frac{\delta}{2} \sup_{[-\delta,\delta]} |D^2 f| \tag{2.25}$$

Proof:

$$|f(\delta) - f(0) - \delta \mathrm{D}f(0)| = \left| \int_0^\delta \mathrm{D}^2 f(x)(\delta - x)\mathrm{d}x \right| \leq \frac{\delta^2}{2} \sup_{[-\delta,\delta]} |\mathrm{D}^2 f|$$

The first equation uses integration by parts. By adding $|f(\delta) - f(0)|$ to both sides, and using the triangle inequality to make the L.H.S. smaller and the R.H.S. greater, we get

$$|\delta \mathrm{D}f(0)| \leq |f(\delta)| + |f(0)| + \frac{\delta^2}{2} \sup_{[-\delta,\delta]} |\mathrm{D}^2 f|$$

from which one can acquire (2.25) by dividing with δ and making an upper estimate for $|f(\delta)|$ and $|f(0)|$. ■

Theorem 2.3.5 (Berstein-Boas Theorem) *If $f \in C^\infty(-1,1)$ obeys $D^{2k-1}f(x) \geq 0$ $\forall x \in (-1,1)$ $\forall k \in \mathbb{N}^+$,then f is a restriction of a function analytic on $\mathbb{D}$. In particular, $\forall R > 1$ $\exists C_R > 0 : \forall n \in \mathbb{N}$*

$$\left| \frac{D^n f(0)}{n!} \right| \leq C_R R^n \tag{2.26}$$

Proof: For the upper estimate, one can use Cauchy's differentiation formula on the curve $\partial \mathrm{B}_{R^{-1}}(0) \subseteq \mathbb{D}$, which is a compact set, thus $|f|$ reaches its maximum, K on it, hence

$$\left|\frac{\mathrm{D}^n f(0)}{n!}\right| \leq \left|\frac{1}{2\pi i}\right| \int\limits_{\partial \mathrm{B}_{R^{-1}}(0)} \frac{|f(z)|}{|z^{n+1}|} \mathrm{d}z \leq \frac{1}{2\pi} \int\limits_{\partial \mathrm{B}_{R^{-1}}(0)} \frac{K}{R^{-(n+1)}} \mathrm{d}z = (KR)R^n$$

Let $g(x) = f(x) - f(-x)$, so $\forall k \in \mathbb{N}^+ \quad \mathrm{D}^{2k-1} g(x) \geq 0,\ \mathrm{D}^{2k} g(0) = 0$, so we can use the same arguments for g, as in the first half Bernstein's theorem. Firstly, if $x \in (0, 1)$,

$$\sum_{k=1}^{n} \frac{\mathrm{D}^{2k-1} g(0)}{(2k-1)!} x^{2k-1} = \sum_{j=1}^{2n} \frac{\mathrm{D}^j g(0)}{j!} x^j \leq g(x)$$

Notice that $\mathrm{D}^{2k-1} f(0) = \frac{1}{2} \mathrm{D}^{2k-1} g(0)$, from which we can see that

$$\sum_{k=1}^{n} \frac{\mathrm{D}^{2k-1} f(0)}{(2k-1)!} x^{2k-1} < \infty \Rightarrow R^{-1} = \limsup_{k \to \infty} \left(\frac{\mathrm{D}^{2k-1} f(0)}{(2k-1)!} \right)^{\frac{1}{2k-1}} \leq 1$$

where R is the radius of convergence of the power series on the left. The implication follows from the Cauchy-Hadamard Theorem. Let $\alpha \in (0, 1)$, $x_0 \in [-\alpha, \alpha]$, $h(x) = f(x) - f(2x_0 - x)\ \forall x \in (x_0 - (1 - |x_0|), x_0 + (1 - |x_0|))$, so

$$\forall k \in \mathbb{N}^+ \quad \mathrm{D}^{2k-1} h(x_0) \geq 0,\ \mathrm{D}^{2k} h(x_0) = 0$$

thus we can use Taylor's approximation about x_0. If there exists an $x > 0$ such that $x \in \mathbf{Dom}\, h$, we can use the exact same argument, as for g. If $\mathbf{Dom}\, h \subseteq (-3, 0]$, we can find an $x < 0$ in it, thus we can use almost the same reasoning, with Taylor's approximation being a monotone decreasing sequence that has a lower bound $h(x)$. In either case, we arrive at

$$\forall \alpha \in (0, 1),\ x_0 \in [-\alpha, \alpha] \quad \limsup_{k \to \infty} \left(\frac{\mathrm{D}^{2k-1} f(x_0)}{(2k-1)!} \right)^{\frac{1}{2k-1}} \leq (1 - |x_0|)^{-1}$$

By combining the previous lemma and $\mathrm{D}^{2k}f = \mathrm{D}\mathrm{D}^{2k-1}f$ we see that

$\forall x_0 \in (-1,1)\ \forall \delta \in (0, 1-|x_0|)$

$$\begin{aligned}
\limsup_{k\to\infty}\left(\frac{\mathrm{D}^{2k}f(x_0)}{(2k+1)!}\right)^{\frac{1}{2k}} &\le \limsup_{k\to\infty}\left(\frac{2}{\delta}\sup_{\mathrm{B}_\delta(x_0)}\frac{\mathrm{D}^{2k-1}f}{(2k+1)!}+\frac{\delta}{2}\sup_{\mathrm{B}_\delta(x_0)}\frac{\mathrm{D}^{2k+1}f}{(2k+1)!}\right)^{\frac{1}{2k}}\\
&\le \limsup_{k\to\infty}\left(\frac{2}{\delta}\sup_{\mathrm{B}_\delta(x_0)}\frac{\mathrm{D}^{2k-1}f}{(2k-1)!}+\frac{\delta}{2}\sup_{\mathrm{B}_\delta(x_0)}\frac{\mathrm{D}^{2k+1}f}{(2k+1)!}\right)^{\frac{1}{2k}}\\
&\le \limsup_{k\to\infty}\left(\left(\frac{2}{\delta}+\frac{\delta}{2}\right)\max_{i\in\{-1,1\}}\sup_{\mathrm{B}_\delta(x_0)}\frac{\mathrm{D}^{2k+i}f}{(2k+i)!}\right)^{\frac{1}{2k}}\\
&\le \limsup_{k\to\infty}\left(\max_{i\in\{-1,1\}}\sup_{\mathrm{B}_\delta(x_0)}\frac{\mathrm{D}^{2k+i}f}{(2k+i)!}\right)^{\frac{1}{2k}}\left(\frac{2}{\delta}+\frac{\delta}{2}\right)^{\frac{1}{2k}}\\
&\le (1-\delta-|x_0|)^{-1}
\end{aligned}$$

Where the last inequality uses $\frac{2k+i}{2k} \to 1$ in both cases of the maximum. Taking $\delta \searrow 0$, we have the same estimate as for the odd case. If $x \in (-\frac{1}{2}, \frac{1}{2})$, the estimates show us that $|x|^n \sup_{\mathrm{B}_{|x|}(0)} \frac{\mathrm{D}^n f(0)}{n!} \to 0$, which means that the Lagrange remainder of the Taylor's approximation goes to zero uniformly on all $[-\alpha, \alpha] \subseteq (-\frac{1}{2}, \frac{1}{2})$. Thus f is real analytic on $(-\frac{1}{2}, \frac{1}{2})$. Notice, that for all $x \in (-1,1)$, If we restrict f to $(x-(1-|x|), x-(1-|x|))$, the same argument shows us that f is real analytic on $(x-\frac{1}{2}(1-|x|), x-\frac{1}{2}(1-|x|))$. Hence f is real analytic on $(-1,1)$, and by the uniqueness of the analytic continuation, f is the restriction of $g(x) = \sum_{k\in\mathbb{N}} \frac{\mathrm{D}^k f(0)}{k!} x^k$, which is convergent on $\mathbb{D}$, for it is absolute convergent. ■

Corollary 2.3.2 *Every $f \in \mathfrak{M}_\infty(-1,1)$ has an analytic continuation onto $\mathbb{D}$, and if*

$$c_n \equiv \frac{D^{n+1}f(0)}{(n+1)!} \tag{2.27}$$

then $\forall R > 1 \; \exists C_R > 0 : \forall n \in \mathbb{N}$

$$|c_n| \leq C_R R^n \tag{2.28}$$

Proof: By Corollary 2.3.1, we see that f satisfies the condition of the Bernstein-Boas Theorem, which has almost the same consequences, but with $|c_n| \leq \tilde{C_R} R^{n+1}$. By setting $C_R = R\tilde{C_R}$, one can see that (2.28) also holds. ■

Lemma 2.3.6 *Let* $R > 0$*. Then all* $P \in \mathbb{C}[x]$ *polynomial obeying* $P([-R, R]) \subseteq \mathbb{R}_0^+$ *is the finite sum of terms in the form of*

$$Q(x)^2, \qquad (R - x)Q(x)^2, \qquad (R + x)Q(x)^2, \qquad (R^2 - x^2)Q(x)^2 \tag{2.29}$$

where $Q \in \mathbb{R}[x]$ *polynomial.*

In both the statement and the proof we use $\mathbb{R}[x]$ and $\mathbb{C}[x]$, which indicates whether a polynomial has real or complex coefficients. However, this denotation is not accurate, because an element of $\mathbb{C}[x]$ is not a function until it is restricted to a domain. Hence in the following, we will use $\mathbb{R}[x]$ and $\mathbb{C}[x]$ as $\mathbb{R}[x]|_{\mathbb{C}}$ and $\mathbb{C}[x]|_{\mathbb{C}}$.

Proof: Let $P(x) = \sum_{k=0}^{n} a_n x^n$, and notice, that by the condition, $P \in \mathbb{R}[x]$, since $P(0) = a_0 \in \mathbb{R}$ and $F(x) = \frac{P(x)}{x} - a_0 \in \mathbb{C}[x]$ with $F([-R, R]) \in \mathbb{R}$, so $F(0) = a_1 \in \mathbb{R}$, and so on. This implies, that for all $z \in \mathbb{C} \setminus \mathbb{R}$, z and $\overline{z}$ is a root of the same multiplicity, if any of them is a zero of P. Also any $y \in [-R, R]$ has even multiplicity, since if $g(x) = \frac{P(x)}{x - y} \in \mathbb{C}[x]$ then $g((y, R]) \subseteq \mathbb{R}_0^+$, $\quad g([-R, y)) \subseteq \mathbb{R}_0^-$, thus $g(y) = 0$. With these in mind, we can write

$$P(x) = C \prod_{z_j \in \mathbb{C}_+} |x - z_j|^2 \prod_{x_j \in [-R,R]} (x - x_j)^2 \prod_{y_j < -R} (x - y_j) \prod_{w_j > R} (w_j - x)$$

with the indexed numbers being the roots, and $C > 0$. We can punctuate the complex terms by writing $z_j = a_j + ib_j$, $|x - z_j|^2 = (x - a_j)^2 + b_j^2$, hence we can write $P(x)$ as a finite sum of terms in the form of

$$Q(x)^2 \prod_{y_j < -R} (x - y_j) \prod_{w_j > R} (w_j - x)$$

with $Q \in \mathbb{R}[x]$. By expanding $(x - y_j) = (-R - y_j) + (R + x)$ and $(w_j - x) = (R - x) + (w_j - R)$, we get a finite sum of term such as

$$Q(x)^2 (R + x)^n (R - x)^m$$

where n is the number of roots less than $-R$ and m is the number of zeros greater than R. If n and m are even, this simplifies to $Q(x)^2$, $Q(x)^2(R^2 - x^2)$ if both are odd, and $Q(x)^2(R - x)$ or $Q(x)^2(R + x)$ otherwise. ■

Theorem 2.3.6 (Hausdorff Moment Problem) *If* $\{c_n \mid n \in \mathbb{N}\} \subseteq \mathbb{R}$ *and* $R > 0$, *then*

$$\exists! \mu \text{ finite measure on } [-R, R], \text{ such that } \quad c_n = \int_{-R}^{R} x^n \mathrm{d}\mu(x) \qquad (2.30)$$

if and only if

(a) *(2.28) Holds for* R, *and*

(b) *for all* $n \in \mathbb{N}^+$, *the* $n \times n$ *Hankel matrix,* $H_{ij}^{(n)} = c_{i+j-2}$ *is positive.*

Proof: If (2.30) holds, then for any $a \in \mathbb{C}^{n+1}$, define $P_a(x) = \sum_{k=0}^{n} a_k x^k$ so

$$0 \le \langle P_a \,|P_a\rangle_{L^2(\mu)} = \sum_{i,j=0}^{n} \overline{a_i} a_j \int_{-R}^{R} x^{i+j} \mathrm{d}\mu(x) = \sum_{i,j=1}^{n+1} \overline{a_{i-1}} a_{j-1} c_{i+j-2} = \left\langle a \,\middle|\, H^{n+1} a \right\rangle$$

holds, thus $H^{n+1} \geq 0$ for all $n \geq 0$ by definition. We can also notice that

$$|c_n| \leq \int_{-R}^{R} |x^n| \mathrm{d}\mu(x) \leq R^n \int_{-R}^{R} \mathrm{d}\mu(x) \leq R^n c_0$$

It is clear, that $\mathrm{d}\mu$ is unique on the space of polynomials, as

$$\int_{-R}^{R} \sum_{k=0}^{n} a_k x^k \mathrm{d}\mu(x) = \sum_{k=0}^{n} a_k c_k$$

The polynomials are dense in $C[-R, R]$, so $\mathrm{d}\mu$ is unique.

For the converse, we examine $(C[-R, R], \|\cdot\|_\infty)$. Firstly, we define

$$\mathfrak{f}\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=0}^{n} a_k c_k$$

We show, that $P([-R, R]) \subseteq [0, \infty) \Rightarrow \mathfrak{f}(P) \geq 0$
By the previous lemma, such polynomials are a finite sum of terms in the form of (2.29) with $Q \in \mathbb{R}[x]$. $\mathfrak{f}$ defines a scalar product on the real polynomials with $\langle P_1 | P_2 \rangle_\mathfrak{f} = \mathfrak{f}(P_1 P_2)$. One can see that $0 \leq \left\langle q \middle| H^{n+1} q \right\rangle = \langle Q | Q \rangle_\mathfrak{f} = \mathfrak{f}(Q^2)$, where the coefficients of Q are the elements of q. By the Cauchy-Schwartz-Bunyakovsky inequality, we have $\mathfrak{f}(PQ) \leq \mathfrak{f}(P^2)^{1/2} \mathfrak{f}(Q^2)^{1/2}$, and we can use it repeatedly to see, that for any $Q \in \mathbb{R}[x]$

$$\mathfrak{f}(xQ^2) \leq \mathfrak{f}(x^2Q^2)^{1/2}\mathfrak{f}(Q^2)^{1/2} \leq \mathfrak{f}(x^4Q^2)^{\frac{1}{4}}\mathfrak{f}(Q^2)^{\frac{3}{4}} \leq \mathfrak{f}(x^{2^n}Q^2)^{\frac{1}{2^n}}\mathfrak{f}(Q^2)^{1-\frac{1}{2^n}}$$

Note, that if $Q^2(x) = \sum_{k=0}^{m} q_k x^k$, then

$$|\mathfrak{f}(x^{2^n}Q^2)| = \sum_{k=0}^{m} |q_k||c_{k+2^n}| \leq \sum_{k=0}^{m} |q_k| C_R R^{k+2^n} \leq m \max_{j \leq m} |q_j C_R R^j| R^{2^n} = K_{R,Q} R^{2^n}$$

Using (2.28). This implies $\limsup_{n\to\infty} |\mathfrak{f}(x^{2^n}Q^2)|^{\frac{1}{2^n}} \leq R$, which in combination with the iterated C-S-B inequality yields $|\mathfrak{f}(xQ^2)| \leq R\mathfrak{f}(Q^2)$. Since $\mathfrak{f}$ is linear, this means $\mathfrak{f}((R+x)Q^2), \mathfrak{f}((R-x)Q^2) \geq 0$. One can use the same reasoning to arrive at $\mathfrak{f}(x^2Q^2) \leq \mathfrak{f}(x^{2\cdot 2^n}Q^2)^{\frac{1}{2^n}}\mathfrak{f}(Q^2)^{1-\frac{1}{2^n}}$, $|\mathfrak{f}(x^2Q^2)| \leq R^2\mathfrak{f}(Q^2)$ and finally $\mathfrak{f}((R^2-x^2)Q^2) \geq 0$, so we proved our first aim.

For any $P \in \mathbb{R}[x]$ real polynomial, we have $\|P\| \pm P \geq 0$, therefore $\mathfrak{f}(\|P\| \pm P) \geq 0$, hence $|\mathfrak{f}(P)| \leq \|P\|c_0$. This means that $\mathfrak{f}$ is continuous on the space of polynomials, which are dense in $C[-R,R]$. Thus $\mathfrak{f}$ extends to $\mathfrak{F}$ on continuous functions with $f > 0 \Rightarrow \mathfrak{F}(f) \geq 0$, since if $f > 0$, then $\min_{[-R,R]} f = x_0 > 0$, and we can approximate f with polynomials in $\mathrm{B}_{x_0}(f)$ which are all non-negative, and $\mathfrak{F}$ is continuous. Again by the continuity, if we write $f = \lim_{\varepsilon \searrow 0} \varepsilon\mathbf{1} + f$, we see that

$$f \geq 0 \Rightarrow \mathfrak{F}(f) \geq 0$$

In conclusion, $\mathfrak{F}$ is a continuous positive linear functional of $C[-R,R]$, so by the Riesz-Markov theorem

$$\exists\mu \textit{ measure on } [-R,R], \textit{ such that } \quad \mathfrak{F}(f) = \int_{-R}^{R} f \mathrm{d}\mu$$

By the definition of $\mathfrak{f}$, $c_n = \int x^n \mathrm{d}\mu$ holds, and we already concluded, that if there exists such μ, it is unique. Since $c_0 = \mu([-R,R])$, μ is a finite measure.

■

Theorem 2.3.7 $\forall f \in \mathfrak{M}_\infty(-1,1) \quad \exists!\nu$ *finite measure on* $[-1,1]$, *such that*

$$f(x) = f(0) + \int_{-1}^{1} \frac{x}{1+\lambda x} d\nu(\lambda) \tag{2.31}$$

Proof: Let $c_n = \dfrac{\mathrm{D}^{n+1}f(0)}{(n+1)!}$, so $H_{ij}^{(n)} = c_{i+j-2} = \dfrac{\mathrm{D}^{i+j-1}f(0)}{(i+j-1)!} = B_n(0;f)_{ij}$, which is positive by theorem 2.3.3. By corollary 2.3.2, (2.28) is met for every $R > 1$, thus by the previous theorem, there is a unique measure μ_R supported on $[-R, R]$ with

$$c_n = \int_{-R}^{R} x^n \mathrm{d}\mu_R(x)$$

By uniqueness, they agree for all $R > 1$, so there exists a unique finite measure μ supported on $\bigcap_{R>1} [-R, R] = [-1, 1]$ with $c_n = \int_{-1}^{1} x^n \mathrm{d}\mu(x)$. Define $\mathrm{d}\nu(x) = \mathrm{d}\mu(-x)$, and write the Taylor approximation for $z \in \mathrm{B}_1(0)$

$$\begin{aligned} f(z) &= \sum_{k\in\mathbb{N}} \frac{\mathrm{D}^k f(0)}{k!} z^k = f(0) + \sum_{k\in\mathbb{N}} z z^k (-1)^k \int_{-1}^{1} \lambda^k \mathrm{d}\nu(\lambda) \\ &= f(0) + \sum_{k\in\mathbb{N}} \int_{-1}^{1} z(-z)^k \lambda^k \mathrm{d}\nu(\lambda) = f(0) + \int_{-1}^{1} z \sum_{k\in\mathbb{N}} (-\lambda z)^k \mathrm{d}\nu(\lambda) \\ &= f(0) + \int_{-1}^{1} \frac{z}{1+\lambda z} \mathrm{d}\nu(\lambda) \end{aligned}$$

The sum and the integral can be interchanged since the Neumann series $\sum_{k\in\mathbb{N}} x^k$ converges uniformly on $\overline{\mathrm{B}_{|z|}(0)}$, thus for all $\lambda \in [-1, 1]$. ■

Chapter 3

Bounded operators

In this chapter, we shall see that the name "operator monotone" is justified, since an operator monotone function is monotone on the bounded linear operators of any separable Hilbert space. To show this, we have to introduce a continuity notion on $\mathscr{L}(\mathfrak{H})$, that is a topology. We will not use the norm topology, for it has a too fine. In the below introduced topology, any bounded operator of a separable Hilbert space can be approximated by finite rank operators, which is not in the norm topology of an infinite dimensional space, since it would mean that every continuous operator is compact.

Notation 3.0.1 *Let $\mathfrak{H}$ be a Hilbert space. We endow the bounded linear operators of $\mathfrak{H}$, $\mathscr{L}(\mathfrak{H})$ with the topology of pointwise convergence, and call it strong topology to create $(\mathscr{L}(\mathfrak{H}), \tau_s)$. In this notion, for any net $A : \mathfrak{I} \to \mathscr{L}(\mathfrak{H})$ $\lim\limits_{i,\mathfrak{I}} A_i = L \in \mathscr{L}(\mathfrak{H}) \Leftrightarrow \lim\limits_{i,\mathfrak{I}} A_i v = Lv \ \forall v \in \mathfrak{H}$. This will be the convergence used throughout the rest of the thesis, not the norm convergence. If $\mathfrak{H}$ is separable, this topological space is metrizable, because let $\{s_k \mid k \in \mathbb{N}^+\}$*

be a countable dense subset of $\mathfrak{B}$ *and let* $d(A,B) = \sum_{k\in\mathbb{N}^+} \frac{\|(A-B)s_k\|}{1+\|(A-B)s_k\|}$. *In this case, it is enough to check sequence continuity of function as opposed to net continuity, because the space is* M_1.

Lemma 3.0.1 (Continuity) *If* $\mathfrak{H}$ *be a Hilbert space, then* $C^b(\mathbb{R}) \subseteq C((\mathscr{L}(\mathfrak{H})_{sa}, \tau_s), (\mathscr{L}(\mathfrak{H})_{sa}, \tau_s))$

Proof: We start with proving that if $S \subseteq \mathscr{L}(\mathfrak{H})$ is a norm bounded set, then the multiplication (composition) $S \times \mathscr{L}(\mathfrak{H}) \to \mathscr{L}(\mathfrak{H})$ is strongly continuous, where the topology of $S \times \mathscr{L}(\mathfrak{H})$ is $\tau_s \times \tau_s$. For this, we use that

$$\begin{aligned}
\|ABv - CDv\| &\leq \|ABv - ADv\| + \|ADv - CDv\| \\
&\leq \|A\|\|Bv - Dv\| + \|(A-C)Dv\| \\
&\leq K\|Bv - Dv\| + \|(A-C)Dv\|
\end{aligned}$$

where the last equation uses the boundedness of S. If $(A,B) : \mathfrak{I} \to S \times \mathscr{L}(\mathfrak{H})$ is a net converging strongly to (C,D), then A_i and B_i converges to C and D. With this in mind, we can reduce the first and second term of the R.H.S. arbitrarily, thus the L.H.S..

Let $\mathscr{A}$ denote the set of strongly continuous functions, which is a vector space endowed with the pointwise operations. By the above argument, we have that $fh \in \mathscr{A}$ for all $h, f \in \mathscr{A}$, if h is bounded. Firstly we show that $\overline{C_0(\mathbb{R})} \subseteq \mathscr{A}$. Let $\mathscr{A}_0 = \mathscr{A} \cap \overline{C_0(\mathbb{R})}$, which is a closed subalgebra of $\overline{C_0(\mathbb{R})}$. Let z be the indetity on $\mathbb{R}$, $f(x) = (1+x^2)^{-1}$, and $g = fz$. Clearly

$f, g \in \overline{\mathrm{B}_1(0)} \subseteq \overline{C_0(\mathbb{R})}$. If $a, b \in \mathscr{L}(\mathfrak{H})_{sa}$, then

$$\begin{aligned} g(a) - g(b) &= a(1+a^2)^{-1} - b(1+b^2)^{-1} \\ &= (1+a^2)^{-1}[a(1+b^2) - (1+a^2)b](1+b^2)^{-1} \\ &= (1+a^2)^{-1}[a - b + a(b-a)b](1+b^2)^{-1} \end{aligned}$$

so if we take a $v \in \mathfrak{H}$, then

$$\begin{aligned} &\|g(a)v - g(b)v\| \leq \|(1+a^2)^{-1}(a-b)(1+b^2)^{-1}v\| + \|(1+a^2)^{-1}a(b-a)b(1+b^2)^{-1}v\| \\ &\leq \|(1+a^2)^{-1}\|\|(a-b)(1+b^2)^{-1}v\| + \|(1+a^2)^{-1}a\|\|(b-a)b(1+b^2)^{-1}v\| \\ &\leq \|(a-b)(1+b^2)^{-1}v\| + \|(b-a)b(1+b^2)^{-1}v\| \end{aligned}$$

with $\|(1+a^2)^{-1}\|, \|(1+a^2)^{-1}a\| \leq 1$ in consideration. If we substitute a for a net $c : \mathfrak{I} \to \mathscr{L}(\mathfrak{H})_{sa}$ converging to b strongly, the first and second term of the R.H.S. can be arbitrarily small, since $b(1+b^2)^{-1}v$ and $(1+b^2)^{-1}v$ are just fixed vectors in $\mathfrak{H}$. In conclusion, $g \in \mathscr{A}_0$. $z \in \mathscr{A}$ also holds, thus $gz \in \mathscr{A}$, hence by direct computation $f = 1 - gz \in \mathscr{A}$ resulting in $f \in \mathscr{A}_0$. Since $\{f, g\}$ separates the points of $\mathbb{R}$, and vanishes nowhere, so by the Stone-Weierstrass Theorem the closed algebra generated by f and g is $\overline{C_0(\mathbb{R})}$. We already mentioned, that $\mathscr{A}_0 \subseteq \overline{C_0(\mathbb{R})}$, and on the gound of the closed algebra generated by f and g being in $\mathscr{A}_0$, $\mathscr{A}_0 = \overline{C_0(\mathbb{R})}$

Now take any $h \in C_b(\mathbb{R})$. Clearly $hf, hg \in \overline{C_0(\mathbb{R})} = \mathscr{A}_0 \subseteq \mathscr{A}$. If hg is a continuous function vanishing at infinity, it is bounded, so $(hg)z \in \mathscr{A}$, therefore $h = hf + hgz \in \mathscr{A}$ ■

For a set H in a vector space V, let conv$H$ denote the convex hull of H

Lemma 3.0.2 *If $\mathfrak{H}$ is a Hilbert space, $P, A \in \mathscr{L}(\mathfrak{H})_{sa}$ with P being an orthogonal projection and $A_P = PAP|_{\mathbf{Ran}P}$, then $\operatorname{spec} A_P \subseteq \operatorname{conv}(\operatorname{spec} A)$. In particular, if $\operatorname{spec} A \subseteq (a, b)$, then $\operatorname{spec} A_P \subseteq (a, b)$*

Proof: Since the spectrum of a bounded operator is compact and a is self-adjoint so it has a real spectrum, with a greatest element g and a least element l.
This implies $\operatorname{spec} A \subseteq [l, g] \Rightarrow l\mathbf{1} \leq A \leq g\mathbf{1} \Rightarrow l\mathbf{1}_{\mathbf{Ran}A} \leq A_P \leq g\mathbf{1}_{\mathbf{Ran}P} \Rightarrow \operatorname{spec} A_P \subseteq [l, g] \subseteq \operatorname{conv}(\operatorname{spec} A)$ ∎

Theorem 3.0.1 *Let $\mathfrak{H}$ be an infinite dimensional separable Hilbert space, $A, B \in \mathscr{L}(\mathfrak{H})_{sa}$ with spectra in some (a, b), and $f \in \mathfrak{M}_\infty(a, b)$. Then*

$$A \leq B \Rightarrow f(A) \leq f(B) \tag{3.1}$$

Proof: We can assume, that $0 \in (a, b)$, because a we can translate the interval with $T(x) = \frac{2}{b-a}x - \frac{a+b}{b-a}$ and use $T(A), T(B)$, $f \circ T^{-1}$, $(-1, 1)$. If $\mathfrak{H}$ is separable, then it has an orthonormal basis $\{v_k \mid k \in \mathbb{N}^+ \}$. If P_n be the orthogonal projection onto the span of the first n basis vectors, then $P_n \xrightarrow{s} \mathbf{1}$, thus $P_nAP_n \to A$, and $P_nBP_n \to B$, because $P_nAP_n - A = P_nAP_n - P_nA + P_nA - A$.
Since $P_nAP_n = A_{P_n} \oplus 0_{1-P_n}$, $\operatorname{spec} P_nAP_n = \operatorname{spec} A_{P_n} \cup \operatorname{spec} 0_{1-P_n}$
$\subseteq \operatorname{conv}(\operatorname{spec} A \cup 0)$ by the previous lemma and the fact, that $\operatorname{spec} 0 = \{0\}$. This set, and $\operatorname{conv}(\operatorname{spec} B \cup 0)$ are compact subsets of (a, b), so by Urysohn's lemma, there is a continuous $g : (a, b) \to [0, 1]$ function which has $g(x) = 1$ for all $x \in (\operatorname{conv}(\operatorname{spec} A \cup 0) \cup \operatorname{conv}(\operatorname{spec} B \cup 0))$, and $g(y) = 0$ outside of a compact subinterval of (a, b). Since an operator monotone function is continuous, fg

is a bounded continuous function that agrees with f on $(\mathrm{conv}(\operatorname{spec} A \cup 0) \cup \mathrm{conv}(\operatorname{spec} B \cup 0))$. Hence $f(P_nAP_n)$ is well defined, and so is $f(P_nBP_n)$. Since P_nAP_n and P_nBP_n are $n \times n$ matrices with $P_nAP_n \leq P_nBP_n$, we have that $f(P_nAP_n) \leq f(P_nBP_n)$. f is strongly continuous, hence it is strongly sequence continuous, so $f(P_nAP_n) \to f(A)$ and $f(P_nBP_n) \to f(B)$, thus $f(A) \leq f(B)$

■

Chapter 4

Convexity

In this chapter, we prove an analogue of a theorem from real analysis, which is *if* $f : (0, \infty) \to [0, \infty)$ *is a concave function, then* f *is monotone increasing*. With the help of this remark, we can show that not only every operator monotone function on $\mathbb{R}$ is affine, but this is the case for even 2-monotone functions on $\mathbb{R}$.

Definition 4.0.1 *An* $f : (a, b) \to \mathbb{R}$ *function is called concave, if* $f(tx + (1-t)y) \geq tf(x) + (1-t)f(y)$ *holds for all* $x, y \in (a, b)$ *and* $t \in [0, 1]$. f *is referred to as convex, if* $-f$ *is concave.*

Definition 4.0.2 *An* $f : (a, b) \to \mathbb{R}$ *function is referred to as n-concave, if* $f(tA+(1-t)B) \geq tf(A)+(1-t)f(B)$ *holds for all* $t \in [0, 1]$ *and* $A, B \in M_n[\mathbb{C}]$ *that obeys* $\operatorname{spec} A, \operatorname{spec} B \subseteq \mathbf{Dom} f$. *We denote the set of such functions* $\mathfrak{C}_n(a, b)$. f *is called n-convex, if* $-f \in \mathfrak{C}_n(a, b)$.

If $f \in \bigcap_{n \in \mathbb{N}^+} \mathfrak{C}_n(a, b)$, *then* f *is called operator concave. Let* $\mathfrak{C}_\infty(a, b)$ *denote the set of such functions.* f *is called operator concave, if* $-f \in \mathfrak{C}_\infty(a, b)$.

Recall, that if A is a bounded operator on a Hilbert space $\mathfrak{H}$, and P is an ortogonal projection, then $A_P = PAP|_{\mathbf{Ran}P}$.

Lemma 4.0.1 *Let $A \in \mathscr{L}(\mathfrak{H})^+$, where $\mathfrak{H}$ is an arbitrary Hilbert space, and let $P \in \mathscr{L}(\mathfrak{H})^+$ be an ortogonal projection. If $f : (\mathscr{L}(\mathfrak{H})^+, \tau_s) \to (\mathscr{L}(\mathfrak{H}), \tau_s)$ is a continuous and monotone function, then*

$$f(A)_P \leq f(A_P) \tag{4.1}$$

Proof: By writing $\mathfrak{H} = P(\mathfrak{H}) \oplus (\mathbf{1} - P)(\mathfrak{H})$, we can decompose

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Where $a_{11} = A_P$. Given an $\varepsilon > 0$, we define $T_\varepsilon = \begin{pmatrix} \sqrt{\varepsilon} & 0 \\ 0 & -\sqrt{\varepsilon}^{-1} \end{pmatrix}$, thus

$$T_\varepsilon^* A T_\varepsilon = \begin{pmatrix} \varepsilon a_{11} & -a_{12} \\ -a_{21} & \varepsilon^{-1} a_{22} \end{pmatrix}$$

Notice, that $A \geq 0 \Rightarrow T_\varepsilon^* A T_\varepsilon \geq 0$, since we just made a basis transformation. Hence

$$A \leq A + T_\varepsilon^* A T_\varepsilon = \begin{pmatrix} (1+\varepsilon) a_{11} & 0 \\ 0 & (1+\varepsilon^{-1}) a_{22} \end{pmatrix}$$

For f is monotone,

$$f(A) \leq \begin{pmatrix} f((1+\varepsilon) a_{11}) & 0 \\ 0 & f((1+\varepsilon^{-1}) a_{22}) \end{pmatrix}$$

It is clear that $X \geq Y \Rightarrow X_P \geq Y_P$ holds, thus
$f(A)_P \leq f((1+\varepsilon) a_{22}) = f((1+\varepsilon) A_P)$ for all $\varepsilon > 0$. Since $(1+\varepsilon) A_P$ converges

to A_P strongly as $\varepsilon \searrow 0$ and f is strongly continuous,
$f(A)_P \le \lim_{\varepsilon \searrow 0} f((1+\varepsilon)A_P) = f(A_P)$. ■

Theorem 4.0.1 *If* $f : (0,\infty) \to \mathbb{R}$ *and* $n \in \mathbb{N}^+$, *then*

(***a***) $f \in \mathfrak{C}_n(0,\infty)$ *and* $f \ge 0 \ \Rightarrow f \in \mathfrak{M}_n(0,\infty)$

(***b***) $f \in \mathfrak{M}_{2n}(0,\infty) \ \Rightarrow f \in \mathfrak{C}_n(0,\infty)$

In particular, $f \ge 0 \quad \Rightarrow \quad f \in \mathfrak{C}_\infty(0,\infty) \Leftrightarrow f \in \mathfrak{M}_\infty(0,\infty)$.

Proof: (***a***): If $A, B \in M_n[\mathbb{C}]$ are self-adjoint matrices with $0 \le A \le B$, then

$$B = A + \varepsilon\varepsilon^{-1}(B-A) = (1+\varepsilon)\left[\frac{1}{1+\varepsilon}A + \frac{\varepsilon}{1+\varepsilon}\varepsilon^{-1}(B-A)\right]$$

Dividing both sides by $(1+\varepsilon)$ and using the n-concavity of f, we see that

$$f\left(\frac{B}{1+\varepsilon}\right) \ge \frac{1}{1+\varepsilon}f(A) + \frac{\varepsilon}{1+\varepsilon}f\left(\varepsilon^{-1}(B-A)\right)$$

$f \ge 0$ implies $\frac{\varepsilon}{1+\varepsilon}f\left(\varepsilon^{-1}(B-A)\right) \ge 0$, thus $f\left(\frac{B}{1+\varepsilon}\right) \ge \frac{1}{1+\varepsilon}f(A)$. One can change to a basis in which B is diagonal to see that $f\left(\frac{B}{1+\varepsilon}\right) \to f(B)$, and clearly $\frac{1}{1+\varepsilon}f(A) \to f(A)$, hence by taking $\varepsilon \searrow 0$, we get

$$f(B) \ge f(A)$$

(***b***): Write $\mathbb{C}^{2n} = \mathbb{C}^n \oplus \mathbb{C}^n$, and let P be the orthogonal projection onto $\mathbb{C}^n \oplus \{0\}$. Given $A, B \in M_n[\mathbb{C}]^+$ and $\Xi \in [0,1]$, we define

$$T_\Xi = \begin{pmatrix} \sqrt{\Xi}\mathbf{1} & -\sqrt{1-\Xi}\mathbf{1} \\ \sqrt{1-\Xi}\mathbf{1} & \sqrt{\Xi}\mathbf{1} \end{pmatrix}$$

which is clearly unitary. By direct computation, one can see that $\left[T_\Xi^* \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} T_\Xi\right]_{11} = \Xi A + (1-\Xi)B$. Thus by the preceding proposition

$$\Xi f(A) + (1-\Xi)f(B) = \left[\left(T_\Xi^* \begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix} T_\Xi\right)\right]_{11}$$

$$= \left[f\left(T_\Xi^* \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} T_\Xi\right)\right]_{11} \leq f\left(\left[T_\Xi^* \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} T_\Xi\right]_{11}\right)$$

$$= f(\Xi A + (1-\Xi)B)$$

which yields $f \in \mathfrak{C}_n(a,b)$. ■

Theorem 4.0.2 *If $f \in \mathfrak{M}_2(\mathbb{R})$, then f is an affine function, that is $f(x) = ax + b$ for some $a \geq 0$ and $b \in \mathbb{R}$.*

Proof: If $f \in \mathfrak{M}_2(\mathbb{R})$, then take any $a < 0$ and write $g_a(x) = f(x+a)$. We see that $g_a \in \mathfrak{M}_2(0,\infty)$, thus it is concave on $(0,\infty)$. Hence, f is concave on (a,∞) for all $a < 0$, therefore on all $\mathbb{R}$. $-f(-x) \in \mathfrak{M}_2(\mathbb{R})$ also holds, which implies that $-f(-x)$ is concave as well, which is by definition means that $f(-x)$ is convex. Since $f(-x)$ is convex on a symmetric domain if and only $f(x)$ is, f is also convex. Every function that is both convex and concave is affine, therefore so is f. ■

Chapter 5

Applications

5.1 Operator means

In this chapter, we will discuss operations on $\mathscr{L}(\mathfrak{H})$, the space of bounded linear operators of a separable Hilbert space $\mathfrak{H}$, which is allowed to be infinite dimensional. The subgroup of invertible elements in $\mathscr{L}(\mathfrak{H})$ will be denoted as $\mathscr{G}\mathscr{L}(\mathfrak{H})$. The set positive elements in these spaces will use the $\mathscr{L}(\mathfrak{H})^+$ and $\mathscr{G}\mathscr{L}(\mathfrak{H})^+$ notations. We need symbols as we will discuss multiple functions on these spaces, and the domains will be easier to notate. The convergence of this chapter is the strong convergence.

In this context, the Schur decomposition (2.2.3) can be rewritten, but proven the same way.

Theorem 5.1.1 (Schur complement) *If* $A, X \in \mathscr{L}(\mathfrak{H})$ *and* $B \in \mathscr{G}\mathscr{L}(\mathfrak{H})$

with $A, B \geq 0$, *then*

$$\begin{pmatrix} A & X \\ X^* & B \end{pmatrix} \geq 0 \quad \Leftrightarrow \quad A \geq XB^{-1}X^* \tag{5.1}$$

Corollary 5.1.1 *The funtion* $F : \mathscr{L}(\mathfrak{H}) \times \mathscr{G}\mathscr{L}(\mathfrak{H})^+ \to \mathscr{L}(\mathfrak{H})$, $(X, B) \mapsto XB^{-1}X^*$ *is jointly convex, and* $F_1 : \mathscr{L}(\mathfrak{H}) \times \mathscr{L}(\mathfrak{H}) \times \mathscr{G}\mathscr{L}(\mathfrak{H})^+ \to \mathscr{L}(\mathfrak{H})$, $(A, X, B) \mapsto A - XB^{-1}X^*$ *is jointly concave.*

Proof: The second statement follows from the first. Take any (X, B) and (Y, A) from $\textbf{Dom}\, F$, and a $t \in [0, 1]$. From the predecessing theorem, $\begin{pmatrix} XB^{-1}X^* & X \\ X^* & B \end{pmatrix} \geq 0$, so by the convexity of $\mathscr{L}(\mathfrak{H})^+$,

$$\begin{pmatrix} tXB^{-1}X^* + (1-t)YA^{-1}Y^* & tX + (1-t)Y \\ tX^* + (1-t)Y^* & tB + (1-t)A \end{pmatrix} \geq 0$$

If we use the theorem on this as well, we get
$F(t(X, B) + (1-t)(Y, A)) \leq tF(X, B) + (1-t)F(Y, A)$. ■

Definition 5.1.1 *We define the parallel sum* $\mathscr{G}\mathscr{L}(\mathfrak{H})^+ \times \mathscr{G}\mathscr{L}(\mathfrak{H})^+ \to \mathscr{G}\mathscr{L}(\mathfrak{H})^+$ $(A, B) \mapsto A : B = (A^{-1} + B^{-1})^{-1}$, *and the harmonic mean* $A!B = 2(A : B)$

Proposition 5.1.1 *The parallel sum obeys*

(a)

$$A : B = A - A(A + B)^{-1}A = B - B(A + B)^{-1}B \tag{5.2}$$

(b) $(A, B) \mapsto A : B$ *is jointly monotone and jointly concave*

(c) $A : B \leq A$, $A : B \leq B$

(d)

$$A : B = \max\left\{ Y \in \mathscr{L}(\mathfrak{H})^+ \middle| \begin{pmatrix} A & A \\ A & A+B \end{pmatrix} \geq \begin{pmatrix} Y & 0 \\ 0 & 0 \end{pmatrix} \right\} \tag{5.3}$$

where max *means greatest element.*

(e) $A \leq B \Rightarrow A \leq A!B \leq B$

Proof: (a):

$$(A^{-1} + B^{-1})^{-1} = (A^{-1}(A+B)B^{-1})^{-1} = B(A+B)^{-1}B$$
$$= B(A+B)^{-1}[(A+B) - A] = B - B(A+B)^{-1}B$$

The other equation follows from the obvious symmetry of the parallel sum in the original form.

(b): We have proved in (2.2.4) that if $f : A \mapsto -A^{-1}$, then $f \in \mathfrak{M}_\infty(0, \infty)$, so $A : B = f(f(A) + f(B)) = -(-A^{-1} - B^{-1})^{-1}$ is indeed jointly monotone, since $+$ is. By Corollary 5.1.1, $(A, B) \mapsto F_1(A, A, A+B) = A - A(A+B)^{-1}A = A : B$ is jointly concave.

(c): $A, B \geq 0 \Rightarrow (A+B)^{-1} \geq 0 \Rightarrow A(A+B)^{-1}A, B(A+B)^{-1}B \geq 0$ by definition, hence $A - A(A+B)^{-1}A \leq A,\ B - B(A+B)^{-1}B \leq B.$

(d) By the Schur decomposition,the inequality holds exactly when $(A - Y) \geq A(A+B)^{-1}A$, which is $A : B = A - A(A+B)^{-1}A \geq Y$.

(e) $A \leq B \Rightarrow 2A \leq A + B \leq 2B \Rightarrow B^{-1} \leq 2(A+B)^{-1} \leq A^{-1} \Rightarrow B \leq 2B(A+B)^{-1}B$ and $2A(A+B)^{-1}A \leq A \Rightarrow 2B - B \geq A : B$ and $A : B \geq 2A - A$ ■

Theorem 5.1.2 *If $A_n, B_n \in \mathscr{GL}(\mathfrak{H})^+$, $A, B \in \mathscr{L}(\mathfrak{H})^+$ such that $A_n \searrow A$ and $B_n \searrow B$, then $\lim_{n\to\infty} A_n : B_n$ exists in the strong topology, and $A_n : B_n \searrow A : B$ independent of the choice of A_n and B_n.*

Proof: The monotonicity of the sequence follows from the joint monotonicity of the parallel sum.

First, we assume that A and B are invertible. In this case $(A_n + B_n)^{-1} \leq (A + B)^{-1} \Rightarrow A_n(A_n + B_n)^{-1}A_n \leq A_n(A + B)^{-1}A_n$ Since $\{A_n \mid n \in \mathbb{N}\}$ is pointwise convergent, it is pointwise bounded, thus by the Uniform boundedness principle, it is norm bounded by a $K > 0$. Also $A_n + B_n \geq A + B \Rightarrow 0 \leq (A_n + B_n)^{-1} \leq (A + B)^{-1}$, hence $\|(A_n + B_n)^{-1}\| \leq \|(A + B)^{-1}\|$. Utilising the triangle inequality with, wee see that for all $v \in \mathfrak{H}$

$$\begin{aligned}
&\|A_n(A_n + B_n)^{-1}A_nv - A(A + B)^{-1}Av\| \\
&\leq \|A_n(A_n + B_n)^{-1}A_nv - A_n(A_n + B_n)^{-1}Av\| \\
&+ \|A_n(A_n + B_n)^{-1}Av - A_n(A + B)^{-1}Av\| \\
&+ \|A_n(A + B)^{-1}Av - A(A + B)^{-1}Av\| \\
&\leq K\|(A + B)^{-1}\|\|A_nv - Av\| \\
&+ K\|(A_n + B_n)^{-1}(Av) - (A + B)^{-1}(Av)\| \\
&+ \|A_n((A + B)^{-1}Av) - A((A + B)^{-1}Av)\|
\end{aligned}$$

The first and third term can be arbitrarily small since $A_n \to A$.For the second one we consider $\operatorname{spec}(A+B) \subseteq (0, \infty)$, which a compact set, so it has a least element c. $A_n + B_n \geq A + B \geq c$ thus $\bigcup_{n\in\mathbb{N}} \operatorname{spec}(A_n + B_n) \cup \operatorname{spec}(A + B) \subseteq [c, \infty)$ thus we can truncate the inversion function to create a bounded continuous function $f(x) = x^{-1}\mathbf{1}_{x\geq c} + c^{-1}\mathbf{1}_{x<c}$ which is strongly continuous

by Lemma 3.0.1, thus $(A_n+B_n)^{-1} = f(A_n+B_n) \to f(A+B) = (A+B)^{-1}$, hence the second term vanishes as well. This implies $A_n(A_n+B_n)^{-1}A_n \to A(A+B)^{-1}A$, thus $(A_n : B_n) \to (A : B)$. In this case, the independecy is obvious.

For general $A, B \in \mathscr{L}(\mathfrak{H})^+$, by the joint monotonicity of the parallel sum, $0 \leq A_n : B_n$ is a monotone decreasing sequence, thus it has a strong limit.

For any other $A'_m, B'_m \in \mathscr{GL}(\mathfrak{H})^+$ pair of sequences with $A'_m \searrow A$ and $B'_m \searrow B$, $A_n + A'_m - A \searrow A_n$ and $B_n + B'_m - B \searrow B_n$ as $m \to \infty$ for all $n \in \mathbb{N}$. Thus by the first case, $(A_n + A'_m - A) : (B_n + B'_m - B) \searrow A_n : B_n$ for all $n \in \mathbb{N}$. Since $A_n : B_n \leq (A_n + A'_m - A) : (B_n + B'_m - B)$, taking $n \to \infty$, we see that $\lim\limits_{n\to\mathbb{N}} A_n : B_n \leq A'_m : B'_m$ for all $m \in \mathbb{N}$. Taking limit in m, we get $\lim\limits_{n\to\infty} A_n : B_n \leq \lim\limits_{m\to\infty} A'_m : B'_m$. As the whole argument was symmetric in A_n, B_n and A'_m, B'_m, we conclude that $\lim\limits_{n\to\infty} A_n : B_n \leq \lim\limits_{m\to\infty} A'_m : B'_m$ also holds, which proves independency.

∎

This theorem not only shows us, that the parallel sum is upper semi-continuous, but also

Corollary 5.1.2 *The parallel sum can be extended onto $\mathscr{L}(\mathfrak{H})^+ \times \mathscr{L}(\mathfrak{H})^+$ by $A : B = \lim\limits_{\varepsilon\searrow 0}(A + \varepsilon\mathbf{1}) : (B + \varepsilon\mathbf{1})$*

Proof: For any $A \geq 0$, $A + \varepsilon\mathbf{1} \geq \varepsilon\mathbf{1}$, thus $(A + \varepsilon\mathbf{1})^{-1} \leq \varepsilon^{-1}\mathbf{1}$, hence $(A + \varepsilon\mathbf{1})^{-1} \in \mathscr{GL}(\mathfrak{H})^+$, and clearly $A + \varepsilon\mathbf{1} \searrow A$. The same can be noted about B and B_n.

If A and B are invertible,then the above expresion is the original $A : B$, and still exists otherwise by the preceding theorem. ∎

In the following, when use the parallel sum, we refer to this extension. We will encounter some variational formulas that show a minimality or maximality property of an expression. These formulae are quite useful, as we will see, particularly for proof of convexity, or concavity.

Lemma 5.1.1 $\forall v \in \mathfrak{H} \quad \langle v \,| A : Bv \rangle = \min\{\langle x \,| Ax \rangle + \langle y \,| By \rangle \;\mid v = x + y\}$

Proof: The scalar product is continuous and the parallel sum is upper semi-continuous, so by taking limits we can assume that A and B are invertible. Since $A : B = B - B(A+B)^{-1}B$,

$$
\begin{aligned}
&\langle x \,| Ax \rangle + \langle v - x \,| B(v-x) \rangle - \langle v \,| A : Bv \rangle \\
&= \langle x \,| (A+B)x \rangle + \langle v \,| Bv \rangle - 2\,\mathfrak{Re}\, \langle x \,| Bv \rangle - \langle v \,| A : Bv \rangle \\
&= \langle x \,| (A+B)x \rangle - 2\,\mathfrak{Re}\, \langle x \,| Bv \rangle + \left\langle x \,\middle|\, B(A+B)^{-1}Bx \right\rangle \\
&= \|(A+B)^{1/2}x\|^2 + \|(A+B)^{-1/2}Bv\|^2 - 2\,\mathfrak{Re}\left\langle (A+B)^{1/2}x \,\middle|\, (A+B)^{-1/2}Bv \right\rangle \\
&= \|(A+B)^{1/2}x - (A+B)^{-1/2}Bv\|^2 \geq 0
\end{aligned}
$$

and the equality can be reached with $x = (A+B)^{-1}Bv$. ■

Proposition 5.1.2 *$\forall A, B, C, D \in \mathscr{L}(\mathfrak{H})^+$:*

(a) $\forall T \in \mathscr{L}(\mathfrak{H})$ $T^*(A : B)T \leq (T^*AT) : (T^*AT)$

(b) $A : B + C : D \leq (A + C) : (B + D)$

Proof: (a): By the lemma above, $\forall v \in \mathfrak{H}$ $\exists x, y \in \mathfrak{H}$ such that $v = x + y$ and $\langle x \,| T^*ATx \rangle + \langle y \,| T^*BTy \rangle = \langle v \,| (T^*AT) : (T^*BT)v \rangle$, thus

$$
\begin{aligned}
&\langle v \,| T^*(A : B)Tv \rangle = \langle Tv \,| (A : B)Tv \rangle \\
&\leq \langle Tx \,| ATx \rangle + \langle Ty \,| BTy \rangle = \langle x \,| T^*ATx \rangle + \langle y \,| T^*BTy \rangle
\end{aligned}
$$

Where the inequality uses the lemma as well.

(*b*) Again $\forall v \in \mathfrak{H}\ \exists x, y \in \mathfrak{H}$ such that $v = x + y$ and $\langle x\,|(A+C)x\rangle + \langle y\,|(B+D)y\rangle = \langle v\,|(A+C):(B+D)v\rangle$, thus

$$\langle v\,|(A:B+C:D)v\rangle = \langle v\,|(A:B)v\rangle + \langle v\,|(C:D)v\rangle$$
$$\leq \langle x\,|Ax\rangle + \langle y\,|By\rangle + \langle x\,|Cx\rangle + \langle y\,|Dy\rangle$$
$$= \langle x\,|(A+C)x\rangle + \langle y\,|(B+D)y\rangle$$

■

Definition 5.1.2 *We define the geometric mean,*
$\# : \mathscr{G}\mathscr{L}(\mathfrak{H})^+ \times \mathscr{L}(\mathfrak{H})^+ \to \mathscr{L}(\mathfrak{H})^+$,
$(A, B) \mapsto A\#B = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2}$.

Proposition 5.1.3

$$A\#B = \max\{X \in \mathscr{L}(\mathfrak{H})^+ \mid XA^{-1}X \leq B\} \tag{5.4}$$

where the max *notates the greatest element of the set, not the maximal. Moreover,* $(A\#B)A^{-1}(A\#B) = B$ *holds, which implies that* $X = A\#B$ *is the unique positive solution of* $XA^{-1}X = B$.

Proof: $A\#BA^{-1}A\#B = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{\frac{1}{2}\cdot 2}A^{1/2} = B$

If $XA^{-1}X \leq B$, then $A^{-1/2}XA^{-1}XA^{-1/2} \leq A^{-1/2}BA^{-1/2}$ which is $(A^{-1/2}XA^{-1/2})^2 \leq A^{-1/2}BA^{-1/2}$, thus by the monotonicity of the square root, $A^{-1/2}XA^{-1/2} \leq \left(A^{-1/2}BA^{-1/2}\right)^{1/2}$ which yields $X \leq A\#B$. If $XA^{-1}X = B$, then the inequalities turn into equalities, so $X = A\#B$, hence the positive solution of $X = A\#B$ is unique. ■

Corollary 5.1.3 $\forall A, B \in \mathscr{GL}(\mathfrak{H})^+ \quad A\#B = B\#A$

Proof: $XB^{-1}X = A \Leftrightarrow B^{-1} = X^{-1}AX^{-1} \Leftrightarrow B = XA^{-1}X$, since if both A and B is invertible, then so is X. By the proposition, $A\#B \leq B\#A$ and $B\#A \leq A\#B$. ■

Theorem 5.1.3 $A\#B = \max\left\{ X \in \mathscr{L}(\mathfrak{H})^+ \;\middle|\; \begin{pmatrix} A & X \\ X^* & B \end{pmatrix} \geq 0 \right\}$
Where the max *means greatest element.*

Proof: This follows from Proposition 5.4 and Theorem 5.1.1 of the Schur complement. ■

Corollary 5.1.4 *The geometric mean # is jointly concave.*

Proof:

Let $(A, B), (C, D) \in \mathbf{Dom}\,\#$ and $t \in [0, 1]$. If $X = A\#B$ and $Y = C\#D$, we have that

$$\begin{pmatrix} A & X \\ X & B \end{pmatrix} \geq 0, \ \begin{pmatrix} C & Y \\ Y & D \end{pmatrix} \geq 0 \Rightarrow \begin{pmatrix} tA + (1-t)B & tX + (1-t)Y \\ tX + (1-t)Y & tB + (1-t)D \end{pmatrix} \geq 0$$

Thus $tA\#B + (1-t)C\#D = tX + (1-t)Y \leq (tA + (1-t)B)\#(tB + (1-t)D)$

■

We ca ngive a similar proof for the concavity of the parallel sum using (5.3).

Proposition 5.1.4 *The parallel sum is jointly concave.*

Proof: Let $(A, B), (C, D) \in \mathbf{Dom}$: and $t \in [0, 1]$. If we write $X = A : B$ and $Y = C : D$, we see that

$$\begin{pmatrix} A & A \\ A & A+B \end{pmatrix} \geq \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} C & C \\ C & C+D \end{pmatrix} \geq \begin{pmatrix} Y & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} tA + (1-t)C & tA + (1-t)C \\ tA + (1-t)C & t(A+B) + (1-t)(C+D) \end{pmatrix} \geq \begin{pmatrix} tX + (1-t) & 0 \\ 0 & 0 \end{pmatrix}$$

Thus $t(A : B) + (1-t)(C : D) = tX + (1-t)Y \leq (tA + (1-t)C) : (tB + (1-t)D)$

■

Kubo and Ando generalised the notion of means in [8] using a few of the above presented properties of some natural examples.

Definition 5.1.3 *An operator mean is a map*
$\sigma : \mathscr{L}(\mathfrak{H})^+ \times \mathscr{L}(\mathfrak{H})^+ \to \mathscr{L}(\mathfrak{H})^+$, $(A, B) = A\sigma B$ *which obeys four axioms:* $\forall A, B, C, D \in \mathscr{L}(\mathfrak{H})^+$

(i) Joint monotonicity: $A \leq C \ \wedge \ B \leq D \ \Rightarrow \ A\sigma B \leq C\sigma D$

(ii) Transformer inequality: $C(A\sigma B)C \leq (CAC)\sigma(CBC)$

(iii) Upper semi-continuity: $A_n \searrow A \ \wedge \ B_n \searrow B \ \Rightarrow \ A_n\sigma B_n \searrow A\sigma B$

(iiii) Normalisation: $\mathbf{1}\sigma\mathbf{1} = \mathbf{1}$

During the following discussion, we will often suppose that A and B are invertible, since we can approximate them with $A + n^{-1}\mathbf{1}$ and $B + n^{-1}\mathbf{1}$ which are, and use the upper semi-continuity. Moreover, if C is invertible, the transformer inequality tells us $C^{-1}[(CAC)\sigma(CBC)]C^{-1} \leq A\sigma B$ resulting in

$$C \in \mathscr{GL}(\mathfrak{H})^+ \;\Rightarrow\; C(A\sigma B)C = (CAC)\sigma(CBC) \tag{5.5}$$

Notice that this is a quite broad definition, since even $(A, B) \mapsto A$ is an operator mean.

We can also see that

$$A \leq B \Rightarrow A \leq A\sigma B \leq B \tag{5.6}$$

since we can assume that A and B are invertible, $A = A^{1/2}\mathbf{1}A^{1/2} = A^{1/2}(\mathbf{1}\sigma\mathbf{1})A^{1/2} = (A^{1/2}\mathbf{1}A^{1/2})\sigma(A^{1/2}\mathbf{1}A^{1/2}) = A\sigma A \leq A\sigma B \leq B\sigma B = B$. During this proof, we also concluded that

$$\forall A \in \mathscr{L}(\mathfrak{H})^+ \quad A = A\sigma A$$

In retrospection, we have proven that the harmonic mean is an operator mean, since the parallel sum obeys axioms $i - iii$.

5.2 Kubo-Ando Theorem

For any $f \in \mathfrak{M}_\infty(0, \infty)$ obeying $f(1) = 1$ and $f \geq 0$, by Loewner's theorem and a $\lambda = -x$ substitution, we have a Herglotz representation

$$f(z) = a + bz + \int_0^\infty \frac{z(1+\lambda)}{z+\lambda} \mathrm{d}\nu(\lambda) \tag{5.7}$$

$$f(1) = a + b + \int_0^\infty \mathrm{d}\nu(\lambda) \tag{5.8}$$

with $a, b \geq 0$

Theorem 5.2.1 (Kubo-Ando Theorem) *If $(A, B) \mapsto A\sigma B$ is an operator mean, then $\mathbf{1}\sigma(t\mathbf{1})$ is the scalar multiple of $\mathbf{1}$ for all $t \in [0, \infty)$, henceforth there exists an $f : [0, \infty) \to [0, \infty)$ operator monotone function, such that*

$$f(t)\mathbf{1} = \mathbf{1}\sigma(t\mathbf{1}) \tag{5.9}$$

and $f(1) = 1$ holds. If f is represented by (5.7), then for all $A, B \in \mathscr{L}(\mathfrak{H})^+$

$$A\sigma B = aA + bB + \int_0^\infty \frac{1+\lambda}{\lambda}((\lambda A) : B)\mathrm{d}\nu(\lambda) \tag{5.10}$$

Conversely, if $f : [0, \infty) \to [0, \infty)$ is an operator monotone function with $f(1) = 1$ represented by (5.7), then (5.10) defines an operator mean that obeys (5.9). Furthermore:

(a) If $A \in \mathscr{G}\mathscr{L}(\mathfrak{H})^+$, then

$$A\sigma B = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2}$$

(b) $A\sigma A = A$ for all $A \in \mathscr{L}(\mathfrak{H})^+$

From the integral representation of σ, one can see that every operator mean is jointly concave, since for all $\lambda > 0$, $(A, B) \mapsto (\lambda A) : B$ is for the parallel sum is by Proposition (5.1.4). Utilising this, we can see that the theorem gives another proof for $f \geq 0$ *and* $f \in \mathfrak{M}_\infty(0, \infty) \;\Rightarrow\; f \in \mathfrak{C}_\infty(0, \infty)$ by (5.9)

If we introduce an order on the set of non-negative operator monotone functions on $[0, \infty)$ with $f(1) = 1$ by $f \leq g \;\Leftrightarrow\; f(t) \leq g(t)\ \forall t > 0$, and on the set of operator means by $\sigma_1 \leq \sigma_2 \;\Leftrightarrow\; A\sigma_1 B \leq A\sigma_2 B\ \forall A, B \in \mathscr{L}(\mathfrak{H})^+$, from (5.9), it is clear that the mapping $f \mapsto \sigma_f$ is monotone, an order isomorphism even.

From (5.9) we can find the representing functions of the geometric (#), harmonic (!) and arithmetic ($\triangledown$) means. $\mathbf{1}\#(t\mathbf{1}) = \sqrt{t}\mathbf{1}$, thus the function g of the geometric mean is the square root. $\mathbf{1}!(t\mathbf{1}) = 2(\mathbf{1} + t^{-1}\mathbf{1})^{-1}$, hence $h(t) = \dfrac{2t}{1+t}$. $\mathbf{1}\triangledown(t\mathbf{1}) = \dfrac{\mathbf{1}+t\mathbf{1}}{2}$, so $a(t) = \dfrac{1+t}{2}$.

Corollary 5.2.1 $\forall A, B \in \mathscr{L}(\mathfrak{H})^+ \quad A!B \leq A\#B \leq A\triangledown B$

Proof: Since $f \mapsto \sigma_f$ is an order isomorphism, we only have to prove that $h(t) \leq g(t) \leq a(t)$ for all $t > 0$, that is

$$\frac{2t}{1+t} \leq \sqrt{t} \leq \frac{1+t}{2}$$

The right inequality is just the regular inequality of geometric and arithmetic means. If we divide the left inequality by $t > 0$, we get $\dfrac{2}{1+t} \leq \sqrt{t}^{-1}$, hence $\dfrac{1+t}{2} \geq \sqrt{t}$, which the same inequality once again. ∎

Lemma 5.2.1 *Let σ be an operator mean and $A, B \in \mathscr{L}(\mathfrak{H})$. For any orthogonal projection P commuting with both A and B, we have that*

$$[(AP)\sigma(BP)]P = (A\sigma B)P \tag{5.11}$$

and P commutes with $A\sigma B$.

Proof: By the transformer inequality and monotonicity, we note that

$$P(A\sigma B)P \leq (PAP)\sigma(PBP) \leq A\sigma B \tag{5.12}$$

which means that $C \equiv A\sigma B - P(A\sigma B)P \geq 0$. Since $P^2 = P$, $0 = PCP = (C^{1/2}P)^*C^{1/2}P$ so $CP = C^{1/2}(C^{1/2}P) = C^{1/2}0 = 0$, thus

$(A\sigma B)P = P(A\sigma B)P$ so P commutes with $A\sigma B$ for the right side is self-adjoint. Since P commutes with AP and BP, it commutes with $(AP)\sigma(BP)$ as well. If we turn back to (5.12) now, we see that

$$\begin{aligned} P(A\sigma B)P &\le (PAP)\sigma(PBP) \le A\sigma B \\ P(A\sigma B)P &\le \quad (AP)\sigma(BP) \quad \le A\sigma B \\ P(A\sigma B)P &\le [(AP)\sigma(BP)]P \le (A\sigma B)P \\ (A\sigma B)P &\le \quad (AP)\sigma(BP) \quad \le (A\sigma B)P \end{aligned}$$

which is just (5.11). ■

In the following proof, we will use some properties of the Dirac notation of the scalar product. So far, the only difference between this notation and the the standard one has been that the former uses a "$\langle \cdot \, | \cdot \rangle$" instead of "$\langle \cdot, \cdot \rangle$". However, in the Dirac notation, $|\cdot\rangle$ is a linear function from $\mathfrak{H}$ onto $\mathscr{L}(\mathbb{C}, \mathfrak{H})$ defined by $|v\rangle : \mathbb{C} \to \mathfrak{H}$, $\lambda \mapsto \lambda v$ for all $v \in \mathfrak{H}$. $\langle \cdot|$ is a conjugate linear function from $\mathfrak{H}$ to the space of linear functionals of $\mathfrak{H}$ defined by $\langle v| : \mathfrak{H} \to \mathbb{C}$, $u \mapsto \langle v \, | u \rangle$ for all $v \in \mathfrak{H}$.

In this notation, the ortogonal projection onto a unit vector e is given by $|e\rangle\langle e|$. We will utilize this and the associativity of the operator composition.

Proposition 5.2.1 *If σ is an operator mean, then $\mathbf{1}\sigma(t\mathbf{1}) = f(t)\mathbf{1}$ is a multiple of the identity. For any $A \in \mathscr{L}(\mathfrak{H})^+$*

$$\mathbf{1}\sigma A = f(A) \tag{5.13}$$

In particular, f is operator monotone with $f(1) = 1$.

Proof: for any operator C, $C = \lambda$ if and only if C commutes with every projection, since if C commutes with every rank one projection P, then $(1 - P)CP = 0$, so $\mathbf{Ran}P$, so every one dimensional subspace of $\mathfrak{H}$ is an invariant subspace of C. This means that we can use the former lemma, f is a non-negative function with $f(1) = 1$ by the normalisation axiom, and f is monotone by the monotonicity axiom.

If that $A = \sum_{k=1}^{n} \alpha_k P_k$ where $\{P_k \mid k \leq n\}$ are mutually orthogonal projections with $1 = \sum_{k=1}^{n} P_k$, then

$$\mathbf{1}\sigma A = \sum_{k=1}^{n} (\mathbf{1}\sigma A)P_k = \sum_{k=1}^{n} (P_k \sigma(AP_k))P_k \tag{5.14}$$

$$= \sum_{k=1}^{n} (P_k \sigma(\alpha_k P_k))P_k = \sum_{k=1}^{n} (\mathbf{1}\sigma(\alpha_k \mathbf{1}))P_k = \sum_{k=1}^{n} f(\alpha_k)P_k = f(A) \tag{5.15}$$

where the second and fourth equation both used lemma 5.2.1.

We can approximate any self-adjoint operator with a decreasing sequence of such maps, because if $\mathfrak{H}$ is infinite dimensional, let $\{e_i \mid in\}$ be an ortonormal basis of it, and let P_i be the orthogonal projection onto $< e_i >$. Define $\pi_k = \sum_{k=1}^{n} P_k$ and $A_n = \pi_k A \pi_k + (\mathbf{1} - \pi_n)\|A\| = \sum_{k=1}^{n} P_k A P_k + (\mathbf{1} - \pi_n)\|A\|$, which is a map in the desired form, since

$$\begin{aligned} P_k A P_k &= (|e_k\rangle\langle e_k|)A(|e_k\rangle\langle e_k|) = |e_k\rangle\langle e_k|(A|e_k\rangle)\langle e_k| \\ &= |e_k\rangle(\langle e_k||Ae_k\rangle)\langle e_k| = |e_k\rangle \left\langle e_k \left| Ae_k \right\rangle \right\langle e_k| \\ &= \left\langle e_k \left| Ae_k \right\rangle \right. P_k \end{aligned}$$

Moreover, $(A_{n-1})_{\pi_n} = (A_n)_{\pi_n} = A_{\pi_n}$, $(A_{n-1})_{\mathbf{1}-\pi_n} \geq (A_n)_{\mathbf{1}-\pi_n} \geq A_{\mathbf{1}-\pi_n}$ with the notation as in Lemma 3.0.2. Since $A_n = (A_n)_{\pi_k} \oplus (A_n)_{\mathbf{1}-\pi_k}$ and

$A = A_{\pi_k} \oplus A_{\mathbf{1}-\pi_k}$, $A_{n-1} \geq A_n \geq A$ holds. Finally, $\pi_k \to \mathbf{1}$ implies $\pi_k A \pi_k \to A$ and $(\mathbf{1} - \pi_k)\|A\| \to 0$, thus $A_n \searrow A$. ■

Proposition 5.2.2 *For any operator mean σ, define f by $f(t) = \mathbf{1}\sigma(t\mathbf{1})$. Then $\forall A \in \mathscr{GL}(\mathfrak{H})^+ \; \forall B \in \mathscr{L}(\mathfrak{H})^+$*

$$A\sigma B = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2} \tag{5.16}$$

Furthermore, (5.10) *holds.*

Proof: With $C = A^{1/2}$ which is invertible for A is, we have by (5.5), that $C[\mathbf{1}\sigma(A^{-1/2}BA^{-1/2})]C$, which is (5.16) by the preceding proposition. If B is invertible, notice that

$$f(B) = a + bB + \int_0^\infty \frac{B}{B + \lambda\mathbf{1}}(1+\lambda)\mathrm{d}\nu(\lambda)$$

$f(B)$ is well-defined, since the invertibility of B implies its lower boundedness, that is we have $c\mathbf{1} \leq B \leq \|B\|\mathbf{1}$. The norm of the integrand can be upper estimated by $\frac{\|B\|}{c}\frac{c}{c+\lambda}(1+\lambda)$, which has an integral $\frac{\|B\|}{c}(f(c) - a - bc)$. Considering $\lambda\frac{B}{B+\lambda\mathbf{1}} = [\lambda^{-1}B^{-1}(B+\lambda\mathbf{1})]^{-1} = (B^{-1} + \lambda^{-1}\mathbf{1})^{-1} = \lambda\mathbf{1} : B$, we see that

$$f(B) = a + bB + \int_0^\infty \frac{1+\lambda}{\lambda}(\lambda\mathbf{1} : B)\mathrm{d}\nu(\lambda)$$

proving (5.10), when $A = \mathbf{1}$ and B is invertible.

If A is invertible as well, considering $\lambda\mathbf{1} = A^{1/2}[\lambda\mathbf{1} : (A^{-1/2}BA^{-1/2})]A^{1/2}$, this integral is combined with (5.16) is (5.10) in the case, when both A and B are invertible. Using the upper semi-continuity of σ and $A + n^{-1}\mathbf{1}, B + n^{-1}\mathbf{1}$, we get the result for general $A, B \in \mathscr{L}(\mathfrak{H})$.

■

With this Propositions (5.2.2) and (5.2.1) proven, the proof of one direction of the Kubo-Ando theorem is done. The converse requires less machinery, hence we can prove it in one proposition.

Proposition 5.2.3 *Let f be a non-negative function on $(0,\infty)$ with $f(1) = 1$, and has the form*

$$f(z) = a + bz + \int_0^\infty \frac{z(1+\lambda)}{z+\lambda} d\nu(\lambda)$$

If σ is defined by

$$A\sigma B = aA + bB + \int_0^\infty \frac{1+\lambda}{\lambda}((\lambda A) : B) d\nu(\lambda)$$

then σ is an operator mean obeying $\mathbf{1}\sigma(t\mathbf{1}) = f(t)\mathbf{1}$.

Proof: For every $\lambda > 0$, $(A, B) \mapsto (\lambda A) : B$ obeys axioms $i - iii$ of operator means.

In order to prove that the integral is well defined, we separate $(0,\infty)$ into $(0,1)$ and $[1,\infty)$. Using the monotonicity of $A \mapsto -A^{-1}$, we show that both integrals are bounded. By the upper semi-continuity of :, we can assume that both A and B are invertible. If $\lambda \in (0,1]$,

$$\begin{aligned}\frac{1+\lambda}{\lambda}((\lambda A) : B) &= \frac{1+\lambda}{\lambda}(\lambda^{-1}A^{-1} + B^{-1})^{-1} \leq \frac{1+\lambda}{\lambda}(\lambda^{-1}A^{-1})^{-1} \\ &= (1+\lambda)A \leq 2A \leq 2\|A\|\end{aligned}$$

If $\lambda \geq 1$

$$\begin{aligned}\frac{1+\lambda}{\lambda}((\lambda A) : B) &= \frac{1+\lambda}{\lambda}(\lambda^{-1}A^{-1} + B^{-1})^{-1} \leq \frac{1+\lambda}{\lambda}(B^{-1})^{-1} \\ &= \frac{1+\lambda}{\lambda}B \leq 2B \leq 2\|B\|\end{aligned}$$

Thus

$$0 \leq \int_0^\infty \frac{1+\lambda}{\lambda}((\lambda A) : B)\mathrm{d}\nu(\lambda) \leq \int_0^1 2\|A\|\mathrm{d}\nu(\lambda) + \int_1^\infty \|B\|\mathrm{d}\nu(\lambda)$$

$$\leq 2\max\{\|A\|, \|B\|\} \int_0^\infty \mathrm{d}\nu(\lambda) = 2\max\{\|A\|, \|B\|\}(f(1) - a - b) < \infty$$

Thus the integral is well defined for every $A, B \in \mathscr{L}(\mathfrak{H})$, and σ obeys axiom $i - iii$. What is more, since $\lambda\mathbf{1} : t\mathbf{1} = \frac{\lambda t}{\lambda + t}$, we conclude that

$$\mathbf{1}\sigma(t\mathbf{1}) = a\mathbf{1} + bt\mathbf{1} + \int_0^\infty \frac{t(1+\lambda)}{\lambda + t}\mathbf{1}\mathrm{d}\nu(\lambda) = f(t)\mathbf{1}$$

Since $f(1) = 1$, σ obeys the normalisation axiom as well. ■

Chapter 6

Abstract

Loewner's Theorem gives a handy condition for deciding whether a function is operator monotone or not. It also asserts that any operator monotone function is real analytic. Even n-monotone functions admit pleasant regularity properties by the behavior of the Dobsch matrix. Although the notion of operator monotonicity comes from finite dimensional definitions, it also inherits the monotonicity in the infinite dimensional case. non-negative operator concave functions on $(0, \infty)$ are precisely the non-negative operator monotone functions on $(0, \infty)$. The operator means are in order isomorphism with the normalised non-negative operator monotone functions on $[0, \infty)$.

Bibliography

[1] K. Löwner, *Über monotone Matrixfunktionen* , Mathematische Zeitschrift 38: 177–216 ,1934

[2] B. Simon *Loewner's Theorem on Monotone Matrix Functions*, Grundlehren der mathematischen Wissenschaften 354, Springer International Publishing, 2019

[3] J. Conway, *A Course in Functional Analysis*, Graduate Texts in Mathematics 96. Springer Verlag, 1990

[4] Otte Heinävaara *Characterizing matrix monotonicity of fixed order on general sets*, Cornell University Thesis, arXiv:1906.06155, 2019

[5] F. Kraus, *Über konvexe Matrixfunktionen*, Mathematische Zeitschrift, 41: 18–42, 1936

[6] J. von Neumann and E. P. Wigner, *Significance of Loewner's theorem in the quantum theory of collisions*, Annals of Mathematics, 59(2): 418-433, 1954

[7] A. Korányi, *On Some Classes of Analytic Functions of Several Variables*, Transactions of the American Mathematical Society, 101(3), 520-554, 1961

[8] F. Kubo and T. Ando, *Means of positive linear operators*, Mathematische Annalen, 246, 205–224, 1980

[9] Sifis Petrakis, *Introduction to Banach Algebras and the Gelfand-Naimark theorems*, Aristotle University of Thessaloniki, 2008