V

Operator Monotone and Operator Convex Functions

In this chapter we study an important and useful class of functions called operator monotone functions. These are real functions whose extensions to Hermitian matrices preserve order. Such functions have several special properties, some of which are studied in this chapter. They are closely related to properties of operator convex functions. We shall study both of these together.

V.1 Definitions and Simple Examples

Let f be a real function defined on an interval I. If $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix whose diagonal entries λ_j are in I, we define $f(D) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$. If A is a Hermitian matrix whose eigenvalues λ_j are in I, we choose a unitary U such that $A = UDU^*$, where D is diagonal, and then define $f(A) = Uf(D)U^*$. In this way we can define $f(A)$ for all Hermitian matrices (of any order) whose eigenvalues are in I. In the rest of this chapter, it will always be assumed that our functions are real functions defined on an interval (finite or infinite, closed or open) and are extended to Hermitian matrices in this way.

We will use the notation $A \leq B$ to mean A and B are Hermitian and $B - A$ is positive. The relation $\leq$ is a partial order on Hermitian matrices.

A function f is said to be **matrix monotone of order n** if it is monotone with respect to this order on $n \times n$ Hermitian matrices, i.e., if $A \leq B$ implies $f(A) \leq f(B)$. If f is matrix monotone of order n for all n we say f is **matrix monotone** or **operator monotone**.

A function f is said to be **matrix convex of order n** if for all $n \times n$ Hermitian matrices A and B and for all real numbers $0 \leq \lambda \leq 1$,

$$f((1-\lambda)A + \lambda B) \leq (1-\lambda)f(A) + \lambda f(B). \tag{V.1}$$

If f is matrix convex of all orders, we say that f is **matrix convex** or **operator convex.**

(Note that if the eigenvalues of A and B are all in an interval I, then the eigenvalues of any convex combination of A, B are also in I. This is an easy consequence of results in Chapter III.)

We will consider continuous functions only. In this case, the condition (V.1) can be replaced by the more special condition

$$f\left(\frac{A+B}{2}\right) \leq \frac{f(A)+f(B)}{2}. \tag{V.2}$$

(Functions satisfying (V.2) are called **mid-point operator convex**, and if they are continuous, then they are convex.)

A function f is called **operator concave** if the function $-f$ is operator convex.

It is clear that the set of operator monotone functions and the set of operator convex functions are both closed under positive linear combinations and also under (pointwise) limits. In other words, if f, g are operator monotone, and if α, β are positive real numbers, then $\alpha f + \beta g$ is also operator monotone. If f_n are operator monotone, and if $f_n(x) \to f(x)$, then f is also operator monotone. The same is true for operator convex functions.

Example V.1.1 *The function $f(t) = \alpha + \beta t$ is operator monotone (on every interval) for every $\alpha \in \mathbb{R}$ and $\beta \geq 0$. It is operator convex for all $\alpha, \beta \in \mathbb{R}$.*

The first surprise is in the following example.

Example V.1.2 *The function $f(t) = t^2$ on $[0, \infty)$ is not operator monotone. In other words, there exist positive matrices A, B such that $B - A$ is positive but $B^2 - A^2$ is not. To see this, take*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Example V.1.3 *The function $f(t) = t^2$ is operator convex on every interval. To see this, note that for any Hermitian matrices A, B,*

$$\frac{A^2+B^2}{2} - \left(\frac{A+B}{2}\right)^2 = \frac{1}{4}(A^2 + B^2 - AB - BA) = \frac{1}{4}(A-B)^2 \geq 0.$$

This shows that the function $f(t) = \alpha + \beta t + \gamma t^2$ is operator convex for all $\alpha, \beta \in \mathbb{R}$, $\gamma \geq 0$.

Example V.1.4 *The function $f(t) = t^3$ on $[0,\infty)$ is not operator convex. To see this, let*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then,

$$\frac{A^3 + B^3}{2} - \left(\frac{A+B}{2}\right)^3 = \begin{pmatrix} 6 & 1 \\ 1 & 0 \end{pmatrix},$$

and this is not positive.

Examples V.1.2 and V.1.4 show that very simple functions which are monotone (convex) as real functions need not be operator monotone (operator convex). A complete description of operator monotone and operator convex functions will be given in later sections. It is instructive to study a few more examples first. The operator monotonicity or convexity of some functions can be proved by special arguments that are useful in other contexts as well.

spectral radius

We will repeatedly use two simple facts. If A is positive, then $A \leq I$ if and only if $\operatorname{spr}(A) \leq 1$. An operator A is a contraction ($\|A\| \leq 1$) if and only if $A^*A \leq I$. This is also equivalent to the condition $AA^* \leq I$.

The following elementary lemma is also used often.

Lemma V.1.5 *If $B \geq A$, then for every operator X we have $X^*BX \geq X^*AX$.*

Proof. For every vector u we have,

$$\langle u, X^*BXu\rangle = \langle Xu, BXu\rangle \geq \langle Xu, AXu\rangle = \langle u, X^*AXu\rangle.$$

This proves the lemma.

An equally brief proof goes as follows. Let C be the positive square root of the positive operator $B - A$. Then

$$X^*(B-A)X = X^*CCX = (CX)^*CX \geq 0.$$

■

Proposition V.1.6 *The function $f(t) = -\frac{1}{t}$ is operator monotone on $(0,\infty)$.*

Proof. Let $B \geq A > 0$. Then, by Lemma V.1.5, $I \geq B^{-1/2}AB^{-1/2}$. Since the map $T \to T^{-1}$ is order-reversing on commuting positive operators, we have $I \leq B^{1/2}A^{-1}B^{1/2}$. Again, using Lemma V.1.5 we get from this $B^{-1} \leq A^{-1}$. ■

Lemma V.1.7 *If $B \geq A \geq 0$ and B is invertible, then $\|A^{1/2}B^{-1/2}\| \leq 1$.*

Proof. If $B \geq A \geq 0$, then $I \geq B^{-1/2}AB^{-1/2} = (A^{1/2}B^{-1/2})^*A^{1/2}B^{-1/2}$, and hence $\|A^{1/2}B^{-1/2}\| \leq 1$. ■

Proposition V.1.8 *The function* $f(t) = t^{1/2}$ *is operator monotone on* $[0, \infty)$.

Proof. Let $B \geq A \geq 0$. Suppose B is invertible. Then, by Lemma V.1.7,

$$1 \geq \|A^{1/2}B^{-1/2}\| \geq \operatorname{spr}(A^{1/2}B^{-1/2}) = \operatorname{spr}(B^{-1/4}A^{1/2}B^{-1/4}).$$

Since $B^{-1/4}AB^{-1/4}$ is positive, this implies that $I \geq B^{-1/4}A^{1/2}B^{-1/4}$. Hence, by Lemma V.1.5, $B^{1/2} \geq A^{1/2}$. This proves the proposition under the assumption that B is invertible. If B is not strictly positive, then for every $\varepsilon > 0, B + \varepsilon I$ is strictly positive. So, $(B + \varepsilon I)^{1/2} \geq A^{1/2}$. Let $\varepsilon \to 0$. This shows that $B^{1/2} \geq A^{1/2}$. ■

Theorem V.1.9 *The function* $f(t) = t^r$ *is operator monotone on* $[0, \infty)$ *for* $0 \leq r \leq 1$.

Proof. Let r be a dyadic rational, i.e., a number of the form $r = \frac{m}{2^n}$, where n is any positive integer and $1 \leq m \leq 2^n$. We will first prove the assertion for such r. This is done by induction on n.

Proposition V.1.8 shows that the assertion of the theorem is true when $n = 1$. Suppose it is also true for all dyadic rationals $\frac{m}{2^j}$, in which $1 \leq j \leq n-1$. Let $B \geq A$ and let $r = \frac{m}{2^n}$. Suppose $m \leq 2^{n-1}$. Then, by the induction hypothesis, $B^{m/2^{n-1}} \geq A^{m/2^{n-1}}$. Hence, by Proposition V.1.8, $B^{m/2^n} \geq A^{m/2^n}$. Suppose $m > 2^{n-1}$. If $B \geq A > 0$, then $A^{-1} \geq B^{-1}$. Using Lemma V.1.5, we have $B^{m/2^n}A^{-1}\,B^{m/2^n} \geq B^{m/2^n}\,B^{-1}B^{m/2^n} = B^{(m/2^{n-1}-1)}$. By the same argument,

$$\begin{aligned} A^{-1/2}\,B^{m/2^n}A^{-1}B^{m/2^n}A^{-1/2} &\geq A^{-1/2}\,B^{(m/2^{n-1}-1)}A^{-1/2} \\ &\geq A^{-1/2}\,A^{(m/2^{n-1}-1)}A^{-1/2} \end{aligned}$$

(by the induction hypothesis). This can be written also as

$$(A^{-1/2}B^{m/2^n}A^{-1/2})^2 \geq A^{(m/2^{n-1}-2)}.$$

So, by the operator monotonicity of the square root,

$$A^{-1/2}B^{m/2^n}A^{-1/2} \geq A^{(m/2^n-1)}.$$

Hence, $B^{m/2^n} \geq A^{m/2^n}$.

We have shown that $B \geq A > 0$ implies $B^r \geq A^r$ for all dyadic rationals r in $[0, 1]$. Such r are dense in $[0, 1]$. So we have $B^r \geq A^r$ for all r in $[0, 1]$. By continuity this is true even when A is positive semidefinite . ■

Exercise V.1.10 *Another proof of Theorem V.1.9 is outlined below. Fill in the details.*

(i) *The composition of two operator monotone functions is operator monotone. Use this and Proposition V.1.6 to prove that the function* $f(t) = \frac{t}{1+t}$ *is operator monotone on* $(0,\infty)$.

(ii) *For each* $\lambda > 0$, *the function* $f(t) = \frac{t}{\lambda+t}$ *is operator monotone on* $(0,\infty)$.

(iii) *One of the integrals calculated by contour integration in Complex Analysis is*

$$\int_0^\infty \frac{\lambda^{r-1}}{1+\lambda} d\lambda = \pi \operatorname{cosec} r\pi, \quad 0 < r < 1. \tag{V.3}$$

By a change of variables, obtain from this the formula

$$t^r = \frac{\sin r\pi}{\pi} \int_0^\infty \frac{t}{\lambda + t} \lambda^{r-1} d\lambda \tag{V.4}$$

valid for all $t > 0$ *and* $0 < r < 1$.

(iv) *Thus, we can write*

$$t^r = \int_0^\infty \frac{t}{\lambda+t}\, d\mu(\lambda), \quad 0 < r < 1, \tag{V.5}$$

where μ *is a positive measure on* $(0,\infty)$. *Now use (ii) to conclude that the function* $f(t) = t^r$ *is operator monotone on* $(0,\infty)$ *for* $0 \le r \le 1$.

Example V.1.11 *The function* $f(t) = |t|$ *is not operator convex on any interval that contains 0. To see this, take*

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$|A| = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad |A| + |B| = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}.$$

But $|A+B| = \sqrt{2}\, I$. *So* $|A|+|B|-|A+B|$ *is not positive. (See also Exercise III.5.7.)*

Example V.1.12 *The function* $f(t) = t \vee 0$ *is not operator convex on any interval that contains 0. To see this, take* A, B *as in Example V.1.11. Since the eigenvalues of* A *are* -2 *and 0,* $f(A) = 0$. *So* $\frac{1}{2}(f(A) + f(B)) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. *Any positive matrix dominated by this must have* $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ *as an eigenvector with 0 as the corresponding eigenvalue. Since* $\frac{1}{2}(A + B)$ *does not have* $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ *as an eigenvector, neither does* $f(\frac{A+B}{2})$.

Exercise V.1.13 *Let I be any interval. For $a \in I$, let $f(t) = (t-a) \vee 0$. Then f is called an "angle function" angled at a. If I is a finite interval, then every convex function on I is a limit of positive linear combinations of linear functions and angle functions. Use this to show that angle functions are not operator convex.*

Exercise V.1.14 *Show that the function $f(t) = t \vee 0$ is not operator monotone on any interval that contains 0.*

Exercise V.1.15 *Let A, B be positive. Show that*

$$\frac{A^{-1}+B^{-1}}{2} - \left(\frac{A+B}{2}\right)^{-1} = \frac{(A^{-1}-B^{-1})(A^{-1}+B^{-1})^{-1}(A^{-1}-B^{-1})}{2}.$$

Therefore, the function $f(t) = \frac{1}{t}$ is operator convex on $(0,\infty)$.

V.2 Some Characterisations

There are several different notions of averaging in the space of operators. In this section we study the relationship between some of these operations and operator convex functions. This leads to some characterisations of operator convex and operator monotone functions and to the interrelations between them.

In the proofs that are to follow, we will frequently use properties of operators on the direct sum $\mathcal{H} \oplus \mathcal{H}$ to draw conclusions about operators on $\mathcal{H}$. This technique was outlined briefly in Section I.3.

Let K be a contraction on $\mathcal{H}$. Let $L = (I-KK^*)^{1/2}$, $M = (I-K^*K)^{1/2}$. Then the operators U, V defined as

$$U = \begin{pmatrix} K & L \\ M & -K^* \end{pmatrix}, \quad V = \begin{pmatrix} K & -L \\ M & K^* \end{pmatrix} \tag{V.6}$$

are unitary operators on $\mathcal{H} \oplus \mathcal{H}$. (See Exercise I.3.6.) More specially, for each $0 \le \lambda \le 1$, the operator

$$W = \begin{pmatrix} \lambda^{1/2} I & -(1-\lambda)^{1/2} I \\ (1-\lambda)^{1/2} I & \lambda^{1/2} I \end{pmatrix} \tag{V.7}$$

is a unitary operator on $\mathcal{H} \oplus \mathcal{H}$.

Theorem V.2.1 *Let f be a real function on an interval I. Then the following two statements are equivalent:*

(i) f is operator convex on I.

(ii) $f(\mathcal{C}(A)) \le \mathcal{C}(f(A))$ for every Hermitian operator A (on a Hilbert space $\mathcal{H}$) whose spectrum is contained in I and for every pinching $\mathcal{C}$ (in the space $\mathcal{H}$).

Proof. (i) $\Rightarrow$(ii): Every pinching is a product of pinchings by two complementary projections. (See Problems II.5.4 and II.5.5.) So we need to prove this implication only for pinchings $\mathcal{C}$ of the form

$$\mathcal{C}(X) = \frac{X + U^*XU}{2}, \quad \text{where } U = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

For such a $\mathcal{C}$

$$\begin{aligned} f(\mathcal{C}(A)) &= f\left(\frac{A+U^*AU}{2}\right) \le \frac{f(A)+f(U^*AU)}{2} \\ &= \frac{f(A)+U^*f(A)U}{2} = \mathcal{C}(f(A)). \end{aligned}$$

(ii) $\Rightarrow$ (i): Let A, B be Hermitian operators on $\mathcal{H}$, both having their spectrum in I. Consider the operator $T = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ on $\mathcal{H} \oplus \mathcal{H}$. If W is the unitary operator defined in (V.7), then the diagonal entries of W^*TW are $\lambda A + (1-\lambda)B$ and $(1-\lambda)A + \lambda B$. So if $\mathcal{C}$ is the pinching on $\mathcal{H} \oplus \mathcal{H}$ induced by the projections onto the two summands, then

$$\mathcal{C}(W^*TW) = \begin{pmatrix} \lambda A + (1-\lambda)B & 0 \\ 0 & (1-\lambda)A + \lambda B \end{pmatrix}.$$

By the same argument,

$$\begin{aligned} \mathcal{C}(f(W^*TW)) &= \mathcal{C}(W^*f(T)W) \\ &= \begin{pmatrix} \lambda f(A) + (1-\lambda)f(B) & 0 \\ 0 & (1-\lambda)f(A) + \lambda f(B) \end{pmatrix}. \end{aligned}$$

So the condition $f(\mathcal{C}(W^*TW)) \le \mathcal{C}(f(W^*TW))$ implies that

$$f(\lambda A + (1-\lambda)B) \le \lambda f(A) + (1-\lambda)f(B).$$

■

Exercise V.2.2 *The following conditions are equivalent:*

(i) f is operator convex on I.

(ii) $f(A_{\mathcal{M}}) \le (f(A))_{\mathcal{M}}$ for every Hermitian operator A with its spectrum in I, and for every compression $T \to T_{\mathcal{M}}$.

(iii) $f(V^*AV) \le V^*f(A)V$ for every Hermitian operator A (on $\mathcal{H}$) with its spectrum in I, and for every isometry from any Hilbert space into $\mathcal{H}$.

(See Section III.1 for the definition of a compression.)

Theorem V.2.3 *Let I be an interval containing 0 and let f be a real function on I. Then the following conditions are equivalent:*

(i) *f is operator convex on I and $f(0) \leq 0$.*

(ii) *$f(K^*AK) \leq K^*f(A)K$ for every contraction K and every Hermitian operator A with spectrum in I.*

(iii) *$f(K_1^*AK_1 + K_2^*BK_2) \leq K_1^*f(A)K_1 + K_2^*f(B)K_2$ for all operators K_1, K_2 such that $K_1^*K_1 + K_2^*K_2 \leq I$ and for all Hermitian A, B with spectrum in I.*

(iv) *$f(PAP) \leq Pf(A)P$ for all projections P and Hermitian operators A with spectrum in I.*

Proof. (i) $\Rightarrow$ (ii): Let $T = \left(\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and let U, V be the unitary operators defined in (V.6). Then

$$U^*TU = \begin{pmatrix} K^*AK & K^*AL \\ LAK & LAL \end{pmatrix}, \quad V^*TV = \begin{pmatrix} K^*AK & -K^*AL \\ -LAK & LAL \end{pmatrix}.$$

So,

$$\begin{pmatrix} K^*AK & 0 \\ 0 & LAL \end{pmatrix} = \frac{U^*TU + V^*TV}{2}.$$

Hence,

$$\begin{aligned} &\begin{pmatrix} f(K^*AK) & 0 \\ 0 & f(LAL) \end{pmatrix} \\ &= f\left(\frac{U^*TU + V^*TV}{2}\right) \\ &\leq \frac{f(U^*TU) + f(V^*TV)}{2} \\ &= \frac{U^*f(T)U + V^*f(T)V}{2} \\ &= \frac{1}{2}\left\{U^*\begin{pmatrix} f(A) & 0 \\ 0 & f(0) \end{pmatrix}U + V^*\begin{pmatrix} f(A) & 0 \\ 0 & f(0) \end{pmatrix}V\right\} \\ &\leq \frac{1}{2}\left\{U^*\begin{pmatrix} f(A) & 0 \\ 0 & 0 \end{pmatrix}U + V^*\begin{pmatrix} f(A) & 0 \\ 0 & 0 \end{pmatrix}V\right\} \\ &= \begin{pmatrix} K^*f(A)K & 0 \\ 0 & Lf(A)L \end{pmatrix}. \end{aligned}$$

Hence, $f(K^*AK) \leq K^*f(A)K$.

(ii) $\Rightarrow$ (iii): Let $T = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, $K = \begin{pmatrix} K_1 & 0 \\ K_2 & 0 \end{pmatrix}$. Then K is a contraction. Note that

$$K^*TK = \begin{pmatrix} K_1^*AK_1 + K_2^*BK_2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} f(K_1^*AK_1 + K_2^*BK_2) & 0 \\ 0 & f(0) \end{pmatrix} = f(K^*TK) \le K^*f(T)K$$
$$= \begin{pmatrix} K_1^*f(A)K_1 + K_2^*f(B)K_2 & 0 \\ 0 & 0 \end{pmatrix}.$$

(iii) $\Rightarrow$ (iv) obviously.

(iv) $\Rightarrow$(i): Let A, B be Hermitian operators with spectrum in I and let $0 \le \lambda \le 1$. Let $T = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, $P = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$ and let W be the unitary operator defined by (V.7). Then

$$PW^*TWP = \begin{pmatrix} \lambda A + (1-\lambda)B & 0 \\ 0 & 0 \end{pmatrix}.$$

So,

$$\begin{aligned} \begin{pmatrix} f(\lambda A + (1-\lambda)B) & 0 \\ 0 & f(0) \end{pmatrix} &= f(PW^*TWP) \\ &\le Pf(W^*TW)P = PW^*f(T)WP \\ &= \begin{pmatrix} \lambda f(A) + (1-\lambda)f(B) & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Hence, f is operator convex and $f(0) \le 0$. ∎

Exercise V.2.4 *(i) Let λ_1, λ_2 be positive real numbers such that $\lambda_1\lambda_2 \ge C^*C$. Then $\begin{pmatrix} \lambda_1 I & C^* \\ C & \lambda_2 I \end{pmatrix}$ is positive. (Use Proposition I.3.5.)*

(ii) Let $\begin{pmatrix} A & C^* \\ C & B \end{pmatrix}$ be a Hermitian operator. Then for every $\varepsilon > 0$, there exists $\lambda > 0$ such that

$$\begin{pmatrix} A & C^* \\ C & B \end{pmatrix} \le \begin{pmatrix} A + \varepsilon I & 0 \\ 0 & \lambda I \end{pmatrix}.$$

The next two theorems are among the several results that describe the connections between operator convexity and operator monotonicity.

Theorem V.2.5 *Let f be a (continuous) function mapping the positive half-line $[0, \infty)$ into itself. Then f is operator monotone if and only if it is operator concave.*

Proof. Suppose f is operator monotone. If we show that $f(K^*AK) \ge K^*f(A)K$ for every positive operator A and contraction K, then it would follow from Theorem V.2.3 that f is operator concave. Let $T = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$ and let U be the unitary operator defined in (V.6). Then $U^*TU = \begin{pmatrix} K^*AK & K^*AL \\ LAK & LAL \end{pmatrix}$.

By the assertion in Exercise V.2.4(ii), given any $\varepsilon > 0$, there exists $\lambda > 0$ such that

$$U^*TU \leq \begin{pmatrix} K^*AK + \varepsilon & 0 \\ 0 & \lambda I \end{pmatrix}.$$

Replacing T by $f(T)$, we get

$$\begin{aligned} \begin{pmatrix} K^*f(A)K & K^*f(A)L \\ Lf(A)K & Lf(A)L \end{pmatrix} &= U^*f(T)U = f(U^*TU) \\ &\leq \begin{pmatrix} f(K^*AK + \varepsilon) & 0 \\ 0 & f(\lambda)I \end{pmatrix} \end{aligned}$$

by the operator monotonicity of f. In particular, this shows $K^*f(A)K \leq f(K^*AK + \varepsilon)$ for every $\varepsilon > 0$. Hence $K^*f(A)K \leq f(K^*AK)$.

Conversely, suppose f is operator concave. Let $0 \leq A \leq B$. Then for any $0 < \lambda < 1$ we can write

$$\lambda B = \lambda A + (1 - \lambda)\frac{\lambda}{1 - \lambda}(B - A).$$

Since f is operator concave, this gives

$$f(\lambda B) \geq \lambda f(A) + (1 - \lambda)f\left(\frac{\lambda}{1 - \lambda}(B - A)\right).$$

Since $f(X)$ is positive for every positive X, it follows that $f(\lambda B) \geq \lambda f(A)$. Now let $\lambda \to 1$. This shows $f(B) \geq f(A)$. So f is operator monotone. ■

Corollary V.2.6 *Let f be a continuous function from $(0, \infty)$ into itself. If f is operator monotone then the function $g(t) = \frac{1}{f(t)}$ is operator convex.*

Proof. Let A, B be positive operators. Since f is operator concave, $f\left(\frac{A+B}{2}\right) \geq \frac{f(A)+f(B)}{2}$. Since the map $X \to X^{-1}$ is order-reversing and convex on positive operators (see Proposition V.1.6 and Exercise V.1.15), this gives

$$\left[f\left(\frac{A + B}{2}\right)\right]^{-1} \leq \left[\frac{f(A) + f(B)}{2}\right]^{-1} \leq \frac{f(A)^{-1} + f(B)^{-1}}{2}.$$

This is the same as saying g is operator convex. ■

Exercise V.2.7 *Let I be an interval containing 0, and let f be a real function on I with $f(0) \leq 0$. Show that for every Hermitian operator A with spectrum in I and for every projection P*

$$f(PAP) \leq Pf(PAP) = Pf(PAP)P.$$

Exercise V.2.8 *Let f be a continuous real function on $[0,\infty)$. Then for all positive operators A and projections P*

$$f(A^{1/2}PA^{1/2})A^{1/2}P = A^{1/2}Pf(PAP).$$

(Prove this first, by induction, for $f(t) = t^n$. Then use the Weierstrass approximation theorem to show that this is true for all f.)

Theorem V.2.9 *Let f be a (continuous) real function on the interval $[0,\alpha)$. Then the following two conditions are equivalent:*
(i) f is operator convex and $f(0) \le 0$.
(ii) The function $g(t) = f(t)/t$ is operator monotone on $(0,\alpha)$.

Proof. (i) $\Rightarrow$ (ii): Let $0 < A \le B$. Then $0 < A^{1/2} \le B^{1/2}$. Hence, $B^{-1/2}A^{1/2}$ is a contraction by Lemma V.1.7. Therefore, using Theorem V.2.3 we see that

$$f(A) = f(A^{1/2}B^{-1/2}BB^{-1/2}A^{1/2}) \le A^{1/2}B^{-1/2}f(B)B^{-1/2}A^{1/2}.$$

From this, one obtains, using Lemma V.1.5,

$$A^{-1/2}f(A)A^{-1/2} \le B^{-1/2}f(B)B^{-1/2}.$$

Since all functions of an operator commute with each other, this shows that $A^{-1}f(A) \le B^{-1}f(B)$. Thus, g is operator monotone.

(ii) $\Rightarrow$(i): If $f(t)/t$ is monotone on $(0,\alpha)$ we must have $f(0) \le 0$. We will show that f satisfies the condition (iv) of Theorem V.2.3. Let P be any projection and let A be any positive operator with spectrum in $(0,\alpha)$. Then there exists an $\varepsilon > 0$ such that $(1+\varepsilon)A$ has its spectrum in $(0,\alpha)$. Since $P+\varepsilon I \le (1+\varepsilon)I$, we have $A^{1/2}(P+\varepsilon I)A^{1/2} \le (1+\varepsilon)A$. So, by the operator monotonicity of g, we have

$$A^{-1/2}(P+\varepsilon I)^{-1}A^{-1/2}f(A^{1/2}(P+\varepsilon I)A^{1/2}) \le (1+\varepsilon)^{-1}A^{-1}f((1+\varepsilon)A).$$

Multiply both sides on the right by $A^{1/2}(P+\varepsilon I)$ and on the left by its conjugate $(P+\varepsilon I)A^{1/2}$. This gives

$$A^{-1/2}f(A^{1/2}(P+\varepsilon I)A^{1/2})A^{1/2}(P+\varepsilon I) \le (1+\varepsilon)^{-1}(P+\varepsilon I)f((1+\varepsilon)A)(P+\varepsilon I).$$

Let $\varepsilon \to 0$. This gives

$$A^{-1/2}f(A^{1/2}PA^{1/2})A^{1/2}P \le Pf(A)P.$$

Use the identity in Exercise V.2.8 to reduce this to $Pf(PAP) \le Pf(A)P$, and then use the inequality in Exercise V.2.7 to conclude that $f(PAP) \le Pf(A)P$, as desired. ∎

As corollaries to the above results, we deduce the following statements about the power functions .

Theorem V.2.10 *On the positive half-line* $(0, \infty)$ *the functions* $f(t) = t^r$, *where* r *is a real number, are operator monotone if and only if* $0 \le r \le 1$.

Proof. If $0 \le r \le 1$, we know that $f(t) = t^r$ is operator monotone by Theorem V.1.9. If r is not in $[0, 1]$, then the function $f(t) = t^r$ is not concave on $(0, \infty)$. Therefore, it cannot be operator monotone by Theorem V.2.5. ■

Exercise V.2.11 *Consider the functions* $f(t) = t^r$ *on* $(0, \infty)$. *Use Theorems V.2.9 and V.2.10 to show that if* $r \ge 0$, *then* $f(t)$ *is operator convex if and only if* $1 \le r \le 2$. *Use Corollary V.2.6 to show that* $f(t)$, *is operator convex if* $-1 \le r \le 0$. *(We will see later that* $f(t)$ *is not operator convex for any other value of* r.*)*

Exercise V.2.12 *A function* f *from* $(0, \infty)$ *into itself is both operator monotone and operator convex if and only if it is of the form* $f(t) = \alpha + \beta t, \;\; \alpha, \beta \ge 0$.

Exercise V.2.13 *Show that the function* $f(t) = -t \log t$ *is operator concave on* $(0, \infty)$.

V.3 Smoothness Properties

Let I be the open interval $(-1, 1)$. Let f be a continuously differentiable function on I. Then we denote by $f^{[1]}$ the function on $I \times I$ defined as

$$
\begin{aligned}
f^{[1]}(\lambda, \mu) &= \frac{f(\lambda) - f(\mu)}{\lambda - \mu}, \quad \text{if } \lambda \neq \mu \\
f^{[1]}(\lambda, \lambda) &= f'(\lambda).
\end{aligned}
$$

The expression $f^{[1]}(\lambda, \mu)$ is called the **first divided difference** of f at (λ, μ).

If Λ is a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, all of which are in I, we denote by $f^{[1]}(\Lambda)$ the $n \times n$ symmetric matrix whose (i, j)-entry is $f^{[1]}(\lambda_i, \lambda_j)$. If A is Hermitian and $A = U \Lambda U^*$, let $f^{[1]}(A) = U f^{[1]}(\Lambda) U^*$.

Now consider the induced map f on the set of Hermitian matrices with eigenvalues in I. Such matrices form an open set in the real vector space of all Hermitian matrices. The map f is called (Fréchet) **differentiable** at A if there exists a linear transformation $Df(A)$ on the space of Hermitian matrices such that for all H

$$\|f(A + H) - f(A) - Df(A)(H)\| = o(\|H\|). \tag{V.8}$$

The linear operator $Df(A)$ is then called the **derivative** of f at A. Basic rules of the Fréchet differential calculus are summarised in Chapter 10. If

f is differentiable at A, then

$$Df(A)(H) = \frac{d}{dt}\bigg|_{t=0} f(A+tH). \tag{V.9}$$

There is an interesting relationship between the derivative $Df(A)$ and the matrix $f^{[1]}(A)$. This is explored in the next few paragraphs.

Lemma V.3.1 *Let f be a polynomial function. Then for every diagonal matrix Λ and for every Hermitian matrix H,*

$$Df(\Lambda)(H) = f^{[1]}(\Lambda) \circ H, \tag{V.10}$$

where $\circ$ stands for the Schur-product of two matrices.

Proof. Both sides of (V.10) are linear in f. Therefore, it suffices to prove this for the powers $f(t) = t^p, p = 1, 2, 3, \ldots$ For such f, using (V.9) one gets

$$Df(\Lambda)(H) = \sum_{k=1}^{p} \Lambda^{k-1} H \Lambda^{p-k}.$$

This is a matrix whose (i,j)-entry is $\sum\limits_{k=1}^{p} \lambda_i^{k-1}\lambda_j^{p-k} h_{ij}$. On the other hand, the (i,j)-entry of $f^{[1]}(\Lambda)$ is $\sum\limits_{k=1}^{p} \lambda_i^{k-1}\lambda_j^{p-k}$. ■

Corollary V.3.2 *If $A = U\Lambda U^*$ and f is a polynomial function, then*

$$Df(A)(H) = U[f^{[1]}(\Lambda) \circ (U^*HU)]U^*. \tag{V.11}$$

Proof. Note that

$$\frac{d}{dt}\bigg|_{t=0} f(U\Lambda U^* + tH) = U\left[\frac{d}{dt}\bigg|_{t=0} f(\Lambda + tU^*HU)\right]U^*,$$

and use (V.10). ■

Theorem V.3.3 *Let $f \in C^1(I)$ and let A be a Hermitian matrix with all its eigenvalues in I. Then*

$$Df(A)(H) = f^{[1]}(A) \circ H, \tag{V.12}$$

where $\circ$ denotes the Schur-product in a basis in which A is diagonal.

Proof. Let $A = U\Lambda U^*$, where Λ is diagonal. We want to prove that

$$Df(A)(H) = U[f^{[1]}(\Lambda) \circ (U^*HU)]U^*. \tag{V.13}$$

This has been proved for all polynomials f. We will extend its validity to all $f \in C^1$ by a continuity argument.

Denote the right-hand side of (V.13) by $\mathcal{D}f(A)(H)$. For each f in C^1, $\mathcal{D}f(A)$ is a linear map on Hermitian matrices. We have

$$\|\mathcal{D}f(A)(H)\|_2 = \|f^{[1]}(\Lambda) \circ (U^*HU)\|_2.$$

All entries of the matrix $f^{[1]}(\Lambda)$ are bounded by $\max\limits_{|t|\le\|A\|}|f'(t)|$. (Use the mean value theorem.) Hence

$$\|\mathcal{D}f(A)(H)\|_2 \le \max_{|t|\le\|A\|}|f'(t)|\ \|H\|_2. \tag{V.14}$$

Let H be a Hermitian matrix with norm so small that the eigenvalues of $A+H$ are in I. Let $[a,b]$ be a closed interval in I containing the eigenvalues of both A and $A+H$. Choose a sequence of polynomials f_n such that $f_n \to f$ and $f_n' \to f'$ uniformly on $[a,b]$. Let $\mathcal{L}$ be the line segment joining A and $A+H$ in the space of Hermitian matrices. Then, by the mean value theorem (for Fréchet derivatives), we have

$$\begin{aligned} &\|f_m(A+H) - f_n(A+H) - (f_m(A) - f_n(A))\| \\ &\le \quad \|H\| \sup_{X\in\mathcal{L}} \|Df_m(X) - Df_n(X)\| \\ &= \quad \|H\| \sup_{X\in\mathcal{L}} \|\mathcal{D}f_m(X) - \mathcal{D}f_n(X)\|. \end{aligned} \tag{V.15}$$

This is so because we have already shown that $Df_n = \mathcal{D}f_n$ for the polynomial functions f_n.

Let ε be any positive real number. The inequality (V.14) ensures that there exists a positive integer n_0 such that for $m,n \ge n_0$ we have

$$\sup_{X\in\mathcal{L}} \|\mathcal{D}f_m(X) - \mathcal{D}f_n(X)\| \le \frac{\varepsilon}{3} \tag{V.16}$$

and

$$\|\mathcal{D}f_n(A) - \mathcal{D}f(A)\| \le \frac{\varepsilon}{3}. \tag{V.17}$$

Let $m \to \infty$ and use (V.15) and (V.16) to conclude that

$$\|f(A+H) - f(A) - (f_n(A+H) - f_n(A))\| \le \frac{\varepsilon}{3}\|H\|. \tag{V.18}$$

If $\|H\|$ is sufficiently small, then by the definition of the Fréchet derivative, we have

$$\|f_n(A+H) - f_n(A) - \mathcal{D}f_n(A)(H)\| \le \frac{\varepsilon}{3}\|H\|. \tag{V.19}$$

Now we can write, using the triangle inequality,

$$\begin{aligned} &\|f(A+H) - f(A) - \mathcal{D}f(A)(H)\| \\ &\le \quad \|f(A+H) - f(A) - (f_n(A+H) - f_n(A))\| \\ &\quad + \|f_n(A+H) - f_n(A) - \mathcal{D}f_n(A)(H)\| \\ &\quad + \|(\mathcal{D}f(A) - \mathcal{D}f_n(A))(H)\|, \end{aligned}$$

and then use (V.17), (V.18), and (V.19) to conclude that, for $\|H\|$ sufficiently small, we have

$$\|f(A+H)-f(A)-\mathcal{D}f(A)(H)\| \le \varepsilon\|H\|.$$

But this says that $Df(A)=\mathcal{D}f(A)$. ■

Let $t \to A(t)$ be a C^1 map from the interval $[0,1]$ into the space of Hermitian matrices that have all their eigenvalues in I. Let $f \in C^1(I)$, and let $F(t) = f(A(t))$. Then, by the chain rule, $Df(t) = DF(A(t))(A'(t))$. Therefore, by the theorem above, we have

$$F(1)-F(0)=\int_0^1 f^{[1]}(A(t))\circ A'(t)dt, \tag{V.20}$$

where for each t the Schur-product is taken in a basis that diagonalises $A(t)$.

Theorem V.3.4 *Let $f \in C^1(I)$. Then f is operator monotone on I if and only if, for every Hermitian matrix A whose eigenvalues are in I, the matrix $f^{[1]}(A)$ is positive.*

Proof. Let f be operator monotone, and let A be a Hermitian matrix whose eigenvalues are in I. Let H be the matrix all whose entries are 1. Then H is positive. So, $A+tH \ge A$ if $t \ge 0$. Hence, $f(A+tH)-f(A)$ is positive for small positive t. This implies that $Df(A)(H) \ge 0$. So, by Theorem V.3.3, $f^{[1]}(A)\circ H \ge 0$. But, for this special choice of H, this just says that $f^{[1]}(A) \ge 0$.

To prove the converse, let A, B be Hermitian matrices whose eigenvalues are in I, and let $B \ge A$. Let $A(t) = (1-t)A + tB$, $0 \le t \le 1$. Then $A(t)$ also has all its eigenvalues in I. So, by the hypothesis, $f^{[1]}(A(t)) \ge 0$ for all t. Note that $A'(t) = B - A \ge 0$, for all t. Since the Schur-product of two positive matrices is positive, $f^{[1]}(A(t))\circ A'(t)$ is positive for all t. So, by (V.20), $f(B)-f(A) \ge 0$. ■

Lemma V.3.5 *If f is continuous and operator monotone on $(-1,1)$, then for each $-1 \le \lambda \le 1$ the function $g_\lambda(t) = (t+\lambda)f(t)$ is operator convex.*

Proof. We will prove this using Theorem V.2.9. First assume that f is continuous and operator monotone on $[-1,1]$. Then the function $f(t-1)$ is operator monotone on $[0,2)$. Let $g(t) = tf(t-1)$. Then $g(0)=0$ and the function $g(t)/t$ is operator monotone on $(0,2)$. Hence, by Theorem V.2.9, $g(t)$ is operator convex on $[0,2)$. This implies that the function $h_1(t) = g(t+1) = (t+1)f(t)$ is operator convex on $[-1,1)$. Instead of $f(t)$, if the same argument is applied to the function $-f(-t)$, which is also operator monotone on $[-1,1]$, we see that the function $h_2(t) = -(t+1)f(-t)$ is operator convex on $[-1,1)$. Changing t to $-t$ preserves convexity. So the function $h_3(t) = h_2(-t) = (t-1)f(t)$ is also operator convex. But for $|\lambda| \le 1$, $g_\lambda(t) = \frac{1+\lambda}{2}h_1(t) + \frac{1-\lambda}{2}h_3(t)$ is a convex combination of h_1 and h_3. So g_λ is also operator convex.

Now, given f continuous and operator monotone on $(-1,1)$, the function $f((1-\varepsilon)t)$ is continuous and operator monotone on $[-1,1]$ for each $\varepsilon > 0$. Hence, by the special case considered above, the function $(t+\lambda)f((1-\varepsilon)t)$ is operator convex. Let $\varepsilon \to 0$, and conclude that the function $(t+\lambda)f(t)$ is operator convex. ■

The next theorem says that every operator monotone function on I is in the class C^1. Later on, we will see that it is actually in the class C^∞. (This is so even if we do not assume that it is continuous to begin with.) In the proof we make use of some differentiability properties of convex functions and smoothing techniques. For the reader's convenience, these are summarised in Appendices 1 and 2 at the end of the chapter.

Theorem V.3.6 *Every operator monotone function f on I is continuously differentiable.*

Proof. Let $0 < \varepsilon < 1$, and let f_ε be a regularisation of f of order ε. (See Appendix 2.) Then f_ε is a C^∞ function on $(-1+\varepsilon, 1-\varepsilon)$. It is also operator monotone. Let $\tilde{f}(t) = \lim\limits_{\varepsilon\to 0} f_\varepsilon(t)$. Then $\tilde{f}(t) = \frac{1}{2}[f(t+) + f(t-)]$.

Let $g_\varepsilon(t) = (t+1)f_\varepsilon(t)$. Then, by Lemma V.3.5, g_ε is operator convex. Let $\tilde{g}(t) = \lim\limits_{\varepsilon\to 0} g_\varepsilon(t)$. Then $\tilde{g}(t)$ is operator convex. But every convex function (on an open interval) is continuous. So $\tilde{g}(t)$ is continuous. Since $\tilde{g}(t) = (t+1)\tilde{f}(t)$ and $t+1 > 0$ on I, this means that $\tilde{f}(t)$ is continuous. Hence $\tilde{f}(t) = f(t)$. We thus have shown that f is continuous.

Let $g(t) = (t+1)f(t)$. Then g is a convex function on I. So g is left and right differentiable and the one-sided derivatives satisfy the properties

$$g'_-(t) \le g'_+(t), \quad \lim_{s\downarrow t} g'_\pm(s) = g'_+(t), \quad \lim_{s\uparrow t} g'_\pm(s) = g'_-(t). \tag{V.21}$$

But $g'_\pm(t) = f(t) + (t+1)f'_\pm(t)$. Since $t+1 > 0$, the derivatives $f'_\pm(t)$ also satisfy relations like (V.21).

Now let $A = \begin{pmatrix} s & 0 \\ 0 & t \end{pmatrix}$, $s,t \in (-1,1)$. If ε is sufficiently small, s,t are in $(-1+\varepsilon, 1-\varepsilon)$. Since f_ε is operator monotone on this interval, by Theorem V.3.4, the matrix $f_\varepsilon^{[1]}(A)$ is positive. This implies that

$$\left(\frac{f_\varepsilon(s) - f_\varepsilon(t)}{s-t}\right)^2 \le f'_\varepsilon(s)f'_\varepsilon(t).$$

Let $\varepsilon \to 0$. Since $f_\varepsilon \to f$ uniformly on compact sets, $f_\varepsilon(s) - f_\varepsilon(t)$ converges to $f(s) - f(t)$. Also, $f'_\varepsilon(t)$ converges to $\frac{1}{2}[f'_+(t) + f'_-(t)]$. Therefore, the above inequality gives, in the limit, the inequality

$$\left(\frac{f(s)-f(t)}{s-t}\right)^2 \leq \frac{1}{4}[f'_+(s)+f'_-(s)][f'_+(t)+f'_-(t)].$$

Now let $s \downarrow t$, and use the fact that the derivatives of f satisfy relations like (V.21). This gives

$$[f'_+(t)]^2 \leq \frac{1}{4}[f'_+(t)+f'_+(t)][f'_+(t)+f'_-(t)],$$

which implies that $f'_+(t) = f'_-(t)$. Hence f is differentiable. The relations (V.21), which are satisfied by f too, show that f' is continuous. ■

Just as monotonicity of functions can be studied via first divided differences, convexity requires **second divided differences**. These are defined as follows. Let f be twice continuously differentiable on the interval I. Then $f^{[2]}$ is a function defined on $I \times I \times I$ as follows. If $\lambda_1, \lambda_2, \lambda_3$ are distinct

$$f^{[2]}(\lambda_1,\lambda_2,\lambda_3) = \frac{f^{[1]}(\lambda_1,\lambda_2) - f^{[1]}(\lambda_1,\lambda_3)}{\lambda_2-\lambda_3}.$$

For other values of $\lambda_1, \lambda_2, \lambda_3$, $f^{[2]}$ is defined by continuity; e.g.,

$$f^{[2]}(\lambda,\lambda,\lambda) = \frac{1}{2}f''(\lambda).$$

Exercise V.3.7 *Show that if $\lambda_1, \lambda_2, \lambda_3$ are distinct, then $f^{[2]}(\lambda_1,\lambda_2,\lambda_3)$ is the quotient of the two determinants*

$$\begin{vmatrix} f(\lambda_1) & f(\lambda_2) & f(\lambda_3) \\ \lambda_1 & \lambda_2 & \lambda_3 \\ 1 & 1 & 1 \end{vmatrix} \quad \textit{and} \quad \begin{vmatrix} \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ 1 & 1 & 1 \end{vmatrix}.$$

Hence the function $f^{[2]}$ is symmetric in its three arguments.

Exercise V.3.8 *If $f(t) = t^m$, $m = 2, 3, \ldots$, show that*

$$f^{[2]}(\lambda_1,\lambda_2,\lambda_3) = \sum_{\substack{0 \leq p,q,r \\ p+q+r=m-2}} \lambda_1^p \lambda_2^q \lambda_3^r.$$

Exercise V.3.9 *(i) Let $f(t) = t^m, m \geq 2$. Let A be an $n \times n$ diagonal matrix; $A = \sum\limits_{i=1}^{n} \lambda_i P_i$, where P_i are the projections onto the coordinate axes. Show that for every H*

$$\begin{aligned} \left.\frac{d^2}{dt^2}\right|_{t=0} f(A+tH) &= 2 \sum_{p+q+r=m-2} A^p H A^q H A^r \\ &= 2 \sum_{p+q+r=m-2} \sum_{1 \leq i,j,k \leq n} \lambda_i^p \lambda_j^q \lambda_k^r P_i H P_j H P_k, \end{aligned}$$

and

$$\left.\frac{d^2}{dt^2}\right|_{t=0} f(A+tH) = 2\sum_{i,j,k} f^{[2]}(\lambda_i,\lambda_j,\lambda_k)P_iHP_jHP_k. \tag{V.22}$$

(ii) Use a continuity argument, like the one used in the proof of Theorem V.3.3, to show that this last formula is valid for all C^2 functions f.

Theorem V.3.10 *If $f \in C^2(I)$ and f is operator convex, then for each $\mu \in I$ the function $g(\lambda) = f^{[1]}(\mu, \lambda)$ is operator monotone.*

Proof. Since f is in the class C^2, g is in the class C^1. So, by Theorem V.3.4, it suffices to prove that, for each n, the $n \times n$ matrix with entries $g^{[1]}(\lambda_i, \lambda_j)$ is positive for all $\lambda_1, \ldots, \lambda_n$ in I.

Fix n and choose any $\lambda_1, \ldots, \lambda_{n+1}$ in I. Let A be the diagonal matrix with entries $\lambda_1, \ldots \lambda_{n+1}$. Since f is operator convex and is twice differentiable, for every Hermitian matrix H, the matrix $\left.\frac{d^2}{dt^2}\right|_{t=0} f(A+tH)$ must be positive. If we write $P_1, \ldots, P_{n+1}$ for the projections onto the coordinate axes, we have an explicit expression for this second derivative in (V.22). Choose H to be of the form

$$H = \begin{pmatrix} 0 & 0 & \cdots & & \bar{\xi}_1 \\ 0 & 0 & \cdots & & \bar{\xi}_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \xi_1 & \xi_2 & \cdots & \xi_n & 0 \end{pmatrix},$$

where $\xi_1, \ldots, \xi_n$ are any complex numbers. Let x be the $(n+1)$-vector $(1, 1, \ldots, 1, 0)$. Then

$$\langle x, P_iHP_jHP_kx\rangle = \xi_k\bar{\xi}_i\delta_{j,n+1} \tag{V.23}$$

for $1 \le i,j,k \le n+1$, where $\delta_{j,n+1}$ is equal to 1 if $j = n+1$, and is equal to 0 otherwise. So, using the positivity of the matrix (V.22) and then (V.23), we have

$$\begin{aligned} 0 &\le \sum_{1\le i,j,k\le n+1} f^{[2]}(\lambda_i,\lambda_j,\lambda_k)\langle x, P_iHP_jHP_kx\rangle \\ &= \sum_{1\le i,k\le n} f^{[2]}(\lambda_i,\lambda_{n+1},\lambda_k)\xi_k\bar{\xi}_i. \end{aligned}$$

But,

$$\begin{aligned} f^{[2]}(\lambda_i,\lambda_{n+1},\lambda_k) &= \frac{f^{[1]}(\lambda_{n+1},\lambda_i) - f^{[1]}(\lambda_{n+1},\lambda_k)}{\lambda_i - \lambda_k} \\ &= g^{[1]}(\lambda_i,\lambda_k) \end{aligned}$$

(putting $\lambda_{n+1} = \mu$ in the definition of g). So we have

$$0 \leq \sum_{1 \leq i,k \leq n} g^{[1]}(\lambda_i, \lambda_k)\xi_k \, \bar{\xi}_i.$$

Since ξ_i are arbitrary complex numbers, this is equivalent to saying that the $n \times n$ matrix $[g^{[1]}(\lambda_i, \lambda_k)]$ is positive. ■

Corollary V.3.11 *If $f \in C^2(I)$, $f(0) = 0$, and f is operator convex, then the function $g(t) = \frac{f(t)}{t}$ is operator monotone.*

Proof. By the theorem above, the function $f^{[1]}(0, t)$ is operator monotone. But this is just the function $f(t)/t$ in this case. ■

Corollary V.3.12 *If f is operator monotone on I and $f(0) = 0$, then the function $g(t) = \frac{t+\lambda}{t} f(t)$ is operator monotone for $|\lambda| \leq 1$.*

Proof. First assume that $f \in C^2(I)$. By Lemma V.3.5, the function $g_\lambda(t) = (t + \lambda)f(t)$ is operator convex. By Corollary V.3.11, therefore, $g(t)$ is operator monotone.

If f is not in the class C^2, consider its regularisations f_ε. These are in C^2. Apply the special case of the above paragraph to the functions $f_\varepsilon(t) - f_\varepsilon(0)$, and then let $\varepsilon \to 0$. ■

Corollary V.3.13 *If f is operator monotone on I and $f(0) = 0$, then f is twice differentiable at 0.*

Proof. By Corollary V.3.12, the function $g(t) = (1 + \frac{1}{t})f(t)$ is operator monotone, and by Theorem V.3.6, it is continuously differentiable. So the function h defined as $h(t) = \frac{1}{t}f(t)$, $h(0) = f'(0)$ is continuously differentiable. This implies that f is twice differentiable at 0. ■

Exercise V.3.14 *Let f be a continuous operator monotone function on I. Then the function $F(t) = \int_0^t f(s)ds$ is operator convex.*

Exercise V.3.15 *Let $f \in C^1(I)$. Then f is operator convex if and only if for all Hermitian matrices A, B with eigenvalues in I we have*

$$f(A) - f(B) \geq f^{[1]}(B) \circ (A - B),$$

where $\circ$ denotes the Schur-product in a basis in which B is diagonal.

V.4 Loewner's Theorems

Consider all functions f on the interval $I = (-1, 1)$ that are operator monotone and satisfy the conditions

$$f(0) = 0, \qquad f'(0) = 1. \tag{V.24}$$

Let K be the collection of all such functions. Clearly, K is a convex set. We will show that this set is compact in the topology of pointwise convergence and will find its extreme points. This will enable us to write an integral representation for functions in K.

Lemma V.4.1 *If $f \in K$, then*

$$\begin{aligned} f(t) &\leq \frac{t}{1-t} \quad \text{for} \quad 0 \leq t < 1, \\ f(t) &\geq \frac{t}{1+t} \quad \text{for} -1 < t < 0, \\ |f''(0)| &\leq 2. \end{aligned}$$

Proof. Let $A = \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}$. By Theorem V.3.4, the matrix

$$f^{[1]}(A) = \begin{pmatrix} f'(t) & f(t)/t \\ f(t)/t & 1 \end{pmatrix}$$

is positive. Hence,

$$\frac{f(t)^2}{t^2} \leq f'(t). \tag{V.25}$$

Let $g_\pm(t) = (t \pm 1) f(t)$. By Lemma V.3.5, both functions $g_\pm$ are convex. Hence their derivatives are monotonically increasing functions. Since $g'_\pm(t) = f(t) + (t \pm 1) f'(t)$ and $g'_\pm(0) = \pm 1$, this implies that

$$f(t) + (t-1) f'(t) \geq -1 \quad \text{for} \quad t > 0 \tag{V.26}$$

and

$$f(t) + (t+1) f'(t) \leq 1 \quad \text{for} \quad t < 0. \tag{V.27}$$

From (V.25) and (V.26) we obtain

$$f(t) + 1 \geq \frac{(1-t) f(t)^2}{t^2} \quad \text{for} \quad t > 0. \tag{V.28}$$

Now suppose that for some $0 < t < 1$ we have $f(t) > \frac{t}{1-t}$. Then $f(t)^2 > \frac{t}{1-t} f(t)$. So, from (V.28), we get $f(t) + 1 > \frac{f(t)}{t}$. But this gives the inequality $f(t) < \frac{t}{1-t}$, which contradicts our assumption. This shows that $f(t) \leq \frac{t}{1-t}$ for $0 \leq t < 1$. The second inequality of the lemma is obtained by the same argument using (V.27) instead of (V.26).

We have seen in the proof of Corollary V.3.13 that

$$f'(0) + \frac{1}{2}f''(0) = \lim_{t\to 0} \frac{(1+t^{-1})f(t) - f'(0)}{t}.$$

Let $t \downarrow 0$ and use the first inequality of the lemma to conclude that this limit is smaller than 2. Let $t \uparrow 0$, and use the second inequality to conclude that it is bigger than 0. Together, these two imply that $|f''(0)| \le 2$. ■

Proposition V.4.2 *The set K is compact in the topology of pointwise convergence.*

Proof. Let $\{f_i\}$ be any net in K. By the lemma above, the set $\{f_i(t)\}$ is bounded for each t. So, by Tychonoff's Theorem, there exists a subnet $\{f_i\}$ that converges pointwise to a bounded function f. The limit function f is operator monotone, and $f(0) = 0$. If we show that $f'(0) = 1$, we would have shown that $f \in K$, and hence that K is compact.

By Corollary V.3.12, each of the functions $(1 + \frac{1}{t})f_i(t)$ is monotone on $(-1, 1)$. Since for all i, $\lim_{t\to 0}(1 + \frac{1}{t})f_i(t) = f_i'(0) = 1$, we see that $(1 + \frac{1}{t})f_i(t) \ge 1$ if $t \ge 0$ and is ≤ 1 if $t \le 0$. Hence, if $t > 0$, we have $(1 + \frac{1}{t})f(t) \ge 1$; and if $t < 0$, we have the opposite inequality. Since f is continuously differentiable, this shows that $f'(0) = 1$. ■

Proposition V.4.3 *All extreme points of the set K have the form*

$$f(t) = \frac{t}{1 - \alpha t}, \quad \textit{where} \quad \alpha = \frac{1}{2}f''(0).$$

Proof. Let $f \in K$. For each $\lambda, -1 < \lambda < 1$, let

$$g_\lambda(t) = (1 + \frac{\lambda}{t})f(t) - \lambda.$$

By Corollary V.3.12, g_λ is operator monotone. Note that $g_\lambda(0) = 0$, since $f(0) = 0$ and $f'(0) = 1$. Also, $g_\lambda'(0) = 1 + \frac{1}{2}\lambda f''(0)$. So the function h_λ defined as

$$h_\lambda(t) = \frac{1}{1 + \frac{1}{2}\lambda f''(0)}[(1 + \frac{\lambda}{t})f(t) - \lambda]$$

is in K. Since $|f''(0)| \le 2$, we see that $|\frac{1}{2}\lambda f''(0)| < 1$. We can write

$$f = \frac{1}{2}(1 + \frac{1}{2}\lambda f''(0))h_\lambda + \frac{1}{2}(1 - \frac{1}{2}\lambda f''(0))h_{-\lambda}.$$

So, if f is an extreme point of K, we must have $f = h_\lambda$. This says that

$$(1 + \frac{1}{2}\lambda f''(0))f(t) = (1 + \frac{\lambda}{t})f(t) - \lambda,$$

from which we can conclude that

$$f(t) = \frac{t}{1 - \frac{1}{2}f''(0)t}.$$

■

Theorem V.4.4 *For each f in K there exists a unique probability measure μ on $[-1,1]$ such that*

$$f(t) = \int_{-1}^{1} \frac{t}{1-\lambda t} d\mu(\lambda). \tag{V.29}$$

Proof. For $-1 \leq \lambda \leq 1$, consider the functions $h_\lambda(t) = \frac{t}{1-\lambda t}$. By Proposition V.4.3, the extreme points of K are included in the family $\{h_\lambda\}$. Since K is compact and convex, it must be the closed convex hull of its extreme points. (This is the Krein-Milman Theorem.) Finite convex combinations of elements of the family $\{h_\lambda : -1 \leq \lambda \leq 1\}$ can also be written as $\int h_\lambda d\nu(\lambda)$, where ν is a probability measure on $[-1,1]$ with finite support. Since f is in the closure of these combinations, there exists a net $\{\nu_i\}$ of finitely supported probability measures on $[-1,1]$ such that the net $f_i(t) = \int h_\lambda(t)d\nu_i(\lambda)$ converges to $f(t)$. Since the space of the probability measures is weak* compact, the net ν_i has an accumulation point μ. In other words, a subnet of $\int h_\lambda d\nu_i(\lambda)$ converges to $\int h_\lambda d\mu(\lambda)$. So $f(t) = \int h_\lambda(t) d\mu(\lambda) = \int \frac{t}{1-\lambda t} d\mu(\lambda)$.

Now suppose that there are two measures μ_1 and μ_2 for which the representation (V.29) is valid. Expand the integrand as a power series $\frac{t}{1-\lambda t} = \sum_{n=0}^{\infty} t^{n+1}\lambda^n$ convergent uniformly in $|\lambda| \leq 1$ for every fixed t with $|t| < 1$. This shows that

$$\sum_{n=0}^{\infty} t^{n+1} \int_{-1}^{1} \lambda^n d\mu_1(\lambda) = \sum_{n=0}^{\infty} t^{n+1} \int_{-1}^{1} \lambda^n d\mu_2(\lambda)$$

for all $|t| < 1$. The identity theorem for power series now shows that

$$\int_{-1}^{1} \lambda^n d\mu_1(\lambda) = \int_{-1}^{1} \lambda^n d\mu_2(\lambda), \quad n = 0, 1, 2, \ldots$$

But this is possible if and only if $\mu_1 = \mu_2$. ■

One consequence of the uniqueness of the measure μ in the representation (V.29) is that every function h_{λ_0} is an extreme point of K (because it can be represented as an integral like this with μ concentrated at λ_0).

The normalisations (V.24) were required to make the set K compact. They can now be removed. We have the following result.

Corollary V.4.5 *Let f be a nonconstant operator monotone function on $(-1,1)$. Then there exists a unique probability measure μ on $[-1,1]$ such that*

$$f(t) = f(0) + f'(0) \int_{-1}^{1} \frac{t}{1-\lambda t} d\mu(\lambda). \tag{V.30}$$

Proof. Since f is monotone and is not a constant, $f'(0) \neq 0$. Now note that the function $\frac{f(t)-f(0)}{f'(0)}$ is in K. ∎

It is clear from the representation (V.30) that every operator monotone function on $(-1,1)$ is infinitely differentiable. Hence, by the results of earlier sections, every operator convex function is also infinitely differentiable.

Theorem V.4.6 *Let f be a nonlinear operator convex function on $(-1,1)$. Then there exists a unique probability measure μ on $[-1,1]$ such that*

$$f(t) = f(0) + f'(0)t + \frac{1}{2} f''(0) \int_{-1}^{1} \frac{t^2}{1-\lambda t} d\mu(\lambda). \tag{V.31}$$

Proof. Assume, without loss of generality, that $f(0) = 0$ and $f'(0) = 0$. Let $g(t) = f(t)/t$. Then g is operator monotone by Corollary V.3.11, $g(0) = 0$, and $g'(0) = \frac{1}{2} f''(0)$. So g has a representation like (V.30), from which the representation (V.31) for f follows. ∎

We have noted that the integral representation (V.30) implies that every operator monotone function on $(-1,1)$ is infinitely differentiable. In fact, we can conclude more. This representation shows that f has an analytic continuation

$$f(z) = f(0) + f'(0) \int_{-1}^{1} \frac{z}{1-\lambda z} d\mu(\lambda) \tag{V.32}$$

defined everywhere on the complex plane except on $(-\infty,-1] \cup [1,\infty)$. Note that

$$\mathrm{Im}\ \frac{z}{1-\lambda z} = \frac{\mathrm{Im}\ z}{|1-\lambda z|^2}.$$

So f defined above maps the upper half-plane $H_+ = \{z : \mathrm{Im}\ z > 0\}$ into itself. It also maps the lower half-plane H_- into itself. Further, $f(\bar{z}) = \overline{f(z)}$. In other words, the function f on H_- is an analytic continuation of f on H_+ across the interval $(-1,1)$ obtained by reflection.

This is a very important observation, because there is a very rich theory of analytic functions in a half-plane that we can exploit now. Before doing so, let us now do away with the special interval $(-1,1)$. Note that a function f is operator monotone on an interval (a,b) if and only if the function

$f(\frac{(b-a)t}{2} + \frac{b+a}{2})$ is operator monotone on $(-1, 1)$. So, all results obtained for operator monotone functions on $(-1, 1)$ can be extended to functions on (a, b). We have proved the following.

Theorem V.4.7 *If f is an operator monotone function on (a, b), then f has an analytic continuation to the upper half-plane H_+ that maps H_+ into itself. It also has an analytic continuation to the lower-half plane H_-, obtained by reflection across (a, b).*

The converse of this is also true: if a real function f on (a, b) has an analytic continuation to H_+ mapping H_+ into itself, then f is operator monotone on (a, b). This is proved below.

Let P be the class of all complex analytic functions defined on H_+ with their ranges in the closed upper half-plane $\{z : \text{Im } z \geq 0\}$. This is called the class of **Pick functions**. Since every nonconstant analytic function is an open map, if f is a nonconstant Pick function, then the range of f is contained in H_+. It is obvious that P is a convex cone, and the composition of two nonconstant functions in P is again in P.

Exercise V.4.8 *(i) For $0 \leq r \leq 1$, the function $f(z) = z^r$ is in P.*

(ii) The function $f(z) = \log z$ is in P.

(iii) The function $f(z) = \tan z$ is in P.

(iv) The function $f(z) = -\frac{1}{z}$ is in P.

(v) If f is in P, then so is the function $\frac{-1}{f}$.

Given any open interval (a, b), let $P(a, b)$ be the class of Pick functions that admit an analytic continuation across (a, b) into the lower half-plane and the continuation is by reflection. In particular, such functions take only real values on (a, b), and if they are nonconstant, they assume real values only on (a, b). The set $P(a, b)$ is a convex cone.

Let $f \in P(a, b)$ and write $f(z) = u(z) + iv(z)$, where as usual $u(z)$ and $v(z)$ denote the real and imaginary parts of f. Since $v(x) = 0$ for $a < x < b$, we have $v(x+iy) - v(x) \geq 0$ if $y > 0$. This implies that the partial derivative $v_y(x) \geq 0$ and hence, by the Cauchy-Riemann equations, $u_x(x) \geq 0$. Thus, on the interval $(a, b), f(x) = u(x)$ is monotone. In fact, we will soon see that f is operator monotone on (a, b). This is a consequence of a theorem of Nevanlinna that gives an integral representation of Pick functions. We will give a proof of this now using some elementary results from Fourier analysis. The idea is to use the conformal equivalence between H_+ and the unit disk D to transfer the problem to D, and then study the real part u of f. This is a harmonic function on D, so we can use standard facts from Fourier analysis.

Theorem V.4.9 *Let u be a nonnegative harmonic function on the unit disk $D = \{z : |z| < 1\}$. Then there exists a finite measure m on $[0, 2\pi]$ such* *that*

$$u(re^{i\theta}) = \int_0^{2\pi} \frac{1-r^2}{1+r^2-2r\,\cos(\theta-t)}dm(t). \tag{V.33}$$

Conversely, any function of this form is positive and harmonic on the unit disk D.

Proof. Let u be any continuous real function defined on the closed unit disk that is harmonic in D. Then, by a well-known and elementary theorem in analysis,

$$\begin{aligned} u(re^{i\theta}) &= \frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{1+r^2-2r\,\cos(\theta-t)}u(e^{it})dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} P_r(\theta-t)u(e^{it})dt, \end{aligned} \tag{V.34}$$

where $P_r(\theta)$ is the Poisson kernel (defined by the above equation) for $0 \le r < 1$, $0 \le \theta \le 2\pi$. If u is nonnegative, put $dm(t) = \frac{1}{2\pi}u(e^{it})dt$. Then m is a positive measure on $[0, 2\pi]$. By the mean value property of harmonic functions, the total mass of this measure is

$$\frac{1}{2\pi}\int_0^{2\pi} u(e^{it})dt = u(0). \tag{V.35}$$

So we do have a representation of the form (V.33) under the additional hypothesis that u is continuous on the closed unit disk.

The general case is a consequence of this. Let u be positive and harmonic in D. Then, for $\varepsilon > 0$, the function $u_\varepsilon(z) = u(\frac{z}{1+\varepsilon})$ is positive and harmonic in the disk $|z| < 1+\varepsilon$. Therefore, it can be represented in the form (V.33) with a measure $m_\varepsilon(t)$ of finite total mass $u_\varepsilon(0) = u(0)$. As $\varepsilon \to 0$, u_ε converges to u uniformly on compact subsets of D. Since the measures m_ε all have the same mass, using the weak* compactness of the space of probability measures, we conclude that there exists a positive measure m such that

$$u(re^{i\theta}) = \lim_{\varepsilon\to 0} u_\varepsilon(re^{i\theta}) = \int_0^{2\pi} \frac{1-r^2}{1+r^2-2r\,\cos(\theta-t)}dm(t).$$

Conversely, since the Poisson kernel P_r is nonnegative any function represented by (V.33) is nonnegative. ∎

Theorem V.4.9 is often called the **Herglotz Theorem**. It says that *every nonnegative harmonic function on the unit disk is the Poisson integral of a positive measure.*

Recall that two harmonic functions u, v are called **harmonic conjugates** if the function $f(z) = u(z) + iv(z)$ is analytic. Every harmonic function u has a harmonic conjugate that is uniquely determined up to an additive constant.

Theorem V.4.10 *Let $f(z) = u(z) + iv(z)$ be analytic on the unit disk D. If $u(z) \geq 0$, then there exists a finite positive measure m on $[0, 2\pi]$ such that*

$$f(z) = \int_0^{2\pi} \frac{e^{it} + z}{e^{it} - z}\, dm(t) + iv(0). \tag{V.36}$$

Conversely, every function of this form is analytic on D and has a positive real part.

Proof. By Theorem V.4.9, the function u can be written as in (V.33). The Poisson kernel P_r, $0 \leq r < 1$, can be written as

$$P_r(\theta) = \frac{1 - r^2}{1 + r^2 - 2r \cos\theta} = \sum_{-\infty}^{\infty} r^{|n|} e^{in\theta} = \mathrm{Re}\, \frac{1 + re^{i\theta}}{1 - re^{i\theta}}.$$

Hence,

$$P_r(\theta - t) = \mathrm{Re}\, \frac{1 + re^{i(\theta - t)}}{1 - re^{i(\theta - t)}} = \mathrm{Re}\, \frac{e^{it} + re^{i\theta}}{e^{it} - re^{i\theta}},$$

and

$$u(z) = \mathrm{Re} \int_0^{2\pi} \frac{e^{it} + z}{e^{it} - z}\, dm(t).$$

So, $f(z)$ differs from this last integral only by an imaginary constant. Putting $z = 0$, one sees that this constant is $iv(0)$.

The converse statement is easy to prove. ■

Next, note that the disk D and the half-plane H_+ are conformally equivalent,, i.e., there exists an analytic isomorphism between these two spaces. For $z \in D$, let

$$\zeta(z) = \frac{1}{i}\, \frac{z + 1}{z - 1}. \tag{V.37}$$

Then $\zeta \in H_+$. The inverse of this map is given by

$$z(\zeta) = \frac{\zeta - i}{\zeta + i}. \tag{V.38}$$

Using these transformations, we can establish an equivalence between the class P and the class of analytic functions on D with positive real part. If f is a function in the latter class, let

$$\varphi(\zeta) = if(z(\zeta)). \tag{V.39}$$

Then $\varphi \in P$. The inverse of this transformation is

$$f(z) = -i\varphi(\zeta(z)). \tag{V.40}$$

Using these ideas we can prove the following theorem, called **Nevanlinna's Theorem.**

Theorem V.4.11 *A function φ is in the Pick class if and only if it has a representation*

$$\varphi(\zeta) = \alpha + \beta\zeta + \int_{-\infty}^{\infty} \frac{1+\lambda\zeta}{\lambda-\zeta} d\nu(\lambda), \tag{V.41}$$

where α is a real number, $\beta \geq 0$, and ν is a positive finite measure on the real line.

Proof. Let f be the function on D associated with φ via the transformation (V.40). By Theorem V.4.10, there exists a finite positive measure m on $[0, 2\pi]$ such that

$$f(z) = \int_0^{2\pi} \frac{e^{it}+z}{e^{it}-z} dm(t) - i\alpha.$$

If $f(z) = u(z) + iv(z)$, then $\alpha = -v(0)$, and the total mass of m is $u(0)$. If the measure m has a positive mass at the singleton $\{0\}$, let this mass be β. Then the expression above reduces to

$$f(z) = \int_{(0,2\pi)} \frac{e^{it}+z}{e^{it}-z}\, dm(t) + \beta\, \frac{1+z}{1-z} - i\alpha.$$

Using the transformations (V.38) and (V.39), we get from this

$$\varphi(\zeta) = \alpha + \beta\zeta + i \int_{(0,2\pi)} \frac{e^{it} + \frac{\zeta-i}{\zeta+i}}{e^{it} - \frac{\zeta-i}{\zeta+i}} dm(t).$$

The last term above is equal to

$$\int_{(0,2\pi)} \frac{\zeta\, \cos\frac{t}{2} - \sin\frac{t}{2}}{\zeta\, \sin\frac{t}{2} + \cos\frac{t}{2}} dm(t).$$

Now, introduce a change of variables $\lambda = -\cot\frac{t}{2}$. This maps $(0, 2\pi)$ onto $(-\infty, \infty)$. The measure m is transformed by the above map to a finite measure ν on $(-\infty, \infty)$ and the above integral is transformed to

$$\int_{-\infty}^{\infty} \frac{1+\lambda\zeta}{\lambda-\zeta}\, d\nu(\lambda).$$

This shows that φ can be represented in the form (V.41).

It is easy to see that every function of this form is a Pick function. ∎

There is another form in which it is convenient to represent Pick functions. Note that

$$\frac{1+\lambda\zeta}{\lambda-\zeta} = \left(\frac{1}{\lambda-\zeta} - \frac{\lambda}{\lambda^2+1}\right)(\lambda^2+1).$$

So, if we write $d\mu(\lambda) = (\lambda^2+1)d\nu(\lambda)$, then we obtain from (V.41) the representation

$$\varphi(\zeta) = \alpha + \beta\zeta + \int_{-\infty}^{\infty} \left[\frac{1}{\lambda-\zeta} - \frac{\lambda}{\lambda^2+1}\right] d\mu(\lambda), \tag{V.42}$$

where μ is a positive Borel measure on $\mathbb{R}$, for which $\int \frac{1}{\lambda^2+1} d\mu(\lambda)$ is finite. (A Borel measure on $\mathbb{R}$ is a measure defined on Borel sets that puts finite mass on bounded sets.)

Now we turn to the question of uniqueness of the above representations. It is easy to see from (V.41) that

$$\alpha = \operatorname{Re} \varphi(i). \tag{V.43}$$

Therefore, α is uniquely determined by φ. Now let η be any positive real number. From (V.41) we see that

$$\frac{\varphi(i\eta)}{i\eta} = \frac{\alpha}{i\eta} + \beta + \int_{-\infty}^{\infty} \frac{1+\lambda^2+i\lambda(\eta-\eta^{-1})}{\lambda^2+\eta^2} d\nu(\lambda).$$

As $\eta \to \infty$, the integrand converges to 0 for each λ. The real and imaginary parts of the integrand are uniformly bounded by 1 when $\eta > 1$. So by the Lebesgue Dominated Convergence Theorem, the integral converges to 0 as $\eta \to \infty$. Thus,

$$\beta = \lim_{\eta\to\infty} \varphi(i\eta)/i\eta, \tag{V.44}$$

and thus β is uniquely determined by φ.

Now we will prove that the measure $d\mu$ in (V.42), is uniquely determined by φ. Denote by μ the unique right continuous monotonically increasing function on $\mathbb{R}$ satisfying $\mu(0) = 0$ and $\mu((a,b]) = \mu(b) - \mu(a)$ for every interval $(a,b]$. (This is called the **distribution function** associated with $d\mu$.) We will prove the following result, called the **Stieltjes inversion formula**, from which it follows that μ is unique.

Theorem V.4.12 *If the Pick function φ is represented by (V.42), then for any a, b that are points of continuity of the distribution function μ we have*

$$\mu(b) - \mu(a) = \lim_{\eta\to 0} \frac{1}{\pi} \int_a^b \operatorname{Im} \varphi(x+i\eta)dx. \tag{V.45}$$

Proof. From (V.42) we see that

$$\frac{1}{\pi}\int_a^b \operatorname{Im} \varphi(x+i\eta)dx = \frac{1}{\pi}\int_a^b \left[\beta\eta + \int_{-\infty}^{\infty} \frac{\eta}{(\lambda - x)^2+\eta^2} d\mu(\lambda)\right] dx$$
$$= \frac{1}{\pi}\left[\beta\eta(b-a) + \int_{-\infty}^{\infty}\int_a^b \frac{\eta dx}{(x-\lambda)^2+\eta^2} d\mu(\lambda)\right],$$

the interchange of integrals being permissible by Fubini's Theorem. As $\eta \to 0$, the first term in the square brackets above goes to 0. The inner integral can be calculated by the change of variables $u = \frac{x-\lambda}{\eta}$. This gives

$$\int_a^b \frac{\eta dx}{(x-\lambda)^2+\eta^2} = \int_{\frac{a-\lambda}{\eta}}^{\frac{b-\lambda}{\eta}} \frac{du}{u^2+1}$$
$$= \arctan\left(\frac{b-\lambda}{\eta}\right) - \arctan\left(\frac{a-\lambda}{\eta}\right).$$

So to prove (V.45), we have to show that

$$\mu(b)-\mu(a) = \lim_{\eta\to 0}\frac{1}{\pi}\int_{-\infty}^{\infty}\left[\arctan\left(\frac{b-\lambda}{\eta}\right) - \arctan\left(\frac{a-\lambda}{\eta}\right)\right] d\mu(\lambda).$$

We will use the following properties of the function arctan. This is a monotonically increasing odd function on $(-\infty,\infty)$ whose range is $(-\frac{\pi}{2}, \frac{\pi}{2})$. So,

$$0 \le \arctan\left(\frac{b-\lambda}{\eta}\right) - \arctan\left(\frac{a-\lambda}{\eta}\right) \le \pi.$$

If $(b-\lambda)$ and $(a-\lambda)$ have the same sign, then by the addition law for arctan we have,

$$\arctan\left(\frac{b-\lambda}{\eta}\right) - \arctan\left(\frac{a-\lambda}{\eta}\right) = \arctan\frac{\eta(b-a)}{\eta^2+(b-\lambda)(a-\lambda)}.$$

If x is positive, then

$$\arctan x = \int_0^x \frac{dt}{1+t^2} \le \int_0^x dt = x.$$

Now, let ε be any given positive number. Since a and b are points of continuity of μ, we can choose δ such that

$$\mu(a+\delta)-\mu(a-\delta) \le \varepsilon/5,$$
$$\mu(b+\delta)-\mu(b-\delta) \le \varepsilon/5.$$

We then have,

$$
\begin{aligned}
& \left| \mu(b) - \mu(a) - \frac{1}{\pi} \int_{-\infty}^{\infty} \left[\arctan\left(\frac{b-\lambda}{\eta}\right) - \arctan\left(\frac{a-\lambda}{\eta}\right)\right] d\mu(\lambda) \right| \\
\leq\ & \frac{1}{\pi} \int_{b}^{\infty} \left[\arctan\left(\frac{b-\lambda}{\eta}\right) - \arctan\left(\frac{a-\lambda}{\eta}\right)\right] d\mu(\lambda) \\
& + \frac{1}{\pi} \int_{a}^{b} \left[\pi - \arctan\left(\frac{b-\lambda}{\eta}\right) + \arctan\left(\frac{a-\lambda}{\eta}\right)\right] d\mu(\lambda) \\
& + \frac{1}{\pi} \int_{-\infty}^{a} \left[\arctan\left(\frac{b-\lambda}{\eta}\right) - \arctan\left(\frac{a-\lambda}{\eta}\right)\right] d\mu(\lambda) \\
\leq\ & \frac{2\varepsilon}{5} + \frac{1}{\pi} \int_{b+\delta}^{\infty} \arctan\left(\frac{\eta(b-a)}{\eta^2 + (b-\lambda)(a-\lambda)}\right) d\mu(\lambda) \\
& + \frac{1}{\pi} \int_{a+\delta}^{b-\delta} \left[\pi - \arctan\left(\frac{b-\lambda}{\eta}\right) + \arctan\left(\frac{a-\lambda}{\eta}\right)\right] d\mu(\lambda) \\
& + \frac{1}{\pi} \int_{-\infty}^{a-\delta} \arctan\left(\frac{\eta(b-a)}{\eta^2 + (b-\lambda)(a-\lambda)}\right) d\mu(\lambda).
\end{aligned}
$$

Note that in the two integrals with infinite limits, the arguments of arctan are positive. In the middle integral the variable λ runs between $a+\delta$ and $b-\delta$. For such λ, $\frac{b-\lambda}{\eta} \geq \frac{\delta}{\eta}$ and $\frac{a-\lambda}{\eta} \leq -\frac{\delta}{\eta}$. So the right-hand side of the above inequality is dominated by

$$
\begin{aligned}
\frac{2\varepsilon}{5} & + \frac{\eta}{\pi} \int_{b+\delta}^{\infty} \frac{b-a}{\eta^2 + (b-\lambda)(a-\lambda)} d\mu(\lambda) \\
& + \frac{\eta}{\pi} \int_{-\infty}^{a-\delta} \frac{b-a}{\eta^2 + (b-\lambda)(a-\lambda)} d\mu(\lambda) \\
& + \frac{1}{\pi} \int_{a+\delta}^{b-\delta} \left[\pi - 2 \arctan \frac{\delta}{\eta}\right] d\mu(\lambda).
\end{aligned}
$$

The first two integrals are finite (because of the properties of $d\mu$). The third one is dominated by $2(\frac{\pi}{2} - \arctan \frac{\delta}{\eta})[\mu(b) - \mu(a)]$. So we can choose η small enough to make each of the last three terms smaller than $\varepsilon/5$. This proves the theorem. ■

We have shown above that all the terms occurring in the representation (V.42) are uniquely determined by the relations (V.43), (V.44), and (V.45).

Exercise V.4.13 *We have proved the relations (V.33), (V.36), (V.41) and (V.42) in that order. Show that all these are, in fact, equivalent. Hence, each of these representations is unique.*

Proposition V.4.14 *A Pick function φ is in the class $P(a,b)$ if and only if the measure μ associated with it in the representation (V.42) has zero mass on (a,b).*

Proof. Let $\varphi(x+i\eta) = u(x+i\eta) + iv(x+i\eta)$, where u, v are the real and imaginary parts of φ. If φ can be continued across (a,b), then as $\eta \downarrow 0$, on any closed subinterval $[c,d]$ of $(a,b), v(x+i\eta)$ converges uniformly to a bounded continuous function $v(x)$ on $[c,d]$. Hence,

$$\mu(d) - \mu(c) = \frac{1}{\pi}\int_c^d v(x)dx,$$

i.e., $d\mu(x) = \frac{1}{\pi}v(x)dx$. If the analytic continuation to the lower half-plane is by reflection across (a,b), then v is identically zero on $[c,d]$ and hence so is μ.

Conversely, if μ has no mass on (a,b), then for ζ in (a,b) the integral in (V.42) is convergent, and is real valued. This shows that the function φ can be continued from H_+ to H_- across (a,b) by reflection. ■

The reader should note that the above proposition shows that the converse of Theorem V.4.7 is also true.

It should be pointed out that the formula (V.42) defines two analytic functions, one on H_+ and the other on H_-. If these are denoted by φ and ψ, then $\varphi(\zeta) = \overline{\psi(\bar{\zeta})}$. So φ and ψ are reflections of each other. But they need not be analytic continuations of each other. For this to be the case, the measure μ should be zero on an interval (a,b) across which the function can be continued analytically.

Exercise V.4.15 *If a function f is operator monotone on the whole real line, then f must be of the form $f(t) = \alpha + \beta t$, $\alpha \in \mathbb{R}$, $\beta \geq 0$.*

Let us now look at a few simple examples.

Example V.4.16 *The function $\varphi(\zeta) = -\frac{1}{\zeta}$ is a Pick function. For this function, we see from (V.43) and (V.44) that $\alpha = \beta = 0$. Since φ is analytic everywhere in the plane except at 0, Proposition V.4.14 tells us that the measure μ is concentrated at the single point 0.*

Example V.4.17 *Let $\varphi(\zeta) = \zeta^{1/2}$ be the principal branch of the square root function. This is a Pick function. From (V.43) we see that*

$$\alpha = \text{Re } \varphi(i) = \text{Re } e^{i\pi/4} = \frac{1}{\sqrt{2}}.$$

From (V.44) we see that $\beta = 0$. If $\zeta = \lambda + i\eta$ is any complex number, then

$$\zeta^{1/2} = \left(\frac{|\zeta| + \lambda}{2}\right)^{1/2} + i \operatorname{sgn} \eta \left(\frac{|\zeta| - \lambda}{2}\right)^{1/2},$$

where sgn η *is the sign of η, defined to be 1 if $\eta \geq 0$ and -1 if $\eta < 0$. So if $\eta \geq 0$, we have* Im $\varphi(\zeta) = \left(\frac{|\zeta|-\lambda}{2}\right)^{1/2}$. *As $\eta \downarrow 0$, $|\zeta|$ comes closer to $|\lambda|$. So,* Im $\varphi(\lambda + i\eta)$ *converges to 0 if $\lambda > 0$ and to $|\lambda|^{1/2}$ if $\lambda < 0$. Since φ is positive on the right half-axis, the measure μ has no mass at 0. The measure can now be determined from (V.45). We have, then*

$$\zeta^{1/2} = \frac{1}{\sqrt{2}} + \int_{-\infty}^{0} \left(\frac{1}{\lambda - \zeta} - \frac{\lambda}{\lambda^2 + 1}\right) \frac{|\lambda|^{1/2}}{\pi} d\lambda. \tag{V.46}$$

Example V.4.18 *Let $\varphi(\zeta) = \text{Log } \zeta$, where* Log *is the principal branch of the logarithm, defined everywhere except on $(-\infty, 0]$ by the formula* Log $\zeta = \ln|\zeta| + i$ Arg ζ. *The function* Arg ζ *is the principal branch of the argument, taking values in $(-\pi, \pi]$. We then have*

$$\begin{aligned} \alpha &= \text{Re}(\text{Log}\, i) = 0 \\ \beta &= \lim_{\eta \to \infty} \frac{\text{Log}(i\eta)}{i\eta} = 0. \end{aligned}$$

As $\eta \downarrow 0$, Im (Log$(\lambda + i\eta)$) *converges to π if $\lambda < 0$ and to 0 if $\lambda > 0$. So from (V.45) we see that, the measure μ is just the restriction of the Lebesgue measure to $(-\infty, 0]$. Thus,*

$$\text{Log } \zeta = \int_{-\infty}^{0} \left(\frac{1}{\lambda - \zeta} - \frac{\lambda}{\lambda^2 + 1}\right) d\lambda. \tag{V.47}$$

Exercise V.4.19 *For $0 < r < 1$, let ζ^r denote the principal branch of the function $\varphi(\zeta) = \zeta^r$. Show that*

$$\zeta^r = \cos \frac{r\pi}{2} + \frac{\sin r\pi}{\pi} \int_{-\infty}^{0} \left(\frac{1}{\lambda - \zeta} - \frac{\lambda}{\lambda^2 + 1}\right) |\lambda|^r d\lambda. \tag{V.48}$$

This includes (V.46) as a special case.

Let now f be any operator monotone function on $(0,\infty)$. We have seen above that f must have the form

$$f(t)=\alpha+\beta t+\int\limits_{-\infty}^{0}\left(\frac{1}{\lambda-t}-\frac{\lambda}{\lambda^2+1}\right)d\mu(\lambda).$$

By a change of variables we can write this as

$$f(t)=\alpha+\beta t+\int\limits_{0}^{\infty}\left(\frac{\lambda}{\lambda^2+1}-\frac{1}{\lambda+t}\right)d\mu(\lambda), \tag{V.49}$$

where $\alpha\in\mathbb{R}$, $\beta\geq 0$ and μ is a positive measure on $(0,\infty)$ such that

$$\int\limits_{0}^{\infty}\frac{1}{\lambda^2+1}d\mu(\lambda)<\infty. \tag{V.50}$$

Suppose f is such that

$$f(0):=\lim_{t\to 0}f(t)>-\infty. \tag{V.51}$$

Then, it follows from (V.49) that μ must also satisfy the condition

$$\int\limits_{0}^{1}\frac{1}{\lambda}d\mu(\lambda)<\infty. \tag{V.52}$$

We have from (V.49)

$$\begin{aligned} f(t)-f(0) &= \beta t+\int\limits_{0}^{\infty}\left(\frac{1}{\lambda}-\frac{1}{\lambda+t}\right)d\mu(\lambda)\\ &= \beta t+\int_0^{\infty}\frac{t}{(\lambda+t)\lambda}\,d\mu(\lambda). \end{aligned}$$

Hence, we can write f in the form

$$f(t)=\gamma+\beta t+\int_0^{\infty}\frac{\lambda t}{\lambda+t}dw(\lambda), \tag{V.53}$$

where $\gamma=f(0)$ and $dw(\lambda)=\frac{1}{\lambda^2}d\mu(\lambda)$. From (V.50) and (V.52), we see that the measure w satisfies the conditions

$$\int\limits_{0}^{\infty}\frac{\lambda^2}{\lambda^2+1}dw(\lambda)<\infty \text{ and } \int\limits_{0}^{1}\lambda dw(\lambda)<\infty. \tag{V.54}$$

These two conditions can, equivalently, be expressed as a single condition

$$\int_0^\infty \frac{\lambda}{1+\lambda} dw(\lambda) < \infty. \tag{V.55}$$

We have thus shown that an operator monotone function on $(0, \infty)$ satisfying the condition (V.51) has a canonical representation (V.53), where $\gamma \in \mathbb{R}, \beta \geq 0$ and w is a positive measure satisfying (V.55).

The representation (V.53) is often useful for studying operator monotone functions on the positive half-line $[0, \infty)$.

Suppose that we are given a function f as in (V.53). If μ satisfies the conditions (V.54) then

$$\int_0^\infty \left(\frac{\lambda}{\lambda^2+1} - \frac{1}{\lambda}\right) \lambda^2 dw(\lambda) > -\infty,$$

and we can write

$$f(t) = \{\gamma - \int_0^\infty \left(\frac{\lambda}{\lambda^2+1} - \frac{1}{\lambda}\right) \lambda^2 dw(\lambda)\} + \beta t + \int_0^\infty \left(\frac{\lambda}{\lambda^2+1} - \frac{1}{\lambda+t}\right) \lambda^2 dw(\lambda).$$

So, if we put the number in braces above equal to α and $d\mu(\lambda) = \lambda^2 dw(\lambda)$, then we have a representation of f in the form (V.49).

Exercise V.4.20 *Use the considerations in the preceding paragraphs to show that, for* $0 < r \leq 1$ *and* $t > 0$, *we have*

$$t^r = \frac{\sin r\pi}{\pi} \int_0^\infty \frac{\lambda t}{\lambda + t} \lambda^{r-2} d\lambda. \tag{V.56}$$

(See Exercise V.1.10 also.)

Exercise V.4.21 *For* $t > 0$, *show that*

$$\log(1+t) = \int_1^\infty \frac{\lambda t}{\lambda + t} \lambda^{-2} d\lambda. \tag{V.57}$$

Appendix 1. Differentiability of Convex Functions

Let f be a real valued convex function defined on an interval I. Then f has some smoothness properties, which are listed below.

The function f is Lipschitz on any closed interval $[a, b]$ contained in I^0, the interior of I. So f is continuous on I^0.

At every point x in I^0, the right and left derivatives of f exist. These are defined, respectively, as

$$f'_+(x) := \lim_{y \downarrow x} \frac{f(y) - f(x)}{y - x},$$

$$f'_-(x) := \lim_{y \uparrow x} \frac{f(y) - f(x)}{y - x}.$$

Both these functions are monotonically increasing on I^0. Further,

$$\begin{aligned} \lim_{x \downarrow w} f'_\pm(x) &= f'_+(w), \\ \lim_{x \uparrow w} f'_\pm(x) &= f'_-(w). \end{aligned}$$

The function f is differentiable except on a countable set E in I^0, i.e., at every point x in $I^0 \backslash E$ the left and right derivatives of f are equal. Further, the derivative f' is continuous on $I^0 \backslash E$.

If a sequence of convex functions converges at every point of I, then the limit function is convex. The convergence is uniform on any closed interval $[a, b]$ contained in I^0.

Appendix 2. Regularisation of Functions

The convolution of two functions leads to a new function that inherits the stronger of the smoothness properties of the two original functions. This is the idea behind "regularisation" of functions.

Let φ be a real function of class C^∞ with the following properties: $\varphi \geq 0, \varphi$ is even, the support supp $\varphi = [-1, 1]$, and $\int \varphi = 1$. For each $\varepsilon > 0$, let $\varphi_\varepsilon(x) = \frac{1}{\varepsilon}\varphi(\frac{x}{\varepsilon})$. Then supp $\varphi_\varepsilon = [-\varepsilon, \varepsilon]$ and φ_ε has all the other properties of φ listed above. The functions φ_ε are called **mollifiers** or **smooth approximate identities**.

If f is a locally integrable function, we define its **regularisation of order** ε as the function

$$\begin{aligned} f_\varepsilon(x) = (f * \varphi_\varepsilon)(x) &:= \int f(x - y)\varphi_\varepsilon(y)dy \\ &= \int f(x - \varepsilon t)\varphi(t)dt. \end{aligned}$$

The family f_ε has the following properties.

1. Each f_ε is a C^∞ function.

2. If the support of f is contained in a compact set K, then the support of f_ε is contained in an ε-neighbourhood of K.

3. If f is continuous at x_0, then $\lim_{\varepsilon \downarrow 0} f_\varepsilon(x_0) = f(x_0)$.

4. If f has a discontinuity of the first kind at x_0, then $\lim_{\varepsilon \downarrow 0} f_\varepsilon(x_0) = 1/2\,[f(x_0+) + f(x_0-)]$. (A point x_0 is a point of discontinuity of the first kind if the left and right limits of f at x_0 exist; these limits are denoted as $f(x_0-)$ and $f(x_0+)$, respectively.)

5. If f is continuous, then $f_\varepsilon(x)$ converges to $f(x)$ as $\varepsilon \to 0$. The convergence is uniform on every compact set.

6. If f is differentiable, then, for every $\varepsilon > 0, (f_\varepsilon)' = (f')_\varepsilon$.

7. If f is monotone, then, as $\varepsilon \to 0, f_\varepsilon'(x)$ converges to $f'(x)$ at all points x where $f'(x)$ exists. (Recall that a monotone function can have discontinuities of the first kind only and is differentiable almost everywhere.)

V.5 Problems

Problem V.5.1. Show that the function $f(t) = \exp t$ is neither operator monotone nor operator convex on any interval.

Problem V.5.2. Let $f(t) = \frac{at+b}{ct+d}$, where a, b, c, d are real numbers such that $ad - bc > 0$. Show that f is operator monotone on every interval that does not contain the point $\frac{-d}{c}$.

Problem V.5.3. Show that the derivative of an operator convex function need not be operator monotone.

Problem V.5.4. Show that for $r < -1$, the function $f(t) = t^r$ on $(0, \infty)$ is not operator convex. (Hint: The function $f^{[1]}(1, t)$ cannot be continued analytically to a Pick function.) Together with the assertion in Exercise V.2.11, this shows that on the half-line $(0, \infty)$ the function $f(t) = t^r$ is operator convex if $-1 \le r \le 0$ or if $1 \le r \le 2$; and it is not operator convex for any other real r.

Problem V.5.5. A function g on $[0, \infty)$ is operator convex if and only if it is of the form

$$g(t) = \alpha + \beta t + \gamma t^2 + \int_0^\infty \frac{\lambda t^2}{\lambda + t}\, d\mu(\lambda),$$

where α, β are real numbers, $\gamma \ge 0$, and μ is a positive finite measure.

Problem V.5.6. Let f be an operator monotone function on $(0, \infty)$. Then $(-1)^{n-1} f^{(n)}(t) \geq 0$ for $n = 1, 2, \ldots$. [A function g on $(0, \infty)$ is said to be **completely monotone** if for all $n \geq 0$, $(-1)^n g^{(n)}(t) \geq 0$. There is a theorem of S.N. Bernstein that says that a function g is completely monotone if and only if there exists a positive measure μ such that $g(t) = \int_0^\infty e^{-\lambda t} d\mu(\lambda)$.] The result of this problem says that the derivative of an operator monotone function on $(0, \infty)$ is completely monotone. Thus, f has a Taylor expansion $f(t) = \sum_{n=0}^{\infty} a_n (t-1)^n$, in which the coefficients a_n are positive for all odd n and negative for all even n.

Problem V.5.7. Let f be a function mapping $(0, \infty)$ into itself. Let $g(t) = [f(t^{-1})]^{-1}$. Show that if f is operator monotone, then g is also operator monotone. If f is operator convex and $f(0) = 0$, then g is operator convex.

Problem V.5.8. Show that the function $f(\zeta) = -\cot \zeta$ is a Pick function. Show that in its canonical representation (V.42), $\alpha = \beta = 0$ and the measure μ is atomic with mass 1 at the points $n\pi$ for every integer n. Thus, we have the familiar series expansion

$$-\cot \zeta = \sum_{n=-\infty}^{\infty} \left[\frac{1}{n\pi - \zeta} - \frac{n\pi}{n^2\pi^2 + 1} \right].$$

Problem V.5.9. The aim of this problem is to show that if a Pick function φ satisfies the growth restriction

$$\sup_{\eta \to \infty} |\eta \varphi(i\eta)| < \infty, \tag{V.58}$$

then its representation (V.42) takes the simple form

$$\varphi(\zeta) = \int_{-\infty}^{\infty} \frac{1}{\lambda - \zeta} d\mu(\lambda), \tag{V.59}$$

where μ is a finite measure.

To see this, start with the representation (V.41). The condition (V.58) implies the existence of a constant M that bounds, for all $\eta > 0$, the quantity $\eta\varphi(i\eta)$, and hence also its real and imaginary parts. This gives two inequalities:

$$|\alpha\eta + \int_{-\infty}^{\infty} \frac{\eta(1-\eta^2)\lambda}{\lambda^2 + \eta^2} d\nu(\lambda)| \leq M,$$

$$|\beta\eta^2 + \eta^2 \int_{-\infty}^{\infty} \frac{1+\lambda^2}{\lambda^2 + \eta^2} d\nu(\lambda)| \leq M.$$

From the first, conclude that

$$\alpha = \lim_{\eta \to \infty} \int_{-\infty}^{\infty} \frac{(\eta^2 - 1)\lambda}{\lambda^2 + \eta^2} d\nu(\lambda) = \int_{-\infty}^{\infty} \lambda \, d\nu(\lambda).$$

From the second, conclude that $\beta = 0$ and

$$\int_{-\infty}^{\infty} \frac{\eta^2}{\lambda^2 + \eta^2} (1 + \lambda^2) d\nu(\lambda) \leq M.$$

Taking limits as $\eta \to \infty$, this gives

$$\int_{-\infty}^{\infty} (1 + \lambda^2) d\nu(\lambda) = \int_{-\infty}^{\infty} d\mu(\lambda) \leq M.$$

Thus, μ is a finite measure. From (V.41), we get

$$\varphi(\zeta) = \int_{-\infty}^{\infty} \lambda \, d\nu(\lambda) + \int_{-\infty}^{\infty} \frac{1 + \lambda\zeta}{\lambda - \zeta} d\nu(\lambda).$$

This is the same as (V.59).

Conversely, observe that if φ has a representation like (V.59), then it must satisfy the condition (V.58).

Problem V.5.10. Let f be a function on $(0, \infty)$ such that

$$f(t) = \alpha + \beta t - \int_{0}^{\infty} \frac{1}{\lambda + t} d\mu(\lambda),$$

where $\alpha \in \mathbb{R}$, $\beta \geq 0$ and μ is a positive measure such that $\int \frac{1}{\lambda} d\mu(\lambda) < \infty$. Then f is operator monotone. Find operator monotone functions that can not be expressed in this form.

V.6 Notes and References

Operator monotone functions were first studied in detail by K.Löwner (C. Loewner) in a seminal paper *Über monotone Matrixfunktionen*, Math. Z., 38 (1934) 177-216. In this paper, he established the connection between operator monotonicity, the positivity of the matrix of divided differences (Theorem V.3.4), and Pick functions. He also noted that the functions $f(t) = t^r$, $0 \leq r \leq 1$, and $f(t) = \log t$ are operator monotone on $(0, \infty)$.

Operator convex functions were studied, soon afterwards, by F. Kraus, *Über konvexe Matrixfunktionen*, Math. Z., 41(1936) 18-42.

In another well-known paper, *Beiträge zur Störungstheorie der Spectralzerlegung*, Math. Ann., 123 (1951) 415-438, E. Heinz used the theory of operator monotone functions to study several problems of perturbation theory for bounded and unbounded operators. The integral representation (V.41) in this context seems to have been first used by him. The operator monotonicity of the map $A \to A^r$ for $0 \leq r \leq 1$ is sometimes called the "Loewner-Heinz inequality", although it was discovered by Loewner.

J. Bendat and S. Sherman, *Monotone and convex operator functions*, Trans. Amer. Math. Soc., 79(1955) 58-71, provided a new perspective on the theorems of Loewner and Kraus. Theorem V.4.4 was first proved by them, and used to give a proof of Loewner's theorems.

A completely different and extremely elegant proof of Loewner's Theorem, based on the spectral theorem for (unbounded) selfadjoint operators was given by A. Korányi, *On a theorem of Löwner and its connections with resolvents of selfadjoint transformations*, Acta Sci. Math. Szeged, 17 (1956) 63-70.

Formulas like (V.13) and (V.22) were proved by Ju. L. Daleckii and S.G. Krein, *Formulas of differentiation according to a parameter of functions of Hermitian operators*, Dokl. Akad. Nauk SSSR, 76 (1951) 13-16. It was pointed out by M.G. Krein that the resulting Taylor formula could be used to derive conditions for operator monotonicity.

A concise presentation of the main ideas of operator monotonicity and convexity, including the approach of Daleckii and Krein, was given by C. Davis, *Notions generalizing convexity for functions defined on spaces of matrices*, in *Convexity: Proceedings of Symposia in Pure Mathematics*, American Mathematical Society, 1963, pp. 187-201. This paper also discussed other notions of convexity, examples and counterexamples, and was very influential.

A full book devoted to this topic is *Monotone Matrix Functions and Analytic Continuation*, by W.F. Donoghue, Springer-Verlag, 1974. Several ramifications of the theory and its connections with classical real and complex analysis are discussed here.

In a set of mimeographed lecture notes, *Topics on Operator Inequalities*, Hokkaido University, Sapporo, 1978, T. Ando provided a very concise modern survey of operator monotone and operator convex functions. Anyone who wishes to learn the Korányi method mentioned above should certainly read these notes.

A short proof of Löwner's Theorem appeared in G. Sparr, *A new proof of Löwner's theorem on monotone matrix functions*, Math. Scand., 47 (1980) 266-274.

In another brief and attractive paper, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann., 258 (1982) 229-241, F. Hansen and G.K. Pedersen provided another approach.

Much of Sections 2, 3, and 4 are based on this paper of Hansen and Pedersen. For the latter parts of Section 4 we have followed Donoghue. We have also borrowed freely from Ando and from Davis. Our proof of Theorem V.1.9 is taken from M. Fujii and T. Furuta, *Löwner-Heinz, Cordes and Heinz-Kato inequalities*, Math. Japonica, 38 (1993) 73-78. Characterisations of operator convexity like the one in Exercise V.3.15 may be found in J.S. Aujla and H.L. Vasudeva, *Convex and monotone operator functions*, Ann. Polonici Math., 62 (1995) 1-11.

Operator monotone and operator convex functions are studied in R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Chapter 6. See also the interesting paper R.A. Horn, *The Hadamard product*, in C.R. Johnson, ed. *Matrix Theory and Applications*, American Mathematical Society, 1990.

A short, but interesting, section of the Marshall-Olkin book (cited in Chapter 2) is devoted to this topic. Especially interesting are some of the examples and connections with statistics that they give.

Among several applications of these ideas, there are two that we should mention here. Operator monotone functions arise often in the study of electrical networks. See, e.g., W.N. Anderson and G.E. Trapp, *A class of monotone operator functions related to electrical network theory*, Linear Algebra Appl., 15(1975) 53-67. They also occur in problems related to elementary particles. See, e.g., E. Wigner and J. von Neumann, *Significance of Löwner's theorem in the quantum theory of collisions*, Ann. of Math., 59 (1954) 418-433.

There are important notions of means of operators that are useful in the analysis of electrical networks and in quantum physics. An axiomatic approach to the study of these means was introduced by F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann., 249 (1980) 205-224. They establish a one-to-one correspondence between the class of operator monotone functions f on $[0, \infty)$ with $f(1) = 1$ and the class of operator means.