THE SPECTRAL SHIFT OPERATOR

FRITZ GESZTESY, KONSTANTIN A. MAKAROV,
AND SERGUEI N. NABOKO

Abstract. We introduce the concept of a spectral shift operator and use it to derive Krein's spectral shift function for pairs of self-adjoint operators. Our principal tools are operator-valued Herglotz functions and their logarithms. Applications to Krein's trace formula and to the Birman-Solomyak spectral averaging formula are discussed.

1. Introduction

Krein's spectral shift function [51], [52], [54] has received enormous attention in the past due to its widespread applications in a variety of fields including scattering theory ([5], Ch. 19, [9], [10], [12], [13], [44], [46], [52], [84], [88], Ch. 8), relative index theory ([3], [14], [15], [38], [59], Ch. X), spectral averaging ([2], [6], [11], [27], [42], [43], [66], [79]–[81]) and its application to localization properties of random Hamiltonians ([17], [20], [21], [22], Ch. VIII, [24], [25], [47]–[49], [64], Ch. V, [78], [79]), eigenvalue counting functions and spectral asymptotics ([10], [23], [29], [30], [67]–[69]), semi-classical approximation ([70]–[72], [73]) and trace formulas for one-dimensional Schrödinger and Jacobi operators ([31]–[33], [35], [39]). Detailed reviews on the spectral shift function and its applications up to 1993 were published by Birman and Yafaev [12], [13]. These two papers contain extensive bibliographies and the interested reader will find many additional references therein. Historically, the concept of a spectral shift function was introduced by I. M. Lifshits [56], [57].

Our main contribution to this circle of ideas is the introduction of a spectral shift operator $\Xi(\lambda, H_0, H)$ for a.e. $\lambda \in \mathbb{R}$, associated with a pair of self-adjoint operators H_0, $H = H_0 + V$ with $V \in \mathcal{B}_1(\mathcal{H})$ ($\mathcal{H}$ a complex separable Hilbert space). In the special cases of sign-definite perturbations $V \geq 0$ and $V \leq 0$, $\Xi(\lambda, H_0, H)$ turns out to be a trace class operator in $\mathcal{H}$, whose trace coincides with Krein's spectral shift function $\xi(\lambda, H_0, H)$ for the pair (H_0, H). While the special case $V \geq 0$ has previously been studied by Carey [18], our aim here is to treat the case of general interactions V by separately introducing the positive and negative parts $V_\pm = (|V| \pm V)/2$ of V. In general, if V is not sign-definite, then $\Xi(\lambda, H_0, H)$ is not necessarily of trace class. However, we will introduce trace class operators $\Xi_\pm(\lambda)$ naturally associated with $V_\pm$, acting in distinct Hilbert spaces

$\mathcal{H}_\pm$, such that

$$\xi(\lambda, H_0, H) = \mathrm{Tr}_{\mathcal{H}_+}(\Xi_+(\lambda)) - \mathrm{Tr}_{\mathcal{H}_-}(\Xi_-(\lambda)) \text{ for a.e. } \lambda \in \mathbb{R}. \tag{1.1}$$

(An alternative approach to $\xi(\lambda, H_0, H)$, which does not rely on separately introducing V_+ and V_-, will be discussed elsewhere [36].)

Our main techniques are based on operator-valued Herglotz functions (continuing some of our recent investigations in this area [34], [37], [40]) and especially, on a detailed study of logarithms of Herglotz operators in Section 2. In Section 3 we introduce the spectral shift operator $\Xi(\lambda, H_0, H)$ associated with the pair (H_0, H) and relate it to Krein's spectral shift function $\xi(\lambda, H_0, H)$ and his celebrated trace formula [51]. Finally, Section 4 provides an application to spectral averaging originally due to Birman and Solomyak [11] and hints at operator-valued generalizations thereof. A number of additional applications of this formalism will appear elsewhere [36].

2. Logarithms of Operator-Valued Herglotz Functions

The principal purpose of this section is to study operator-valued Herglotz functions and particularly their logarithms and associated representation theorems. In this manner we shall obtain operator-valued generalizations of some of the classical results on exponential Herglotz representations studied by Aronszajn and Donoghue [4].

In the following $\mathcal{H}$ denotes a complex separable Hilbert space with scalar product $(\cdot, \cdot)_{\mathcal{H}}$ (linear in the second factor) and norm $\|\cdot\|_{\mathcal{H}}$, $I_{\mathcal{H}}$ the identity operator in $\mathcal{H}$, $\mathcal{B}(\mathcal{H})$ the Banach space of bounded linear operators defined on $\mathcal{H}$, $\mathcal{B}_p(\mathcal{H})$, $p \geq 1$ the standard Schatten-von Neumann ideals of $\mathcal{B}(\mathcal{H})$ (cf., e.g., [41], [77]) and $\mathbb{C}_+$ (resp., $\mathbb{C}_-$) the open complex upper (resp., lower) half-plane. Moreover, real and imaginary parts of a bounded operator $T \in \mathcal{B}(\mathcal{H})$ are defined as usual by $\mathrm{Re}(T) = (T + T^*)/2$, $\mathrm{Im}(T) = (T - T^*)/(2i)$.

Definition 2.1. $M : \mathbb{C}_+ \to \mathcal{B}(\mathcal{H})$ is called an *operator-valued Herglotz function* (in short, a *Herglotz operator*) if M is analytic on $\mathbb{C}_+$ and $\mathrm{Im}(M(z)) \geq 0$ for all $z \in \mathbb{C}_+$.

Theorem 2.2. (Birman and Entina [8], de Branges [26], Naboko [60]–[62].) *Let $M : \mathbb{C}_+ \to \mathcal{B}(\mathcal{H})$ be a Herglotz operator.*
(i) Then there exist bounded self-adjoint operators $A = A^* \in \mathcal{B}(\mathcal{H})$, $0 \leq B \in \mathcal{B}(\mathcal{H})$, a Hilbert space $\mathcal{K} \supseteq \mathcal{H}$, a self-adjoint operator $L = L^*$ in $\mathcal{K}$, a bounded nonnegative operator $0 \leq R \in \mathcal{B}(\mathcal{K})$ with $R|_{\mathcal{K} \ominus \mathcal{H}} = 0$ such that

$$M(z) = A + Bz + R^{1/2}(I_{\mathcal{K}} + zL)(L - z)^{-1} R^{1/2}|_{\mathcal{H}} \tag{2.1a}$$

$$= A + (B + R|_{\mathcal{H}})z + (1 + z^2) R^{1/2}(L - z)^{-1} R^{1/2}|_{\mathcal{H}}. \tag{2.1b}$$

(ii) Let $p \geq 1$. Then $M(z) \in \mathcal{B}_p(\mathcal{H})$ for all $z \in \mathbb{C}_+$ if and only if $M(z_0) \in \mathcal{B}_p(\mathcal{H})$ for some $z_0 \in \mathbb{C}_+$. In this case necessarily $A, B, R \in \mathcal{B}_p(\mathcal{H})$.
(iii) Let $M(z) \in \mathcal{B}_1(\mathcal{H})$ for some (and hence for all) $z \in \mathbb{C}_+$. Then $M(z)$ has normal boundary values $M(\lambda + i0)$ for (Lebesgue) a.e. $\lambda \in \mathbb{R}$ in every $\mathcal{B}_p(\mathcal{H})$-norm,

$p > 1$. Moreover, let $\{E_L(\lambda)\}_{\lambda\in\mathbb{R}}$ be the family of orthogonal spectral projections of L in $\mathcal{K}$. Then $R^{1/2}E_L(\lambda)R^{1/2}$ is $\mathcal{B}_1(\mathcal{H})$-differentiable for a.e. $\lambda\in\mathbb{R}$ and denoting the derivative by $d(R^{1/2}E_L(\lambda)R^{1/2})/d\lambda$, $\mathrm{Im}(M(z))$ has normal boundary values $\mathrm{Im}(M(\lambda+i0))$ for a.e. $\lambda\in\mathbb{R}$ in $\mathcal{B}_1(\mathcal{H})$-norm given by

$$\lim_{\varepsilon\downarrow 0}\|\pi^{-1}\mathrm{Im}(M(\lambda+i\varepsilon)) - d(R^{1/2}E_L(\lambda)R^{1/2}|_{\mathcal{H}})/d\lambda\|_{\mathcal{B}_1(\mathcal{H})} = 0 \ \text{a.e.} \tag{2.2}$$

Actually, the normal boundary values $M(\lambda+i0)$ in Theorem 2.2 (iii) can be replaced by nontangential ones. Since this distinction will play no role in the remainder of this paper we omit the corresponding details.

Theorem 2.2 (i) follows by considering the quadratic form $(\varphi, M(z)\varphi)_{\mathcal{H}}$, resulting in a scalar Herglotz function, in combination with Naimark's dilation theorem (cf., e.g., Theorem 1 in Appendix I of [16]). Details can be found in [60]. For Theorem 2.2 (ii), (iii) we refer to [61], [62]. In particular, p cannot be chosen equal to 1 in Theorem 2.2 (iii) (cf. [61]). Moreover, if $M(z_0)\in\mathcal{B}_p(\mathcal{H})$ for some $z_0\in\mathbb{C}_+$ and some $p>1$, then $M(z)$ need not even have boundary values $M(\lambda+i0)$ in the weak topology of $\mathcal{H}$ for a.e. $\lambda\in\mathbb{R}$. In fact, the quadratic form $(f, M(\lambda+i\varepsilon)f)_{\mathcal{H}}$ may converge for f in a fixed set $\mathcal{D}\subset\mathcal{H}$ (independent of λ) to $(f, M(\lambda+i0)f)_{\mathcal{H}}$ for a.e. $\lambda\in\mathbb{R}$ as $\varepsilon\downarrow 0$, however, $M(\lambda+i0)$ may be a densely defined unbounded operator in $\mathcal{H}$ for a.e. $\lambda\in\mathbb{R}$ (cf. [61]).

Originally, the existence of normal limits $M(\lambda+i0)$ for a.e. $\lambda\in\mathbb{R}$ in $\mathcal{B}_2(\mathcal{H})$-norm, in the special case $A=0$, $B=-R|_{\mathcal{H}}$, assuming $M(z)\in\mathcal{B}_1(\mathcal{H})$, was proved by de Branges [26] in 1962. (The more general case in (2.1) can easily be reduced to this special case.) In his paper [26], de Branges also proved the existence of normal limits $\mathrm{Im}(M(\lambda+i0))$ for a.e. $\lambda\in\mathbb{R}$ in $\mathcal{B}_1(\mathcal{H})$-norm and obtained (2.2). These results and their implications on stationary scattering theory were subsequently studied in detail by Birman and Entina [7], [8]. (Textbook representations of this material can also be found in [5], Ch. 3.)

Invoking the family of orthogonal spectral projections $\{E_L(\lambda)\}_{\lambda\in\mathbb{R}}$ of L, (2.1) then yields the familiar representation

$$M(z) = A + Bz + \int_{\mathbb{R}} (1+\lambda^2)d(R^{1/2}E_L(\lambda)R^{1/2}|_{\mathcal{H}})((\lambda-z)^{-1} - \lambda(1+\lambda^2)^{-1}), \tag{2.3}$$

where for our purpose it suffices to interpret the integral in (2.3) in the weak sense. Further results on representations of the type (2.3) can be found in [16], Sect. I.4 and [76].

Before we continue our investigations on Herglotz operators a few comments concerning our terminology are perhaps in order. The representation (2.3) in the special case of scalar Herglotz functions is due to Riesz and Herglotz, respectively, Nevanlinna. The former authors discussed the analog of (2.3) in the open unit disk $D\subset\mathbb{C}$, whereas the latter studied (2.3) in $\mathbb{C}_+$. As a consequence, functions of the type (2.3) are frequently called Herglotz or Nevanlinna functions. Moreover, researchers in the former Soviet Union coined the term R-functions. Here we follow the traditional terminology in mathematical physics which seems to prefer the notion of Herglotz functions.

Since we are interested in logarithms of Herglotz operators, questions of their invertibility naturally arise. The following result clarifies the situation.

Lemma 2.3. *Suppose M is a Herglotz operator with values in $\mathcal{B}(\mathcal{H})$. If $M(z_0)^{-1} \in \mathcal{B}(\mathcal{H})$ for some $z_0 \in \mathbb{C}_+$ then $M(z)^{-1} \in \mathcal{B}(\mathcal{H})$ for all $z \in \mathbb{C}_+$.*

Proof. Suppose there is a $z_1 \in \mathbb{C}_+$ and a sequence $\{e_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$ such that $||e_n||_{\mathcal{H}} = 1$, $n \in \mathbb{N}$ but $\lim_{n\to\infty} \|M(z_1)e_n\|_{\mathcal{H}} = 0$. Then $\lim_{n\to\infty}(e_n, \mathrm{Im}(M(z_1))e_n)_{\mathcal{H}} = \lim_{n\to\infty} ||(\mathrm{Im}(M(z_1)))^{1/2}e_n||_{\mathcal{H}} = 0$ and hence

$$\lim_{n\to\infty} \|\mathrm{Im}(M(z_1))e_n\|_{\mathcal{H}} = 0. \tag{2.4}$$

By (2.1a),

$$\begin{aligned}\mathrm{Im}(M(z)) &= \mathrm{Im}(z)(B + R^{1/2}(I_{\mathcal{K}} + L^2)((L - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1}R^{1/2}|_{\mathcal{H}}) \\ &= \mathrm{Im}(z)(B + R^{1/2}C(z)R^{1/2}|_{\mathcal{H}}), \end{aligned} \tag{2.5}$$

where $C(z) = (I_{\mathcal{K}}+L^2)((L-\mathrm{Re}(z))^2+(\mathrm{Im}(z))^2)^{-1} > 0$ is invertible in $\mathcal{K}$, $C(z)^{-1} \in \mathcal{B}(\mathcal{K})$, $z \in \mathbb{C}_+$. Thus, (2.4) implies $\lim_{n\to\infty}((e_n, Be_n)_{\mathcal{H}} + (R^{1/2}|_{\mathcal{H}}e_n, C(z_1)R^{1/2}|_{\mathcal{H}} e_n)_{\mathcal{H}}) = 0$ and hence $\lim_{n\to\infty} \|B^{1/2}e_n\|_{\mathcal{H}} = 0$, and $\lim_{n\to\infty} \|Be_n\|_{\mathcal{H}} = 0$, $\lim_{n\to\infty} \|R^{1/2}|_{\mathcal{H}}e_n\|_{\mathcal{H}} = 0$. Consequently,

$$M(z_1)e_n = Ae_n + z_1Be_n + (R^{1/2}(I_{\mathcal{K}} + z_1L)(L - z_1)^{-1})R^{1/2}|_{\mathcal{H}}e_n \tag{2.6}$$

with $R^{1/2}(I_{\mathcal{K}} + z_1L)(L - z_1)^{-1} \in \mathcal{B}(\mathcal{K})$ yields $\lim_{n\to\infty} \|Ae_n\|_{\mathcal{H}} = 0$. Applying (2.1a) again we infer $\lim_{n\to\infty} \|M(z)e_n\|_{\mathcal{H}} = 0$ for all $z \in \mathbb{C}_+$ contradicting our hypothesis $M(z_0)^{-1} \in \mathcal{B}(\mathcal{H})$. This argument shows two facts. First, by choosing $e_n = e \in \ker(M(z_1))$ it yields $\ker(M(z)) = \{0\}$ for all $z \in \mathbb{C}_+$, and second, it proves the boundedness of $M(z)^{-1} : \mathrm{ran}(M(z)) \to \mathcal{H}$, $z \in \mathbb{C}_+$. In particular, one infers that $\mathrm{ran}(M(z)) = \mathrm{dom}(M(z)^{-1})$ is a closed subspace of $\mathcal{H}$. By exactly the same argument one derives $\{0\} = \ker(M(z)^*) = (\mathrm{ran}(M(z))^{\perp}, \quad z \in \mathbb{C}_+$ and thus, $M(z)^{-1} \in \mathcal{B}(\mathcal{H})$ for all $z \in \mathbb{C}_+$. ☐

Lemma 2.3 admits the following generalization in connection with parametrices, that is, generalized inverses, familar from treatments of the Calkin algebra. In the following we denote by $\mathcal{F}(\mathcal{H})$ the algebra of finite-rank operators on $\mathcal{H}$ and recall that T is called a left (resp., right) parametrix of $M(z_0)$ if $(TM(z_0) - I_{\mathcal{H}}) \in \mathcal{F}(\mathcal{H})$ (resp., $(M(z_0)T - I_{\mathcal{H}}) \in \mathcal{F}(\mathcal{H})$).

Lemma 2.4. *Suppose M is a Herglotz operator with values in $\mathcal{B}(\mathcal{H})$. If $M(z_0)$ has a left* (*resp., right*) *parametrix for some $z_0 \in \mathbb{C}_+$, then $M(z)$ has a left and right parametrix for all $z \in \mathbb{C}_+$.*

Proof. From the proof of Lemma 2.3 one infers

$$\ker(M(z_1)) = \ker(M(z_2)) \text{ for all } z_1, z_2 \in \mathbb{C}_+, \tag{2.7a}$$

$$\ker(M(z)^*) = \ker(M(z)) = (\mathrm{ran}(M(z)))^{\perp}, \quad z \in \mathbb{C}_+ \tag{2.7b}$$

and hence $M(z)$ can be decomposed as

$$M(z) = \begin{pmatrix} \widetilde{M}(z) & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{H} = \widetilde{\mathcal{H}} \oplus \ker(M(i)), \tag{2.8}$$

where

$$\ker(\widetilde{M}(z)) = \{0\}, \quad \overline{\operatorname{ran}(\widetilde{M}(z))} = \widetilde{\mathcal{H}}, \quad z \in \mathbb{C}_+. \tag{2.9}$$

From $TM(z_0) - I_{\mathcal{H}} = F \in \mathcal{F}(\mathcal{H})$ one concludes $\ker(M(z_0)) \subseteq \ker(I_{\mathcal{H}} + F)$ and hence

$$\dim(\ker(M(z))) < \infty \text{ for all } z \in \mathbb{C}_+. \tag{2.10}$$

Denote by $\widetilde{P}$ the orthogonal projection onto $\widetilde{\mathcal{H}}$, $\widetilde{\mathcal{H}} = \widetilde{P}\mathcal{H}$ then $\widetilde{P}TM(z_0)\widetilde{P} = \widetilde{P}T\widetilde{P}\widetilde{M}(z_0) = \widetilde{P}(I_{\mathcal{H}} + F)\widetilde{P} = I_{\tilde{\mathcal{H}}} + \widetilde{F}$. In order to prove that $\widetilde{M}(z_0)^{-1} \in \mathcal{B}(\widetilde{\mathcal{H}})$ we argue as follows: Suppose there is a sequence $\{e_n\}_{n\in\mathbb{N}} \subset \widetilde{\mathcal{H}}$, $\|e_n\|_{\tilde{\mathcal{H}}} = 1$ such that

$$\lim_{n\to\infty} \|\widetilde{M}(z_0)e_n\|_{\tilde{\mathcal{H}}} = 0. \tag{2.11}$$

Hence $\lim_{n\to\infty} \|\widetilde{P}T\widetilde{P}\widetilde{M}(z_0)e_n\|_{\tilde{\mathcal{H}}} = \lim_{n\to\infty} \|(I_{\tilde{\mathcal{H}}} + \widetilde{F})e_n\|_{\tilde{\mathcal{H}}} = 0$. Next, let $\widetilde{E}_{-1}$ be the spectral projection onto the geometric eigenspace of $\widetilde{F}$ corresponding to the eigenvalue -1, and

$$\widetilde{E}_{-1} \dotplus \widetilde{E}_{-1}^{\perp} = I_{\tilde{\mathcal{H}}} \tag{2.12}$$

(here $\dotplus$ denotes the direct sum but not necessarily the orthogonal direct sum $\oplus$). Then

$$\lim_{n\to\infty} \|(I_{\tilde{\mathcal{H}}} + \widetilde{F})\widetilde{E}_{-1}^{\perp}e_n\|_{\tilde{\mathcal{H}}} = 0. \tag{2.13}$$

Since $I_{\tilde{\mathcal{H}}} + \widetilde{F}$ is boundedly invertible on $\widetilde{E}_{-1}^{\perp}\widetilde{\mathcal{H}}$, one infers

$$\lim_{n\to\infty} \|\widetilde{E}_{-1}^{\perp}e_n\|_{\tilde{\mathcal{H}}} = 0. \tag{2.14}$$

Introducing $\tilde{e}_n = \widetilde{E}_{-1}e_n$, $n \in \mathbb{N}$, one concludes from (2.11), (2.12)–(2.14) that $\lim_{n\to\infty} \|\tilde{e}_n\|_{\tilde{\mathcal{H}}} = 1$, $\lim_{n\to\infty} \|\widetilde{M}(z_0)\tilde{e}_n\|_{\tilde{\mathcal{H}}} = 0$. Since $\{\tilde{e}_n\}_{n\in\mathbb{N}}$ is compact, there exists a subsequence $\{\tilde{e}_{n_k}\}_{k\in\mathbb{N}} \subseteq \{\tilde{e}_n\}_{n\in\mathbb{N}}$ such that $\lim_{k\to\infty} \|\tilde{e}_{n_k} - \tilde{e}_0\|_{\tilde{\mathcal{H}}} = 0$, $\tilde{e}_0 \in \operatorname{ran}(\widetilde{E}_{-1})$, $\|\tilde{e}_0\|_{\tilde{\mathcal{H}}} = 1$. Hence $\lim_{k\to\infty} \|\widetilde{M}(z_0)\tilde{e}_{n_k}\|_{\tilde{\mathcal{H}}} = 0$ implying the contradiction $\widetilde{M}(z_0)\tilde{e}_0 = 0$. Thus, $\widetilde{M}(z_0)^{-1} \in \mathcal{B}(\widetilde{\mathcal{H}})$ and hence $\widetilde{M}(z)^{-1} \in \mathcal{B}(\widetilde{\mathcal{H}})$, $z \in \mathbb{C}_+$ by Lemma 2.3. In particular, $\operatorname{ran}(\widetilde{M}(z)) = \widetilde{\mathcal{H}}$, $z \in \mathbb{C}_+$ and hence $M(z) = \left(\begin{smallmatrix} \tilde{M}(z) & 0 \\ 0 & 0 \end{smallmatrix}\right)$ has left and right parametrices taking into account (2.10). ☐

Concerning boundary values at the real axis we also recall

Lemma 2.5. (Naboko [63].) *Suppose $(M - I_{\mathcal{H}})$ is a Herglotz operator with values in $\mathcal{B}_1(\mathcal{H})$. Then the boundary values $M(\lambda + i0)$ exist for a.e. $\lambda \in \mathbb{R}$ in $\mathcal{B}_p(\mathcal{H})$-norm, $p > 1$ and $M(\lambda + i0)$ is a Fredholm operator for a.e. $\lambda \in \mathbb{R}$ with index zero a.e.,*

$$\operatorname{ind}(M(\lambda + i0)) = 0 \textit{ for a.e. } \lambda \in \mathbb{R}. \tag{2.15}$$

Moreover,

$$\ker(M(\lambda+i0)) = \ker(M(i)) = (\operatorname{ran}(M(\lambda+i0)))^{\perp} \text{ for a.e. } \lambda\in\mathbb{R}. \tag{2.16}$$

In addition, if $M(z_0)^{-1}\in\mathcal{B}(\mathcal{H})$ *for some (and hence for all)* $z_0\in\mathbb{C}_+$, *then*

$$M(\lambda+i0)^{-1}\in\mathcal{B}(\mathcal{H}) \text{ for a.e. } \lambda\in\mathbb{R}. \tag{2.17}$$

Generalizations and a counter example to (2.17) if $(M-I_{\mathcal{H}})\in\mathcal{B}_1(\mathcal{H})$ is replaced by $(M-I_{\mathcal{H}})\in\mathcal{B}_p(\mathcal{H})$, $p>1$ can be found in [63].

Next, let T be a bounded *dissipative* operator, that is,

$$T\in\mathcal{B}(\mathcal{H}), \quad \operatorname{Im}(T)\geq 0. \tag{2.18}$$

In order to define the logarithm of T we use the integral representation

$$\log(z) = -i\int_0^\infty d\lambda\,((z+i\lambda)^{-1}-(1+i\lambda)^{-1}), \quad z\neq -iy,\, y\geq 0, \tag{2.19}$$

with a cut along the negative imaginary z-axis. We use the symbol $\log(\cdot)$ in (2.19) in order to distinguish it from the integral representation

$$\ln(z) = \int_{-\infty}^0 d\lambda\,((\lambda-z)^{-1}-\lambda(1+\lambda^2)^{-1}), \quad z\in\mathbb{C}\backslash(-\infty,0] \tag{2.20}$$

with a cut along the negative real axis. Both representations will be used later and it is easily verified that $\log(\cdot)$ and $\ln(\cdot)$ coincide for $z\in\mathbb{C}_+$. In particular, one verifies that (2.19) and (2.20) are Herglotz functions, that is, they are analytic in $\mathbb{C}_+$ and

$$0<\operatorname{Im}(\log(z)),\ \operatorname{Im}(\ln(z))\ <\pi, \quad z\in\mathbb{C}_+. \tag{2.21}$$

Lemma 2.6. *Suppose* $T\in\mathcal{B}(\mathcal{H})$ *is dissipative and* $T^{-1}\in\mathcal{B}(\mathcal{H})$. *Define*

$$\log(T) = -i\int_0^\infty d\lambda\,((T+i\lambda)^{-1}-(1+i\lambda)^{-1}I_{\mathcal{H}}) \tag{2.22}$$

in the sense of a $\mathcal{B}(\mathcal{H})$*-norm convergent Riemann integral. Then*
(i) $\log(T)\in\mathcal{B}(\mathcal{H})$.
(ii) If $T=zI_{\mathcal{H}}$, $z\in\mathbb{C}_+$, *then* $\log(T)=\log(z)I_{\mathcal{H}}$.
(iii) Suppose $\{P_n\}_{n\in\mathbb{N}}\subset\mathcal{B}(\mathcal{H})$ *is a family of orthogonal finite-rank projections in* $\mathcal{H}$ *with* $\operatorname{s-lim}_{n\to\infty}P_n=I_{\mathcal{H}}$. *Then*

$$\operatorname*{s-lim}_{n\to\infty}((I_{\mathcal{H}}-P_n)+P_nTP_n)=T$$

and

$$\begin{aligned}&\operatorname*{s-lim}_{n\to\infty}\log((I_{\mathcal{H}}-P_n)+P_n(T+i\varepsilon)P_n)\\ &=\operatorname*{s-lim}_{n\to\infty}P_n(\log(P_n(T+i\varepsilon)P_n|_{P_n\mathcal{H}})P_n=\log(T+i\varepsilon I_H), \quad \varepsilon>0.\end{aligned}$$

(iv) $\operatorname{n-lim}_{\varepsilon\downarrow 0}\log(T+i\varepsilon I_{\mathcal{H}})=\log(T)$.
(v) $e^{\log(T)}=T$.

Proof. (i) Clearly $\log(T) \in \mathcal{B}(\mathcal{H})$ since

$$||\log(T)|| \leq \int_0^\delta d\lambda\,(||(T+i\lambda)^{-1}||+1) + \int_\delta^\infty d\lambda\,(||T||+1)\lambda^{-2} \tag{2.23}$$

using $||(T+i\lambda)^{-1}|| \leq \lambda^{-1}$, $\lambda > 0$. Moreover, by $||(T+i\lambda)^{-1}|| \leq ||T^{-1}||(1 - ||T^{-1}||\lambda)^{-1}$ for $0 < \lambda < ||T^{-1}||^{-1}$, choosing $\delta = 2^{-1}||T^{-1}||^{-1}$ is sufficient to bound the first integral in (2.23).

(ii) is obvious from (2.19).

(iii) For any $f \in \mathcal{H}$, $\varepsilon > 0$,

$$\begin{aligned}
&\left\| \int_0^\infty d\lambda\,(((I_{\mathcal{H}}-P_n)+P_n(T+i\varepsilon)P_n+i\lambda)^{-1} - (T+i\varepsilon+i\lambda)^{-1})f \right\| \\
&\leq 2\int_0^\delta d\lambda\,((\varepsilon+\lambda)^{-1})||f|| \\
&+ \int_\delta^N d\lambda\,||((I_{\mathcal{H}}-P_n)+P_nTP_n - T + i\varepsilon(P_n - I_{\mathcal{H}}))(T+i\varepsilon+i\lambda)^{-1}f||\delta^{-1} \\
&+ \int_N^\infty d\lambda\,\lambda^{-2}\|((I_{\mathcal{H}}-P_n)+P_n(T+i\varepsilon)P_n - T - i\varepsilon I_{\mathcal{H}})\|\|f||.
\end{aligned} \tag{2.24}$$

Taking $\delta > 0$ sufficiently small and N sufficiently large such that the first and third integrals in (2.24) are sufficiently small uniformly with respect to $n \in \mathbb{N}$ for fixed $f \in \mathcal{H}$, it suffices to let $n \to \infty$ in the second integral in (2.24) for fixed δ and N.

(iv) One estimates,

$$\begin{aligned}
&||\log(T+i\varepsilon) - \log(T)|| = \left\| \int_0^\infty d\lambda\,((T+i\varepsilon+i\lambda)^{-1} - (T+i\lambda)^{-1}) \right\| \\
&\leq \int_0^\delta d\lambda\,(||T^{-1}||(1-(\varepsilon+\lambda)||T^{-1}||)^{-1} + ||T^{-1}||(1-\lambda||T^{-1}||)^{-1}) \\
&+ \varepsilon\int_\delta^\infty d\lambda\,\lambda^{-2} \leq 2\delta(||T^{-1}||^{-1} - (\varepsilon+\delta))^{-1} + \varepsilon\delta^{-1}
\end{aligned} \tag{2.25}$$

for $\varepsilon + \delta < ||T^{-1}||^{-1}$. Taking $\delta \leq 2^{-1}||T^{-1}||^{-1}$, $\varepsilon < \delta/2$ and $\varepsilon \downarrow 0$ then yields the result since $\delta > 0$ can be chosen arbitrarily small.

(v) clearly holds for dissipative $n \times n$ matrices and hence

$$e^{\log((I_{\mathcal{H}}-P_n)+P_n(T+i\varepsilon)P_n)} = (I_{\mathcal{H}}-P_n)+P_n(T+i\varepsilon)P_n, \quad \varepsilon > 0 \tag{2.26}$$

upon decomposing $\mathcal{H} = P_n\mathcal{H} \oplus (I_{\mathcal{H}}-P_n)\mathcal{H}$, where P_n are orthogonal rank-n projections. Since $\text{s-lim}_{n\to\infty} e^{A_n} = e^A$ whenever $A_n, A \in \mathcal{B}(\mathcal{H})$ and $\text{s-lim}_{n\to\infty} A_n = A$ (simply use $e^B = \sum_{m=0}^\infty B^m/(m!)$, $B \in \mathcal{B}(\mathcal{H})$ and $A_n^m - A^m = \sum_{\ell=0}^{m-1} A_n^\ell(A_n - A)A^{m-1-\ell}$), one infers from (iii) and (2.26) $e^{\log(T+i\varepsilon)} = T + i\varepsilon$, $\varepsilon > 0$. Together with (iv) this yields (v). $\square$

Lemma 2.7. *Suppose $T \in \mathcal{B}(\mathcal{H})$ is dissipative and $T^{-1} \in \mathcal{B}(\mathcal{H})$. Let L be the minimal self-adjoint dilation of T in the Hilbert space $\mathcal{K} \supseteq \mathcal{H}$. Then*

$$\mathrm{Im}(\log(T)) = \pi P_{\mathcal{H}} E_L((-\infty, 0))|_{\mathcal{H}}, \tag{2.27}$$

where $P_{\mathcal{H}}$ is the orthogonal projection in $\mathcal{K}$ onto $\mathcal{H}$ and $\{E_L(\lambda)\}_{\lambda\in\mathbb{R}}$ is the family of orthogonal spectral projections of L in $\mathcal{K}$. In particular,

$$0 \leq \mathrm{Im}(\log(T)) \leq \pi I_{\mathcal{H}}. \tag{2.28}$$

Proof. By Sz.-Nagy's dilation theorem (see the corresponding result in [86], Ch. III, Sect. 2, Theorem 2.1 for contractions), one infers

$$(T + i\lambda)^{-1} = P_{\mathcal{H}}(L + i\lambda)^{-1}|_{\mathcal{H}}, \quad \lambda > 0, \tag{2.29}$$

where L is the minimal self-adjoint dilation of T in $\mathcal{K}$. Then the existence of $T^{-1} \in \mathcal{B}(\mathcal{H})$ and (2.29) yield

$$E_L(\{0\}) = 0. \tag{2.30}$$

In order to prove (2.30) one can argue as follows. Consider the contraction $S = (T-i)(T+i)^{-1}$. According to Theorem 3.2, Ch. I, Sect. 3 in [86], every contraction on the Hilbert space $\mathcal{H}$ corresponds to a decomposition of $\mathcal{H}$ into an orthogonal sum $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$ of two reducing subspaces of S such that the part of S on $\mathcal{H}_0$ is unitary, and the part of S on $\mathcal{H}_1$ is completely non-unitary. Moreover, this decomposition is unique. (We recall that a contraction is called completely non-unitary if there are no non-zero subspaces reducing this contraction to a unitary operator.) Since T is invertible, the unitary part of the contraction S does not have the eigenvalue -1. Since the minimal unitary dilation of a completely non-unitary contraction has absolutely continuous spectrum (see [85]), we conclude that the minimal unitary dilation of S does not have the eigenvalue -1 and hence the kernel of the minimal self-adjoint dilation L of the dissipative operator T is trivial, proving (2.30). Next, (2.22) implies

$$\begin{aligned}
\mathrm{Im}(\log(T)) &= -\int_0^\infty d\lambda\, \mathrm{Re}((T+i\lambda)^{-1} - (1+i\lambda)^{-1} I_{\mathcal{H}}) \\
&= -\int_0^\varepsilon d\lambda\, \mathrm{Re}((T+i\lambda)^{-1} - (1+i\lambda)^{-1} I_{\mathcal{H}}) \\
&\quad - P_{\mathcal{H}} \int_\varepsilon^\infty d\lambda\, \mathrm{Re}((L+i\lambda)^{-1} - (1+i\lambda)^{-1} I_{\mathcal{K}})|_{\mathcal{H}} \\
&= -\int_0^\varepsilon d\lambda\, \mathrm{Re}((T+i\lambda)^{-1} - (1+i\lambda)^{-1} I_{\mathcal{H}}) \\
&\quad - P_{\mathcal{H}} \int_\varepsilon^\infty d\lambda\, (L(L^2+\lambda^2)^{-1} - (1+\lambda^2)^{-1} I_{\mathcal{K}})|_{\mathcal{H}}, \quad \varepsilon > 0.
\end{aligned} \tag{2.31}$$

Let us prove that

$$\operatorname*{s-lim}_{\varepsilon\downarrow 0} \int_\varepsilon^\infty d\lambda\, L(L^2+\lambda^2)^{-1} = (\pi/2)(P_{L,+} - P_{L,-}), \tag{2.32}$$

where $P_{L,+} = E_L((0,\infty))$ and $P_{L,-} = E_L((-\infty,0))$ are the spectral projections of L corresponding to the half-lines $(0,\infty)$ and $(-\infty,0)$. Then (2.27) and hence (2.28) follow from (2.31), (2.32) and the fact that

$$\lim_{\varepsilon\downarrow 0}\left\| \int_0^\varepsilon d\lambda\, \mathrm{Re}((T+i\lambda)^{-1} - (1+i\lambda)^{-1}I_{\mathcal H})\right\| = 0. \tag{2.33}$$

Using the estimate $|\int_\varepsilon^\infty d\lambda\, \mu(\mu^2+\lambda^2)^{-1}| \le (\pi/2)$, $\mu\in\mathbb R$, $\varepsilon>0$ by the spectral theorem for the self-adjoint operator L, we infer that the family of operators $\int_\varepsilon^\infty d\lambda\, L(L^2+\lambda^2)^{-1}$ is uniformly bounded in $\varepsilon>0$ and therefore, it suffices to check the convergence (2.32) on a dense set in $\mathcal K$. A natural candidate for this set is $\mathcal D = \{\bigcup_{\delta>0} E_L(\mathbb R\backslash(-\delta,\delta))f \mid f\in\mathcal K\}$, which is dense in $\mathcal K$ since by (2.30) the kernel of L is trivial. For $f\in\mathcal D$ we have

$$\begin{aligned}
&\int_\varepsilon^\infty d\lambda\, L(L^2+\lambda^2)^{-1}f = \int_\varepsilon^\infty d\lambda \int_{\mathbb R\backslash(-\delta,\delta)} \mu(\mu^2+\lambda^2)^{-1} dE_L(\mu)f\\
&= \int_\delta^\infty ((\pi/2) - \arctan(\varepsilon/\mu))dE_L(\mu)f - \int_{-\infty}^{-\delta} ((\pi/2)-\arctan(\varepsilon/|\mu|))dE_L(\mu)f\\
&= (\pi/2)(E_L((\delta,\infty)) - E_L((-\infty,-\delta)))f\\
&- \int_\delta^\infty \arctan(\varepsilon/\mu)dE_L(\mu)f + \int_{-\infty}^{-\delta}\arctan(\varepsilon/|\mu|)dE_L(\mu)f
\end{aligned} \tag{2.34}$$

for all $\delta>0$, $\delta=\delta(f)$ small enough. For fixed $f\in\mathcal D$, going to the limit $\varepsilon\to 0$ in (2.34), we get

$$\begin{aligned}
\lim_{\varepsilon\downarrow 0}\int_\varepsilon^\infty d\lambda\, L(L^2+\lambda^2)^{-1}f &= (\pi/2)(E_L((\delta,\infty)) - E_L((-\infty,\delta)))f\\
&= (\pi/2)(P_{L,+} - P_{L,-})f,
\end{aligned} \tag{2.35}$$

proving (2.32). ☐

Combining Lemmas 2.3 and 2.7 one can prove the following result.

Lemma 2.8. *Suppose $M:\mathbb C_+ \longrightarrow \mathcal B(\mathcal H)$ is a Herglotz operator and $M(z_0)^{-1}\in\mathcal B(\mathcal H)$ for some (and hence for all) $z_0\in\mathbb C_+$. Then $\log(M):\mathbb C_+\longrightarrow\mathcal B(\mathcal H)$ is a Herglotz operator and*

$$0\le \mathrm{Im}(\log(M(z)))\le \pi I_{\mathcal H}, \quad z\in\mathbb C_+. \tag{2.36}$$

Proof. Clearly

$$\log(M(z)) = -i\int_0^\infty d\lambda\, ((M(z)+i\lambda)^{-1} - (1+i\lambda)^{-1}I_{\mathcal H}), \quad z\in\mathbb C_+ \tag{2.37}$$

is analytic for $z\in\mathbb C_+$ since $M(z)^{-1}\in\mathcal B(\mathcal H)$ for all $z\in\mathbb C_+$ by Lemma 2.3. An application of Lemma 2.7 then yields (2.36). ☐

Thus applying (2.1a) to $\log(M(z))$ one infers

$$\log(M(z)) = C + Dz + \widetilde R^{1/2}(I_{\widetilde{\mathcal K}} + z\widetilde L)(\widetilde L - z)^{-1}\widetilde R^{1/2}|_{\mathcal H}, \quad z\in\mathbb C_+ \tag{2.38}$$

for some Hilbert space $\widetilde{\mathcal{K}} \supseteq \mathcal{H}$, some bounded self-adjoint operators C, $0 \leq D \in \mathcal{B}(\mathcal{H})$, a bounded nonnegative operator $0 \leq \widetilde{R} \in \mathcal{B}(\mathcal{H})$ with $\widetilde{R}|_{\widetilde{\mathcal{K}}\ominus\mathcal{H}} = 0$, and a self-adjoint operator $\widetilde{L} = \widetilde{L}^*$ in $\widetilde{\mathcal{K}}$. By comparison with scalar Herglotz functions one would expect that $D = 0$. That this is indeed the case is proved next.

Lemma 2.9. *$D = 0$ in the representation* (2.38) *for* $\log(M(z))$, $z \in \mathbb{C}_+$.

Proof. Consider $z = iy$, $y \uparrow \infty$. Then

$$||y^{-1}\log(M(iy)) - Di|| = O(y^{-1}) \tag{2.39}$$

and by (2.36),

$$\begin{aligned}\pi y^{-1} I_{\mathcal{H}} &\geq y^{-1}\mathrm{Im}(\log(M(iy)))\\ &= D + y^{-1}\mathrm{Im}(\widetilde{R}^{1/2}(I_{\widetilde{\mathcal{K}}} + iy\widetilde{L})(\widetilde{L} - iy)^{-1}\widetilde{R}^{1/2}|_{\mathcal{H}})\\ &= D + y^{-1}(\ldots \geq 0 \ldots) \geq D \geq 0 \end{aligned}\tag{2.40}$$

and hence $D = 0$ since y^{-1} can be chosen arbitrarily small. □

Introducing the family of orthogonal spectral projections $\{E_{\widetilde{L}}(\lambda)\}_{\lambda\in\mathbb{R}}$ of $\widetilde{L}$ in $\widetilde{\mathcal{K}}$ one can thus rewrite (2.38) as

$$\begin{aligned}&\log(M(z)) = C\\ &+ \int_{\mathbb{R}} (1+\lambda^2) d(\widetilde{R}^{1/2} E_{\widetilde{L}}(\lambda)\widetilde{R}^{1/2}|_{\mathcal{H}})((\lambda - z)^{-1} - \lambda(1+\lambda^2)^{-1}), \quad z \in \mathbb{C}_+\end{aligned}\tag{2.41}$$

and hence

$$\mathrm{Im}(\log(M(z))) = \mathrm{Im}(z)\int_{\mathbb{R}} d(\widetilde{R}^{1/2} E_{\widetilde{L}}(\lambda)\widetilde{R}^{1/2}|_{\mathcal{H}})(1+\lambda^2)|\lambda - z|^{-2} \tag{2.42}$$

for $z \in \mathbb{C}_+$, interpreting both integrals in (2.41) and (2.42) in the weak sense for simplicity. Again by comparison with scalar Herglotz functions one expects that $d(\widetilde{R}^{1/2} E_{\widetilde{L}}(\lambda)\ \widetilde{R}^{1/2}|_{\mathcal{H}})$ is a purely absolutely continuous operator-valued measure on $\mathbb{R}$. This is confirmed by the following result ($\mathrm{Tr}_{\mathcal{H}}(\cdot)$ denotes the trace of trace class operators in $\mathcal{H}$).

Theorem 2.10. *Suppose $M : \mathbb{C}_+ \longrightarrow \mathcal{B}(\mathcal{H})$ is a Herglotz operator and $M(z_0)^{-1} \in \mathcal{B}(\mathcal{H})$ for some (and hence for all) $z_0 \in \mathbb{C}_+$. Then there exists a family of bounded self-adjoint weakly (Lebesgue) measurable operators $\{\Xi(\lambda)\}_{\lambda\in\mathbb{R}} \subset \mathcal{B}(\mathcal{H})$,*

$$0 \leq \Xi(\lambda) \leq I_{\mathcal{H}} \text{ for a.e. } \lambda \in \mathbb{R} \tag{2.43}$$

such that

$$\log(M(z)) = C + \int_{\mathbb{R}} d\lambda\, \Xi(\lambda)((\lambda - z)^{-1} - \lambda(1+\lambda^2)^{-1}), \quad z \in \mathbb{C}_+ \tag{2.44}$$

the integral taken in the weak sense, where $C = C^* \in \mathcal{B}(\mathcal{H})$. Moreover, if $\mathrm{Im}(\log(M(z_0))) \in \mathcal{B}_1(\mathcal{H})$ *for some (and hence for all) $z_0 \in \mathbb{C}_+$, then*

$$0 \leq \Xi(\lambda) \in \mathcal{B}_1(\mathcal{H}) \text{ for a.e. } \lambda \in \mathbb{R}, \tag{2.45}$$

$$0 \leq \operatorname{Tr}_{\mathcal{H}}(\Xi(\cdot)) \in L^1_{\mathrm{loc}}(\mathbb{R}; d\lambda), \qquad \int_{\mathbb{R}} d\lambda\, (1+\lambda^2)^{-1} \operatorname{Tr}_{\mathcal{H}}(\Xi(\lambda)) < \infty, \tag{2.46}$$

and

$$\operatorname{Tr}_{\mathcal{H}}(\operatorname{Im}(\log(M(z)))) = \operatorname{Im}(z) \int_{\mathbb{R}} d\lambda\ \operatorname{Tr}_{\mathcal{H}}(\Xi(\lambda))|\lambda - z|^{-2}, \quad z \in \mathbb{C}_+. \tag{2.47}$$

Proof. Let $f \in \mathcal{H}$ and denote

$$d\omega_f(\lambda) = (1+\lambda^2) d(\widetilde{R}^{1/2} f, E_{\widetilde{L}}(\lambda)\widetilde{R}^{1/2} f)_{\mathcal{H}} = (1+\lambda^2) d||E_{\widetilde{L}}(\lambda)\widetilde{R}^{1/2} f||^2_{\mathcal{H}}. \tag{2.48}$$

Then

$$0 \leq (f, \operatorname{Im}(\log(M(z)))f)_{\mathcal{H}} \leq \pi ||f||^2_{\mathcal{H}}, \quad z \in \mathbb{C}_+, \quad f \in \mathcal{H} \tag{2.49}$$

proves that $d\omega_f$ is purely absolutely continuous, $d\omega_f = d\omega_{f,ac}$ by standard arguments (see, e.g., [4]). Thus,

$$d\omega_f(\lambda) = \xi_f(\lambda) d\lambda \text{ with } 0 \leq \xi_f(\lambda) \leq ||f||^2_{\mathcal{H}} \quad \text{ for all } f \in \mathcal{H} \text{ and a.e. } \lambda \in \mathbb{R}. \tag{2.50}$$

By (2.48), $\xi_f(\lambda)$ defines a quadratic form

$$\xi_f(\lambda) = (f, \Xi(\lambda) f)_{\mathcal{H}} \text{ for some } 0 \leq \Xi(\lambda) \leq I_{\mathcal{H}} \text{ and a.e. } \lambda \in \mathbb{R} \tag{2.51}$$

proving (2.43) and (2.44). The representation (2.38) (with $D = 0$) implies

$$\operatorname{Im}(\log(M(z)) = \widetilde{R}^{1/2}(I_{\mathcal{K}} + \widetilde{L}^2)((\widetilde{L} - \operatorname{Re}(z))^2 + (\operatorname{Im}(z))^2)^{-1} \widetilde{R}^{1/2}|_{\mathcal{H}} \tag{2.52}$$

and hence $\operatorname{Im}(\log(M(z_0))) \in \mathcal{B}_1(\mathcal{H})$ for some $z_0 \in \mathbb{C}_+$ implies $\operatorname{Im}(\log(M(z))) \in \mathcal{B}_1(\mathcal{H})$ for all $z \in \mathbb{C}_+$. In particular,

$$\operatorname{Im}(\log(M(i))) = \int_{\mathbb{R}} d\lambda\, (1+\lambda^2)^{-1} \Xi(\lambda) \tag{2.53}$$

then proves

$$\operatorname{Tr}_{\mathcal{H}}(\operatorname{Im}(\log(M(i)))) = \int_{\mathbb{R}} d\lambda\, (1+\lambda^2)^{-1} \operatorname{Tr}_{\mathcal{H}}(\Xi(\lambda)) < \infty \tag{2.54}$$

by the monotone convergence theorem. Hence $0 \leq \operatorname{Tr}_{\mathcal{H}}(\Xi(\lambda)) = ||\Xi(\lambda)||_1 < \infty$ for a.e. $\lambda \in \mathbb{R}$ completes the proof. □

Remark 2.11. For simplicity we focused on dissipative operators thus far. Later we will also encounter operators $S \in \mathcal{B}(\mathcal{H})$ with $-S$ dissipative, that is, $\operatorname{Im}(S) \leq 0$. In this case S^* is dissipative and one can simply define $\log(S)$ by

$$\log(S) = (\log(S^*))^*, \tag{2.55}$$

with $\log(S^*)$ defined as in (2.22). Moreover,

$$\log(\widehat{M}(z)) = \widehat{C} - \int_{\mathbb{R}} d\lambda\, \widehat{\Xi}(\lambda)((\lambda - z)^{-1} - \lambda(1+\lambda^2)^{-1}), \quad z \in \mathbb{C}_+, \tag{2.56}$$

$$\widehat{C} = \widehat{C}^* \in \mathcal{B}(\mathcal{H}) \text{ and } 0 \leq \widehat{\Xi}(\lambda) \leq I_{\mathcal{H}} \text{ for a.e. } \lambda \in \mathbb{R}, \tag{2.57}$$

whenever $\widehat{M}$ is analytic in $\mathbb{C}_+$ and $\operatorname{Im}(\widehat{M}(z)) \leq 0$, $z \in \mathbb{C}_+$.

Remark 2.12. Theorem 2.10 represents the operator-valued generalization of the exponential Herglotz representation for scalar Herglotz functions studied in detail by Aronszajn and Donoghue [4] (see also Carey and Pepe [19]). Theorem 2.10 is not the first attempt in this direction. Carey [18], in 1976, considered the case $M(z) = I_{\mathcal{H}} + K^*(H_0 - z)^{-1}K$ (i.e., $A - R^{1/2}LR^{1/2}|_{\mathcal{H}} = I_{\mathcal{H}}$, $B = 0$, $(1+L^2)^{1/2}R^{1/2}|_{\mathcal{H}} = K$ when compared to (2.1a)) and established

$$M(z) = \exp\left(\int_{\mathbb{R}} d\lambda\, B(\lambda)(\lambda - z)^{-1}\right) \tag{2.58}$$

for a summable operator function $B(\lambda)$, $0 \leq B(\lambda) \leq I_{\mathcal{H}}$ (i.e., the analog of $\Xi(\lambda)$ in (2.44)). Although Carey's proof also uses Naimark's dilation theorem as described in Theorem 1 of Appendix I of [16], it is different from ours and does not utilize the integral representation (2.22) for logarithms.

Remark 2.13. At first glance it may seem that we have been a bit pedantic in introducing various branches of the logarithm in (2.19) and (2.20). Actually, these branches are just a special case of the following family of branches

$$\log_\alpha(z) = \int_{\gamma_\alpha} d\zeta\, ((z-\zeta)^{-1} - (1-\zeta)^{-1}), \quad z \in \mathbb{C}\backslash\gamma_\alpha, \tag{2.59}$$

where γ_α denotes the ray, $\gamma_\alpha = \{\zeta \in \mathbb{C}_+ \mid \zeta = re^{i\alpha}, 0 \leq r < \infty, \alpha \in [\pi, 2\pi]\}$. In particular, $\log(\cdot) = \log_{3\pi/2}(\cdot)$ and since $\int_{-\infty}^0 d\lambda\, ((1-\lambda)^{-1} - \lambda(1+\lambda^2)^{-1}) = 0$, one infers $\ln(\cdot) = \log_\pi(\cdot)$. That some care has to be taken in connection with a consistent choice of branches especially for operator-valued branches of the logarithm is illustrated in the following Remark 2.14 and in Remark 3.3.

Remark 2.14. In view of our applications in Sections 3 and 4 it seems worthwhile to recall in connection with our hypothesis $\mathrm{Im}(\log(M(z_0))) \in \mathcal{B}_1(\mathcal{H})$ in Theorem 2.10 that if $A \in \mathcal{B}_1(\mathcal{H})$ and $\det_{\mathcal{H}}(I_{\mathcal{H}} + A) \neq 0$, then $\log(I_{\mathcal{H}} + A) \in \mathcal{B}_1(\mathcal{H})$ by (2.22) (using $(I_{\mathcal{H}} + A + i\lambda)^{-1} - (1 + i\lambda)^{-1}I_{\mathcal{H}} = (1+\lambda)^{-1}(I_{\mathcal{H}} + A + i\lambda)^{-1}A$.) Moreover, $\det_{\mathcal{H}}(I_{\mathcal{H}} + A) = \prod_{n\in\mathbb{N}}(1 + \lambda_n(A))$, where $\lambda_n(A)$ denote the eigenvalues of A repeated according to their algebraic multiplicity, then shows

$$\mathrm{Tr}_{\mathcal{H}}(\log(I_{\mathcal{H}} + A)) = \log(\det_{\mathcal{H}}(I_{\mathcal{H}} + A)). \tag{2.60}$$

In fact, (2.60) holds for any branch $\log_\alpha(\cdot)$ and hence, in particular, for the branch $\ln(\cdot)$ on either side of (2.60). (Here $\det_{\mathcal{H}}(\cdot)$ denotes the Fredholm determinant for operators in $\mathcal{H}$.)

3. The Spectral Shift Operator

The main purpose of this section is to introduce the concept of a spectral shift operator (cf. Definition 3.5) and a new approach to Krein's basic trace formula [51].

Let $\mathcal{H}$ be a complex separable Hilbert space and assume the following hypothesis for the remainder of this section.

Hypothesis 3.1. *Let H_0 be a self-adjoint operator in $\mathcal{H}$ with domain* $\mathrm{dom}(H_0)$*, J a bounded self-adjoint operator with $J^2 = I_{\mathcal{H}}$, and $K \in \mathcal{B}_2(\mathcal{H})$ a Hilbert-Schmidt operator.*

Introducing

$$V = KJK^* \tag{3.1}$$

we define the self-adjoint operator

$$H = H_0 + V, \quad \mathrm{dom}(H) = \mathrm{dom}(H_0) \tag{3.2}$$

in $\mathcal{H}$.

Given Hypothesis 3.1 we decompose $\mathcal{H}$ and J according to

$$J = \begin{pmatrix} I_+ & 0 \\ 0 & -I_- \end{pmatrix}, \quad \mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-, \tag{3.3}$$

$$J_+ = \begin{pmatrix} I_+ & 0 \\ 0 & 0 \end{pmatrix}, \quad J_- = \begin{pmatrix} 0 & 0 \\ 0 & I_- \end{pmatrix}, \quad J = J_+ - J_-, \tag{3.4}$$

with $I_\pm$ the identity operator in $\mathcal{H}_\pm$. Moreover, we introduce the following bounded operators

$$\Phi(z) = J + K^*(H_0 - z)^{-1}K : \mathcal{H} \to \mathcal{H}, \tag{3.5}$$

$$\Phi_+(z) = I_+ + J_+K^*(H_0 - z)^{-1}K|_{\mathcal{H}_+} : \mathcal{H}_+ \to \mathcal{H}_+, \tag{3.6}$$

$$\widetilde{\Phi}_-(z) = I_- - J_-K^*(H_+ - z)^{-1}K|_{\mathcal{H}_-} : \mathcal{H}_- \to \mathcal{H}_-, \tag{3.7}$$

for $z \in \mathbb{C}\backslash\mathbb{R}$, where

$$V_+ = KJ_+K^*, \tag{3.8}$$

$$H_+ = H_0 + V_+, \quad \mathrm{dom}(H_+) = \mathrm{dom}(H_0). \tag{3.9}$$

Lemma 3.2. *Assume Hypothesis* 3.1*. Then Φ, Φ_+, and $-\widetilde{\Phi}_-$ are Herglotz operators in $\mathcal{H}$, $\mathcal{H}_+$, and $\mathcal{H}_-$, respectively. In addition ($z \in \mathbb{C}\backslash\mathbb{R}$),*

$$\Phi(z)^{-1} = J - JK^*(H - z)^{-1}KJ, \tag{3.10}$$

$$\Phi_+(z)^{-1} = I_+ - J_+K^*(H_+ - z)^{-1}K|_{\mathcal{H}_+}, \tag{3.11}$$

$$\widetilde{\Phi}_-(z)^{-1} = I_- + J_-K^*(H - z)^{-1}K|_{\mathcal{H}_-}. \tag{3.12}$$

Proof. It suffices to consider Φ_+. Since I_+ and H_0 are self-adjoint, $\Phi_+(z) = (J_+ + J_+K^*(H_0 - z)^{-1}KJ_+)|_{\mathcal{H}_+}$ is clearly analytic in $\mathbb{C}\backslash\mathbb{R}$ and satisfies $\mathrm{Im}(\Phi_+(z)) \geq 0$ for $z \in \mathbb{C}_+$. Relation (3.11) is then an elementary consequence of the second resolvent equation,

$$(H_+ - z)^{-1} = (H_0 - z)^{-1} - (H_0 - z)^{-1}V_+(H_+ - z)^{-1} \tag{3.13a}$$

$$= (H_0 - z)^{-1} - (H_+ - z)^{-1}V_+(H_0 - z)^{-1}, \tag{3.13b}$$

the fact $J_+^2 = J_+$, and simply follows by multiplying the right-hand sides of (3.6) and (3.11) in either order. ∎

Remark 3.3. In the following we need to make use of the formula

$$d\,\mathrm{Tr}_{\mathcal{K}}(\log(I_{\mathcal{K}}+F(z)))/dz=\mathrm{Tr}_{\mathcal{K}}(F'(z)(I_{\mathcal{K}}+F(z))^{-1}). \tag{3.14}$$

This result is proven, for instance, in [28], Ch. I, Sect. 6.11 or [41], Sect. IV.1 for analytic $F(\cdot)\in\mathcal{B}_1(\mathcal{K})$ in some region $\Omega\subset\mathbb{C}$ such that $(I_{\mathcal{K}}+F(\cdot))^{-1}\in\mathcal{B}(\mathcal{K})$ in Ω using the standard branch $\ln(\cdot)$. In this case equation (3.14) is first proven in the finite-dimensional case, followed by taking the limit $n\to\infty$ upon replacing $F(z)$ by $P_nF(z)P_n$, with P_n a family of orthogonal projections in $\mathcal{K}$ strongly converging to $I_{\mathcal{K}}$ as $n\to\infty$ (cf., [41], p. 163). This strategy of proof extends to all branches $\log_\alpha(\cdot)$ introduced in Remark 2.13. More generally, we have the following result,

$$d\,\mathrm{Tr}_{\mathcal{K}}(\varphi(F(s)))/ds=\mathrm{Tr}_{\mathcal{K}}(\varphi'(F(s))F'(s)), \tag{3.15}$$

with $F(s)\in\mathcal{B}_1(\mathcal{K})$ defined on an interval $s_1\le s\le s_2$, continuously differentiable in $\mathcal{B}_1(\mathcal{K})$-norm, and $\varphi(z)$ any scalar function, holomorphic in some domain $\mathcal{D}\subset\mathbb{C}$ with a Jordan boundary and $\mathrm{spec}(F(s))\subset\mathcal{D}$ for all $s\in[s_1,s_2]$.

Lemma 3.4. *Assume Hypothesis* 3.1 *and* $\mathbb{C}\backslash\mathbb{R}$. *Then*

$$\mathrm{Tr}_{\mathcal{H}}((H_0-z)^{-1}-(H_+-z)^{-1})=d\,\mathrm{Tr}_{\mathcal{H}_+}(\log(\Phi_+(z)))/dz, \tag{3.16a}$$

$$\mathrm{Tr}_{\mathcal{H}}((H_+-z)^{-1}-(H-z)^{-1})=d\,\mathrm{Tr}_{\mathcal{H}_-}(\log(\widetilde{\Phi}_-(z)))/dz. \tag{3.16b}$$

Proof. It suffices to prove (3.16a). Applying (3.13) repeatedly one infers

$$\begin{aligned}
&\mathrm{Tr}_{\mathcal{H}}((H_0-z)^{-1}-(H_+-z)^{-1})\\
&=\mathrm{Tr}_{\mathcal{H}}((H_0-z)^{-1}V_+[(H_0-z)^{-1}-(H_+-z)^{-1}V_+(H_0-z)^{-1}])\\
&=\mathrm{Tr}_{\mathcal{H}}(J_+K^*(H_0-z)^{-2}KJ_+[J_+-J_+K^*(H_+-z)^{-1}KJ_+])\\
&=\mathrm{Tr}_{\mathcal{H}}(J_+\Phi_+'(z)J_+^2\Phi_+(z)^{-1}J_+)=\mathrm{Tr}_{\mathcal{H}_+}(\Phi_+'(z)\Phi_+(z)^{-1})\\
&=d\,\mathrm{Tr}_{\mathcal{H}_+}(\log(\Phi_+(z)))/dz,\quad z\in\mathbb{C}\backslash\mathbb{R},
\end{aligned} \tag{3.17}$$

where we used $J_+^2=J_+$,

$$\mathrm{Tr}_{\mathcal{H}}(AB)=\mathrm{Tr}_{\mathcal{H}}(BA) \tag{3.18}$$

for $A,B\in\mathcal{B}(\mathcal{H})$ with AB, $BA\in\mathcal{B}_1(\mathcal{H})$ (cf. Corollary 3.8 in [77]), and (3.14). □

Next, applying Theorem 2.10 and Remark 2.11 to $\Phi_+(z)$ and $\widetilde{\Phi}_-(z)$ one infers the existence of two families of bounded operators $\{\Xi_\pm(\lambda)\}_{\lambda\in\mathbb{R}}$ defined for (Lebesgue) a.e. $\lambda\in\mathbb{R}$ and satisfying

$$\begin{aligned}
&0\le\Xi_\pm(\lambda)\le I_\pm,\quad \Xi_\pm(\lambda)\in\mathcal{B}_1(\mathcal{H}_\pm)\text{ for a.e. }\lambda\in\mathbb{R},\\
&\|\Xi_\pm(\cdot)\|_1\in L^1(\mathbb{R};(1+\lambda^2)^{-1}d\lambda)
\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}
\log(\Phi_+(z))&=\log(I_++J_+K^*(H_0-z)^{-1}K|_{\mathcal{H}_+})\\
&=C_++\int_{\mathbb{R}}d\lambda\,\Xi_+(\lambda)((\lambda-z)^{-1}-\lambda(1+\lambda^2)^{-1}),
\end{aligned} \tag{3.20a}$$

$$\log(\widetilde{\Phi}_-(z)) = \log(I_- - J_- K^*(H_+ - z)^{-1} K|_{\mathcal{H}_-})$$
$$= C_- - \int_{\mathbb{R}} d\lambda\, \Xi_-(\lambda)((\lambda - z)^{-1} - \lambda(1+\lambda^2)^{-1}) \tag{3.20b}$$

for $z \in \mathbb{C}\backslash\mathbb{R}$, with $C_\pm = C_\pm^* \in \mathcal{B}_1(\mathcal{H})$.

Equations (3.20) motivate the following

Definition 3.5. $\Xi_+(\lambda)$ (resp., $\Xi_-(\lambda)$) is called the *spectral shift operator* associated with $\Phi_+(z)$ (resp., $\widetilde{\Phi}_-(z)$). Alternatively, we will refer to $\Xi_+(\lambda)$ as the spectral shift operator associated with the pair (H_0, H_+) and occasionally use the notation $\Xi_+(\lambda, H_0, H_+)$ to stress the dependence on (H_0, H_+), etc.

Moreover, we introduce

$$\xi_\pm(\lambda) = \mathrm{Tr}_{\mathcal{H}_\pm}(\Xi_\pm(\lambda)), \quad 0 \le \xi_\pm \in L^1(\mathbb{R}; (1+\lambda^2)^{-1} d\lambda) \text{ for a.e. } \lambda \in \mathbb{R}. \tag{3.21}$$

Actually, taking into account the simple behavior of $\Phi_+(iy)$ and $\widetilde{\Phi}_-(iy)$ as $|y| \to \infty$, one can improve (3.20a) and (3.20b) as follows.

Lemma 3.6. *Assume Hypothesis* 3.1 *and define* $\xi_\pm$ *as in* (3.21)*. Then*

$$0 \le \xi_\pm \in L^1(\mathbb{R}; d\lambda), \tag{3.22}$$

and (3.20a) *and* (3.20b) *simplify to*

$$\log(\Phi_+(z)) = \int_{\mathbb{R}} d\lambda\, \Xi_+(\lambda)(\lambda - z)^{-1}, \tag{3.23a}$$

$$\log(\widetilde{\Phi}_-(z)) = -\int_{\mathbb{R}} d\lambda\, \Xi_-(\lambda)(\lambda - z)^{-1}. \tag{3.23b}$$

Moreover, for a.e. $\lambda \in \mathbb{R}$,

$$\lim_{\varepsilon \downarrow 0} \|\Xi_+(\lambda) - \pi^{-1}\mathrm{Im}(\log(\Phi_+(\lambda + i\varepsilon)))\|_{\mathcal{B}_1(\mathcal{H}_+)} = 0, \tag{3.24a}$$

$$\lim_{\varepsilon \downarrow 0} \|\Xi_-(\lambda) + \pi^{-1}\mathrm{Im}(\log(\widetilde{\Phi}_-(\lambda + i\varepsilon)))\|_{\mathcal{B}_1(\mathcal{H}_-)} = 0. \tag{3.24b}$$

Proof. It suffices to consider $\xi_+(\lambda)$ and $\Phi_+(z)$. Since

$$\|\log(\Phi_+(y))\|_1 = O(|y|^{-1}) \text{ as } |y| \to \infty \tag{3.25}$$

by the Hilbert-Schmidt hypothesis on K and the fact $\|(H_0 - iy)^{-1}\| = O(|y|^{-1})$ as $|y| \to \infty$, the scalar Herglotz function $\mathrm{Tr}_{\mathcal{H}_+}(\log(\Phi_+(z)))$ satisfies

$$|\,\mathrm{Tr}_{\mathcal{H}_+}(\log(\Phi_+(z)))| = O(|y|^{-1}) \text{ as } |y| \to \infty. \tag{3.26}$$

By standard results (see, e.g., [4], [45]), (3.26) yields

$$\mathrm{Tr}_{\mathcal{H}_+}(\log(\Phi_+(z))) = \int_{\mathbb{R}} d\omega_+(\lambda)(\lambda - z)^{-1}, \quad z \in \mathbb{C}\backslash\mathbb{R}, \tag{3.27}$$

where ω_+ is a finite measure,

$$\int_{\mathbb{R}} d\omega_+(\lambda) = -i \lim_{y \uparrow \infty} (y\, \mathrm{Tr}_{\mathcal{H}_+}(\log(\Phi_+(z)))) < \infty. \tag{3.28}$$

Moreover, the fact that $\operatorname{Im}(\log(\Phi_+(z)))$ is uniformly bounded with respect to $z \in \mathbb{C}_+$ yields that ω_+ is purely absolutely continuous,

$$d\omega_+(\lambda) = \xi_+(\lambda)d\lambda, \quad \xi_+ \in L^1(\mathbb{R}; d\lambda), \tag{3.29}$$

where

$$\begin{aligned}\xi_+(\lambda) &= \pi^{-1} \lim_{\varepsilon\downarrow 0}(\operatorname{Im}(\operatorname{Tr}_{\mathcal{H}_+}(\log(\Phi_+(\lambda+i\varepsilon))))) = \operatorname{Tr}_{\mathcal{H}_+}(\Xi_+(\lambda)) \\ &= \pi^{-1} \lim_{\varepsilon\downarrow 0}(\operatorname{Im}(\log(\det{}_{\mathcal{H}_+}(\Phi_+(\lambda+i\varepsilon))))) \text{ for a.e. } \lambda \in \mathbb{R}. \end{aligned}\tag{3.30}$$

In order to prove (3.24a) we first observe that $\operatorname{Im}(\log(\Phi_+(\lambda+i\varepsilon)))$ takes on boundary values $\operatorname{Im}(\log(\Phi_+(\lambda+i0)))$ for a.e. $\lambda \in \mathbb{R}$ in $\mathcal{B}_1(\mathcal{H}_+)$-norm by (2.2). Next, choosing an orthonormal system $\{e_n\}_{n\in\mathbb{N}} \subset \mathcal{H}_+$, we recall that the quadratic form $(e_n, \operatorname{Im}(\log(\Phi(\lambda+i0)))e_n)_{\mathcal{H}_+}$ exists for all $\lambda \in \mathbb{R}\backslash\mathcal{E}_n$, where $\mathcal{E}_n$ has Lebesgue measure zero. Thus one observes,

$$\begin{aligned}&\lim_{\varepsilon\downarrow 0}(e_m, \operatorname{Im}(\log(\Phi_+(\lambda+i\varepsilon)))e_n)_{\mathcal{H}_+} = (e_m, \operatorname{Im}(\log(\Phi_+(\lambda+i0)))e_n)_{\mathcal{H}_+} \\ &= \pi(e_m, \Xi_+(\lambda)e_n)_{\mathcal{H}_+} \text{ for } \lambda \in \mathbb{R}\backslash\{\mathcal{E}_m \cup \mathcal{E}_n\}.\end{aligned}\tag{3.31}$$

Let $\mathcal{E} = \cup_{n\in\mathbb{N}}\mathcal{E}_n$, then $|\mathcal{E}| = 0$ ($|\cdot|$ denoting the Lebesgue measure on $\mathbb{R}$) and hence

$$\begin{aligned}&(f, \operatorname{Im}(\log(\Phi_+(\lambda+i0)))g)_{\mathcal{H}_+} = \pi(f, \Xi_+(\lambda)g)_{\mathcal{H}_+} \\ &\text{for } \lambda \in \mathbb{R}\backslash\mathcal{E} \text{ and } f, g \in \mathcal{D} = \text{lin.span}\{e_n \in \mathcal{H}_+ \,|\, n \in \mathbb{N}\}.\end{aligned}\tag{3.32}$$

Since $\mathcal{D}$ is dense in $\mathcal{H}_+$ and $\Xi_+(\lambda) \in \mathcal{B}(\mathcal{H}_+)$ one infers $\operatorname{Im}(\log(\Phi_+(\lambda+i0))) = \pi\Xi_+(\lambda)$ for a.e. $\lambda \in \mathbb{R}$, completing the proof. $\square$

Of course (3.24a), (3.24b) (and the method of proof) immediately extend to a.e. nontangential limits to the real axis.

Assuming Hypothesis 3.1 we define

$$\xi(\lambda) = \xi_+(\lambda) - \xi_-(\lambda) \text{ for a.e. } \lambda \in \mathbb{R} \tag{3.33}$$

and call $\xi(\lambda)$ (resp., $\xi_+(\lambda)$, $\xi_-(\lambda)$) the spectral shift function associated with the pair (H_0, H) (resp., (H_0, H_+), (H_+, H)), sometimes also denoted by $\xi(\lambda, H_0, H)$, etc., to underscore the dependence on the pair involved.

M. Krein's basic trace formula [51] is now obtained as follows.

Theorem 3.7. *Assume Hypothesis* 3.1. *Then* ($z \in \mathbb{C}\backslash\{\operatorname{spec}(H_0) \cup \operatorname{spec}(H)\}$)

$$\operatorname{Tr}_{\mathcal{H}}((H-z)^{-1} - (H_0-z)^{-1}) = -\int_{\mathbb{R}} d\lambda\, \xi(\lambda)(\lambda-z)^{-2}. \tag{3.34}$$

Proof. Let $z \in \mathbb{C}\backslash\mathbb{R}$. By (3.27) and (3.29) we infer

$$\operatorname{Tr}_{\mathcal{H}_+}(\log(\Phi_+(z))) = \int_{\mathbb{R}} d\lambda\, \xi_+(\lambda)(\lambda-z)^{-1}, \tag{3.35}$$

$$\operatorname{Tr}_{\mathcal{H}_-}(\log(\widetilde{\Phi}_-(z))) = -\int_{\mathbb{R}} d\lambda\, \xi_-(\lambda)(\lambda-z)^{-1}. \tag{3.36}$$

Adding (3.16a) and (3.16b), differentiating (3.35) and (3.36) with respect to z proves (3.34) for $z \in \mathbb{C}\backslash\mathbb{R}$. The result extends to all $z \in \mathbb{C}\backslash\{\mathrm{spec}(H_0) \cup \mathrm{spec}(H)\}$ by continuity of $((H-z)^{-1} - (H_0-z)^{-1})$ in $\mathcal{B}_1(\mathcal{H})$-norm. $\square$

In particular, $\xi(\lambda)$ introduced in (3.33) is Krein's original spectral shift function (up to normalization). As noted in Section 2, the spectral shift operator $\Xi_+(\lambda)$ in the particular case $V = V_+$, and its relation to Krein's spectral shift function $\xi_+(\lambda)$, was first studied by Carey [18] in 1976.

Remark 3.8. (i) As shown originally by M. Krein [51], the trace formula (3.34) extends to

$$\mathrm{Tr}(f(H) - f(H_0)) = \int_{\mathbb{R}} d\lambda\, \xi(\lambda) f'(\lambda) \tag{3.37}$$

for appropriate functions f. This fact has been studied by numerous authors and we refer, for instance, to [5], Ch. 19, [10]–[12], [52], [53], [65], [82], [83], [87], [88], Ch. 8 and the references therein.
(ii) While we focus here on pairs of self-adjoint operators (H_0, H), M. Krein in his original paper [51] also considered pairs of unitary operators. Given the conformal equivalence of $\mathbb{C}_+$ and the open interior of the unit disk D in $\mathbb{C}$, this corresponds precisely to the study of Nevanlinna, respectively, Riesz-Herglotz functions in $\mathbb{C}_+$, respectively, D. Pairs of unitary operators are also studied, for instance, in [11], [12], [58], [82], [83]. The trace formula in the case of non-self-adjoint and non-unitary pairs of operators is a rather challenging problem. The interested reader can get some insight into this matter by consulting [1], [53], [55], [74] and the extensive literature cited therein. Similarly, the case of non-trace class perturbations and associated trace formulas has been studied by various authors. We refer, for instance, to [75], where Hilbert-Schmidt perturbations of unitary operators are treated in depth.
(iii) While scattering theory for the pair (H_0, H) is not discussed in this paper, we remark that $\xi(\lambda)$, for a.e. $\lambda \in \mathrm{spec}_{ac}(H_0)$ (the absolutely continuous spectrum of H_0), is related to the scattering operator at fixed energy λ by the Birman-Krein formula [9],

$$\det_{\mathcal{H}_\lambda}(S(\lambda, H_0, H)) = e^{-2\pi i \xi(\lambda)} \text{ for a.e. } \lambda \in \mathrm{spec}_{ac}(H_0). \tag{3.38}$$

Here $S(\lambda, H_0, H)$ denote the fibers in the direct integral representation of the scattering operator

$$S(H_0, H) = \int_{\mathrm{spec}_{ac}(H_0)}^{\oplus} d\lambda\, S(\lambda, H_0, H) \text{ in } \mathcal{H} = \int_{\mathrm{spec}_{ac}(H_0)}^{\oplus} d\lambda\, \mathcal{H}_\lambda$$

with respect to the absolutely continuous part $H_{0,ac}$ of H_0. This fundamental connection, originally due to Birman and Krein [9], is further discussed in [5], Ch. 19, [12], [13], [18], [46], [52], [84], [88], Ch. 8 and the literature cited therein.
(iv) The standard identity ([41], Sect. IV.3)

$$\mathrm{Tr}_{\mathcal{H}}((H-z)^{-1} - (H_0-z)^{-1}) = -d \log(\det_{\mathcal{H}}(I_{\mathcal{H}} + V(H_0-z)^{-1}))/dz \tag{3.39}$$

together with the trace formula (3.34) yields the well-known connection between perturbation determinants and $\xi(\lambda)$, also due to M. Krein [51]

$$\log(\det_{\mathcal{H}}(I_{\mathcal{H}} + V(H_0 - z)^{-1})) = \int_{\mathbb{R}} d\lambda\, \xi(\lambda)(\lambda - z)^{-1}, \tag{3.40}$$

$$\xi(\lambda) = \lim_{\varepsilon\downarrow 0} \pi^{-1}\mathrm{Im}(\log(\det_{\mathcal{H}}(I_{\mathcal{H}} + V(H_0 - (\lambda + i0))^{-1}))) \text{ for a.e. } \lambda \in \mathbb{R}, \tag{3.41}$$

$$\mathrm{Tr}_{\mathcal{H}}(V) = \int_{\mathbb{R}} d\lambda\, \xi(\lambda), \qquad \int_{\mathbb{R}} d\lambda\, |\xi(\lambda)| \leq ||V||_1. \tag{3.42}$$

This is discussed in more detail, for instance, in [5], Ch. 19, [12], [18], [50], [52], [54], [82]–[84], [88]. Relation (3.41) and the analog of (2.2) for $d(AE_H(\lambda)B)/d\lambda$, where $H = H_0 + V$, $V = B^*A$, $A, B \in \mathcal{B}_2(\mathcal{H})$, $V = V^*$, leads to the expression

$$\xi(\lambda) = (-2\pi i)^{-1}\, \mathrm{Tr}_{\mathcal{H}}(\log(I_{\mathcal{H}} - 2\pi i(I_{\mathcal{H}} - A(H - \lambda - i0)^{-1}B^*)(d(AE_H(\lambda)B^*)/d\lambda)))$$

for a.e. $\lambda \in \mathbb{R}$ (cf., e.g., [5], Sects. 3.4.4 and 19.1.4).
(v) The invariance principle for wave operators of the pair (H_0, H) implies

$$\xi(\lambda) = \xi(\lambda, H_0, H) = \mathrm{sgn}(\Psi'(\lambda))\xi(\Psi(\lambda), \Psi(H_0), \Psi(H)) \tag{3.43}$$

for a certain class of admissible functions Ψ. In certain cases (e.g., if H_0 and H are bounded from below) this can be used to define ξ even though $H - H_0 = V$ is not of trace class as long as $(\Psi(H) - \Psi(H_0)) \in \mathcal{B}_1(\mathcal{H})$. Prime candidates for Ψ in such cases are semigroup ($\Psi(\lambda) = e^{-t\lambda}, t > 0$) and resolvent ($\Psi(\lambda) = (\lambda - z)^{-1}$, $z \in \mathbb{C}\backslash\mathbb{R}$ or $z < E_0$ for some $E_0 \in \mathbb{R}$) functions. Pertinent facts in this connection can be found in [44], [83], and [88], Sect. 8.11.
(vi) For simplicity we chose a single Hilbert space formulation throughout this section. However, every result immediately extends to the case where $K \in \mathcal{B}_2(\mathcal{K}, \mathcal{H})$, $J = J^* \in \mathcal{B}(\mathcal{K})$, $J^2 = I_{\mathcal{K}}$ and $\mathcal{K}$ is another complex separable Hilbert space.

Remark 3.9. Suppose H_0, H_1, and H_2 are self-adjoint operators in $\mathcal{H}$ with $(H_j - H_k) \in \mathcal{B}_1(\mathcal{H})$ for all $j, k \in \{0, 1, 2\}$. Denoting by $\xi(\lambda, H_j, H_k)$ the Krein spectral shift function of the pair (H_j, H_k) such that

$$\mathrm{Tr}_{\mathcal{H}}((H_k - z)^{-1} - (H_j - z)^{-1}) = -\int_{\mathbb{R}} d\lambda\, \xi(\lambda, H_j, H_k)(\lambda - z)^{-2}, \tag{3.44}$$

the chain rule,

$$\xi(\lambda, H_0, H_2) = \xi(\lambda, H_0, H_1) + \xi(\lambda, H_1, H_2) \text{ a.e.}, \tag{3.45}$$

together with

$$\xi(\lambda, H_j, H_k) = -\xi(\lambda, H_k, H_j) \text{ a.e.}, \tag{3.46}$$

and

$$\xi(\lambda, H_j, H_k) \geq 0 \text{ a.e. if } (H_k - H_j) \geq 0, \tag{3.47}$$

imply the monotonicity property

$$\xi(\lambda, H_0, H_2) \geq \xi(\lambda, H_0, H_1) \text{ a.e. if } H_2 \geq H_1. \tag{3.48}$$

Here (3.47) is clear from (3.19), (3.21). Equation (3.46) follows from (3.44), and (3.45) is a consequence of (3.41) observing the facts

$$I_{\mathcal{H}} + V(H_0 - z)^{-1} = (H_0 + V - z)(H_0 - z)^{-1}, \quad V \in \mathcal{B}(\mathcal{H}),$$
$$\det_{\mathcal{H}}((I_{\mathcal{H}} + A)(I_{\mathcal{H}} + B)) = \det_{\mathcal{H}}(I_{\mathcal{H}} + A)\det_{\mathcal{H}}(I_{\mathcal{H}} + B), \quad A, B \in \mathcal{B}_1(\mathcal{H}).$$

Given the monotonicity property (3.48) of Krein's spectral shift function, it is natural to inquire whether or not this property is shared by the spectral shift operator. More precisely, one might ask whether or not

$$\Xi(\lambda, H_0, H_2) \overset{?}{\geq} \Xi(\lambda, H_0, H_1) \text{ a.e. if } H_2 \geq H_1 \geq H_0. \tag{3.49}$$

The following simple counter example destroys such hopes.

Example 3.10. Let $H_0 = 0$, $K \in \mathcal{B}_2(\mathcal{H})$, $J = I_{\mathcal{H}}$, and hence $\Phi(z) = I_{\mathcal{H}} - K^*Kz^{-1}$, $z \in \mathbb{C}\backslash\{0\}$. Then

$$\lim_{\varepsilon\downarrow 0} \| \log(I_{\mathcal{H}} - K^*K(\lambda + i\varepsilon)^{-1}) - \log(I_{\mathcal{H}} - K^*K\lambda^{-1})\|_{\mathcal{B}_1(\mathcal{H})} = 0 \tag{3.50}$$

for $\lambda \in \mathbb{R}\backslash\{\mathrm{spec}(K^*K) \cup \{0\}\}$ and

$$\pi^{-1}\mathrm{Im}(\log(I_{\mathcal{H}} - K^*K\lambda^{-1})) = \Xi(\lambda), \quad \lambda \in \mathbb{R}\backslash\{\mathrm{spec}(K^*K) \cup \{0\}\}. \tag{3.51}$$

Decomposing the self-adjoint operator $I_{\mathcal{H}} - K^*K\lambda^{-1}$, $\lambda \in \mathbb{R}\backslash\{\mathrm{spec}(K^*K) \cup \{0\}\}$ into its positive and negative spectral parts then yields $\Xi(\lambda) = \theta(K^*K - \lambda I_{\mathcal{H}})$, $\lambda \in \mathbb{R}\backslash\{\mathrm{spec}(K^*K) \cup \{0\}\}$, where $\theta(\cdot)$ denotes the usual step function (i.e., $\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ for $x < 0$).

Next we choose $\mathcal{H} = \mathbb{C}^2$, $0 < a < b < c < 1$, $ac - b^2 \geq 0$, and

$$K_1^*K_1 = \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix}, \quad \mathrm{spec}(K_1^*K_1) = \{1 - b, 1 + b\}, \tag{3.52}$$

with eigenvectors $2^{-1/2}(1, \pm 1)^t$ associated to the eigenvalues $1 \pm b$ and

$$K_2^*K_2 = \begin{pmatrix} 1 + a & 0 \\ 0 & 1 + c \end{pmatrix}, \quad \mathrm{spec}(K_2^*K_2) = \{1 + a, 1 + c\}, \tag{3.53}$$

with eigenvectors $(1, 0)^t$ (resp., $(0, 1)^t$) associated to the eigenvalue $1 + a$ (resp., $1 + c$). Finally, choosing $\lambda \in (1 + a, 1 + b)$ then yields

$$\Xi_1(\lambda) = \theta(K_1^*K_1 - \lambda I_{\mathbb{C}^2}) = E_{K_1^*K_1}(\{1 + b\}), \tag{3.54}$$
$$\Xi_2(\lambda) = \theta(K_2^*K_2 - \lambda I_{\mathbb{C}^2}) = E_{K_2^*K_2}(\{1 + c\}), \tag{3.55}$$

where $\{E_{K^*K}(\lambda)\}_{\lambda\in\mathbb{R}}$ denotes the family of orthogonal spectral projections of K^*K. Clearly $H_2 = K_2^*K_2 \geq K_1^*K_1 = H_1 \geq H_0 = 0$ but $\Xi_2(\lambda) \not\geq \Xi_1(\lambda)$ for $\lambda \in (1 + a, 1 + b)$ since one-dimensional self-adjoint projections cannot satisfy an order relation unless one is a real multiple of the other. (Note, however, that $\mathrm{Tr}(\Xi_2(\lambda))= 1 + c > 1 + b = \mathrm{Tr}(\Xi_1(\lambda))$ in accordance with (3.48).)

This example shows, in particular, that the chain rule (3.45) for $\xi(\lambda)$ does not extend to $\Xi(\lambda)$.

4. Spectral Averaging: An Operator-Theoretic Approach

In this section we apply the formalism developed in Sections 2 and 3 to provide an effortless proof of spectral averaging and its relation to Krein's spectral shift function as originally proven by Birman and Solomyak [11].

For the basic setup of this section we assume the following hypothesis.

Hypothesis 4.1. *Let H_0 be a self-adjoint operator in $\mathcal{H}$ with $\mathrm{dom}(H_0)$, and assume $\{V(s)\}_{s\in\Omega} \subset \mathcal{B}_1(\mathcal{H})$ to be a family of self-adjoint trace class operators in $\mathcal{H}$, where $\Omega \subseteq \mathbb{R}$ denotes an open interval. Moreover, suppose that $V(s)$ is continuously differentiable with respect to $s \in \Omega$ in trace norm.*

To begin our discussions we temporarily assume that $V(s) \geq 0$, that is, we suppose

$$V(s) = K(s)K(s)^*, \quad s \in \Omega \tag{4.1}$$

for some $K(s) \in \mathcal{B}_2(\mathcal{H})$, $s \in \Omega$. Given Hypothesis 4.1 we define the self-adjoint operator $H(s)$ in $\mathcal{H}$ by

$$H(s) = H_0 + V(s), \quad \mathrm{dom}(H(s)) = \mathrm{dom}(H_0), \quad s \in \Omega. \tag{4.2}$$

In analogy to (3.5) and (3.6) we introduce in $\mathcal{H}$ ($s \in \Omega$, $z \in \mathbb{C}\backslash\mathbb{R}$),

$$\Phi(z,s) = I_{\mathcal{H}} + K(s)^*(H_0 - z)^{-1}K(s) \tag{4.3}$$

and hence infer from Lemma 3.2 that

$$\Phi(z,s)^{-1} = I_{\mathcal{H}} - K(s)^*(H(s) - z)^{-1}K(s). \tag{4.4}$$

The following is an elementary but useful result needed in the context of Theorem 4.3.

Lemma 4.2. *Assume Hypothesis* 4.1 *and* (4.1)*. Then* ($s \in \Omega$, $z \in \mathbb{C}\backslash\mathbb{R}$),

$$d\,\mathrm{Tr}_{\mathcal{H}}(\log(\Phi(z,s)))/ds = \mathrm{Tr}_{\mathcal{H}}(V'(s)(H(s) - z)^{-1}). \tag{4.5}$$

Proof. By (3.15), (4.3), and (4.4) one infers for $z = iy$, $|y| > 0$ sufficiently large,

$$\begin{aligned}
&d\,\mathrm{Tr}_{\mathcal{H}}(\log(\Phi(z,s)))/ds = d\,\mathrm{Tr}_{\mathcal{H}}(\log(I_{\mathcal{H}} - V(s)(H_0 - z)^{-1}))/ds \\
&= (d/ds)\sum_{j=0}^{\infty}(-1)^j(j+1)^{-1}\,\mathrm{Tr}_{\mathcal{H}}((V(s)(H_0 - z)^{-1})^{j+1}) \\
&= \mathrm{Tr}_{\mathcal{H}}(V'(s)((H_0 - z)^{-1}\sum_{j=0}^{\infty}(-1)^j(V(s)(H_0 - z)^{-1})^j) \\
&= \mathrm{Tr}_{\mathcal{H}}(V'(s)((H(s) - z)^{-1})
\end{aligned} \tag{4.6}$$

by repeated use of (3.18) and (3.13a). Analytic continuation of (4.6) with respect to $z \in \mathbb{C}\backslash\mathbb{R}$ then proves (4.5). ∎

Next, applying Lemma 3.6 to $\Phi(z,s)$ in (4.3) one infers ($s \in \Omega$),

$$\log(\Phi(z,s)) = \int d\lambda\, \Xi(\lambda,s)(\lambda - z)^{-1}, \tag{4.7}$$

$$0 \le \Xi(\lambda,s) \le I_{\mathcal{H}}, \quad \Xi(\lambda,s) \in \mathcal{B}_1(\mathcal{H}) \text{ for a.e. } \lambda \in \mathbb{R}, \tag{4.8}$$

$$||\Xi(\cdot,s)||_1 \in L^1(\mathbb{R}; d\lambda),$$

where $\Xi(\lambda,s)$ is associated with the pair $(H_0, H(s))$, assuming $H(s) \ge H_0$, $s \in \Omega$.

Our principle result on averaging the spectral measure of $\{E_{H(s)}(\lambda)\}_{\lambda\in\mathbb{R}}$ of $H(s)$ then reads as follows.

Theorem 4.3. *Assume Hypothesis* 4.1 *and* $[s_1, s_2] \subset \Omega$. *Let* $\xi(\lambda, s)$ *be the spectral shift function associated with the pair* $(H_0, H(s))$ (*cf.* (3.33)), *where* $H(s)$ *is defined by* (4.2) (*and we no longer suppose* $H(s) \ge H_0$). *Then*

$$\int_{s1}^{s_2} ds\, (d(\mathrm{Tr}_{\mathcal{H}}(V'(s)E_{H(s)}(\lambda)))) = (\xi(\lambda,s_2) - \xi(\lambda,s_1))d\lambda. \tag{4.9}$$

Proof. First we prove (4.9) in the case $V(s) \ge 0$. The monotone convergence theorem, (4.7), and Lemma 4.2 then yield ($z \in \mathbb{C}\backslash\mathbb{R}$),

$$\begin{aligned}
&\mathrm{Tr}_{\mathcal{H}}\left(\int_{\mathbb{R}} d\lambda\, ((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1}\mathrm{Im}(z)(\Xi(\lambda,s_2) - \Xi(\lambda,s_1))\right)\\
&= \int_{\mathbb{R}} d\lambda\, ((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1}\mathrm{Im}(z)(\xi(\lambda,s_2) - \xi(\lambda,s_1))\\
&= \mathrm{Tr}_{\mathcal{H}}(\mathrm{Im}(\log(\Phi(z,s_2)))) - \mathrm{Tr}_{\mathcal{H}}(\mathrm{Im}(\log(\Phi(z,s_1))))\\
&= \int_{s_1}^{s_2} ds\left(\frac{d}{ds}\mathrm{Tr}_{\mathcal{H}}(\mathrm{Im}(\log(\Phi(z,s))))\right) = \int_{s_1}^{s_2} ds\, \mathrm{Tr}_{\mathcal{H}}(V'(s)\mathrm{Im}((H(s)-z)^{-1})).
\end{aligned}$$

By the spectral theorem applied to $H(s)$ one obtains

$$\mathrm{Im}((H(s)-z)^{-1}) = \mathrm{Im}(z)\int_{\mathbb{R}} dE_{H(s)}(\lambda)((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1} \tag{4.10}$$

and hence

$$\begin{aligned}
&\int_{\mathbb{R}} d\lambda\, ((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1}(\xi(\lambda,s_2) - \xi(\lambda,s_1))\\
&= \int_{s_1}^{s_2} ds\, \mathrm{Tr}_{\mathcal{H}}\left(\int_{\mathbb{R}} dE_{H(s)}(\lambda))((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1}V'(s)\right).
\end{aligned} \tag{4.11}$$

Decomposing the self-adjoint trace class operator $V'(s)$ into its positive and negative parts,

$$V'(s) = (V'(s))_+ - (V'(s))_-, \quad 0 \le (V'(s))_\pm \in \mathcal{B}_1(\mathcal{H}), \tag{4.12}$$

the monotone convergence theorem yields

$$\int_{\mathbb{R}} d\lambda\, ((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1}(\xi(\lambda,s_2) - \xi(\lambda,s_1))$$

$$
\begin{aligned}
&= \int_{s_1}^{s_2} ds \int_{\mathbb{R}} ((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1} (d(\mathrm{Tr}_{\mathcal{H}}((V'(s))_+^{1/2} E_{H(s)}(\lambda)(V'(s))_+^{1/2}) \\
&\quad - d(\mathrm{Tr}_{\mathcal{H}}((V'(s))_-^{1/2} E_{H(s)}(\lambda)(V'(s))_-^{1/2}))) \\
&= \int_{s_1}^{s_2} ds \int_{\mathbb{R}} ((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1} d(\mathrm{Tr}_{\mathcal{H}}(V'(s)E_{H(s)}(\lambda))) \\
&= \int_{\mathbb{R}} ((\lambda - \mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2)^{-1} \int_{s_1}^{s_2} ds(d(\mathrm{Tr}_{\mathcal{H}}(V'(s)E_{H(s)}(\lambda)))) \qquad (4.13)
\end{aligned}
$$

using Fubini's theorem in the last step. Thus, by the uniqueness property of Poisson transforms,

$$
(\xi(\lambda, s_2) - \xi(\lambda, s_1))d\lambda = \int_{s_1}^{s_2} ds\,(d(\mathrm{Tr}_{\mathcal{H}}(V'(s)E_{H(s)}(\lambda)))). \qquad (4.14)
$$

In the case of arbitrary $V(s) \in \mathcal{B}_1(\mathcal{H})$ (not necessarily nonnegative), we argue as follows. Define (cf. (4.12))

$$
W = \left(\int_{s_1}^{s_2} ds\,(V'(s))_- \right) - V(s_1), \qquad (4.15)
$$

then $W \in \mathcal{B}_1(\mathcal{H})$ and $V(s) + W \geq 0$ for all $s \in [s_1, s_2]$. Equation (4.9) now follows from the chain rule (3.45) for spectral shift functions

$$
\xi(\lambda, H_0, H(s)) = \xi(\lambda, H_0, H_0 - W) + \xi(\lambda, H_0 - W, H(s)). \qquad (4.16)
$$

Indeed, $(V(s) + W)' = V'(s)$, and $\xi(\lambda, H_0, H_0 - W)$ is independent of s and hence drops out on the right-hand side of (4.9). Moreover, the pair $(H_0 - W, H(s))$ only involves the nonnegative perturbation $V(s) + W$ of $H_0 - W$ in $H(s) = H_0 - W + (V(s) + W)$ so that Lemma 4.2 becomes applicable as in (4.14) in the first part of our proof. □

Remark 4.4. (i) In the special case of averaging over the boundary condition parameter for half-line Sturm-Liouville operators (effectively a rank-one resolvent perturbation problem), Theorem 4.3 has first been derived by Javrjan [42], [43]. The case of rank-one perturbations was recently treated in detail by Simon [79]. The general case of trace class perturbations is due to Birman and Solomyak [11] using an approach of Stieltjes' double operator integrals. Birman and Solomyak treat the case $V(s) = sV$, $V \in \mathcal{B}_1(\mathcal{H})$, $s \in [0, 1]$. As explained in [6], [11], and [12], the authors were interested in a real analysis approach to the spectral shift function in contrast to M. Krein's complex analytic treatment. In this way the local integrability of ξ could not be obtained directly (only its property of being a generalized function was obtained) although it follows of course from the uniqueness ξ (up to additive constants). While our operator theoretic approach is intrinsically complex-analytic, and hence very much in M. Krein's spirit, it leads to a natural proof of the absolute continuity of the (signed) measure on the right-hand side of (4.9). A short proof of (4.9) (assuming $V'(s) \geq 0$) has recently been given by Simon [80].

(ii) We note that variants of (4.9) in the context of one-dimensional Sturm-Liouville operators (i.e., variants of Javrjan's results in [42], [43]) have been repeatedly rediscovered by several authors. In particular, the absolute continuity of averaged spectral measures (with respect to boundary condition parameters or coupling constants of rank-one perturbations) has been used to prove localization properties of one-dimensional random Schrödinger operators (see, e.g., [17], [20], [21], [22], Ch. VIII, [24], [25], [47]–[49], [64], Ch. V, [78], [79]).
(iii) We emphasize that Theorem 4.3 applies to unbounded operators (and hence to random Schrödinger operators bounded from below) as long as appropriate relative trace class conditions (either with respect to resolvent or semigroup perturbations) are satisfied.
(iv) In the special case $V'(s) \geq 0$, the measure

$$d(\mathrm{Tr}_{\mathcal{H}}(V'(s)E_{H(s)}(\lambda))) = d(\mathrm{Tr}_{\mathcal{H}}(V'(s)^{1/2}E_{H(s)}(\lambda)V'(s)^{1/2}))$$

in (4.9) represents a positive measure.
(v) The result (4.9) is not restricted to a one-dimensional parameter space $s \in [s_1, s_2]$. In fact, if $\gamma(\mathbf{s_1}, \mathbf{s_2})$ denotes an oriented piecewise C^1-path connecting $\mathbf{s_1} \in \mathbb{R}^n$ and $\mathbf{s_2} \in \mathbb{R}^n$, one obtains analogously,

$$\int_{\gamma(\mathbf{s_1},\mathbf{s_2})} d\mathbf{s} \cdot (d(\mathrm{Tr}_{\mathcal{H}}((\nabla V)(\mathbf{s})E_{H(s)}(\lambda)))) = (\xi(\lambda, \mathbf{s_2}) - \xi(\lambda, \mathbf{s_1}))d\lambda. \tag{4.17}$$

We omit further details.

In the special case of a sign-definite perturbation of H_0 of the form sKK^*, one can in fact prove an operator-valued averaging formula as follows.

Theorem 4.5. *Assume Hypothesis* 3.1 *and* $J = I_{\mathcal{H}}$. *Then*

$$\int_0^1 ds\, d(K^* E_{H_0+sKK^*}(\lambda)K) = \Xi(\lambda)d\lambda, \tag{4.18}$$

where $\Xi(\cdot)$ *is the spectral shift operator associated with*

$$\Phi(z) = I_{\mathcal{H}} + K^*(H_0 - z)^{-1}K, \quad z \in \mathbb{C}\backslash\mathbb{R}, \tag{4.19}$$

that is,

$$\log(\Phi(z)) = \int_{\mathbb{R}} d\lambda\, \Xi(\lambda)(\lambda - z)^{-1}, \quad z \in \mathbb{C}\backslash\mathbb{R}, \tag{4.20}$$

$$0 \leq \Xi(\lambda) \in \mathcal{B}_1(\mathcal{H}) \text{ for a.e. } \lambda \in \mathbb{R}, \quad \|\Xi(\cdot)\|_1 \in L^1(\mathbb{R}; d\lambda). \tag{4.21}$$

Proof. An explicit computation shows

$$\begin{aligned}(\lambda - \Phi(z))^{-1} &= -(1-\lambda)^{-1} \\ &\times (I_{\mathcal{H}} - (1-\lambda)^{-1}K^*(H_0 + (1-\lambda)^{-1}KK^* - z)^{-1}K) \in \mathcal{B}(\mathcal{H})\end{aligned} \tag{4.22}$$

for all $\lambda < 0$. Since $\log(\Phi(z)) = \ln(\Phi(z))$ for $z \in \mathbb{C}_+$ as a result of analytic continuation, one obtains

$$\log(\Phi(z)) = \int_{\mathbb{R}} d\lambda\, \Xi(\lambda)(\lambda - z)^{-1}$$

$$
\begin{aligned}
= \ln(\Phi(z)) &= \int_{-\infty}^{0} d\lambda \, ((\lambda - \Phi(z))^{-1} - \lambda(1+\lambda^2)^{-1} I_{\mathcal{H}}) \\
&= \int_{-\infty}^{0} d\lambda \, (1-\lambda)^{-2} K^*(H_0 + (1-\lambda)^{-1} KK^* - z)^{-1} K \\
&= \int_0^1 ds \int_{\mathbb{R}} d(K^* E_{H_0 + sKK^*}(\lambda) K)(\lambda - z)^{-1} \\
&= \int_{\mathbb{R}} (\lambda - z)^{-1} \int_0^1 ds \, d(K^* E_{H_0 + sKK^*}(\lambda) K) \qquad (4.23)
\end{aligned}
$$

proving (4.18). (Here the interchange of the λ and s integrals follows from Fubini's theorem considering (4.23) in the weak sense.) $\square$

As a consequence of Theorem 4.5 one obtains

$$
\int_{s_1}^{s_2} ds \, d(K^* E_{H_0 + sKK^*}(\lambda) K) = \Xi(\lambda, s_2) - \Xi(\lambda, s_1), \qquad (4.24)
$$

where $\Xi(\lambda, s)$ is the spectral shift operator associated with $\Phi(z,s) = I_{\mathcal{H}} + sK^*(H_0 - z)^{-1}K$, $s \in [s_1, s_2]$.

Acknowledgments. F. G. would like to thank P. Exner for his kind invitation to take part in this conference, and all organizers for providing a most stimulating meeting. He also thanks H. Holden for the generous invitation to spend seven weeks in the summer of 1998 at the Norwegian University of Science and Technology, NTNU, Trondheim, during the final stages of this work. The extraordinary hospitality at the Department of Mathematical Sciences at NTNU and financial support by the Research Council Norway, Grant No. 107510/410 are gratefully acknowledged.

K. A. M. gratefully acknowledges financial support by the University of Missouri Research Board, Grant No. 98–119.

S. N. N. thanks H. Holden for making it possible to spend a week in July of 1998 at the Norwegian University of Science and Technology, NTNU, Trondheim.

References

[1] V. M. Adamjan and B. S. Pavlov, *A trace formula for dissipative operators,* Vestnik Leningrad. Univ. Math. **12**, 85–91 (1980).

[2] A. B. Aleksandrov, *The multiplicity of the boundary values of inner functions,* Sov. J. Contemp. Math. Anal. **22**, No. 5, 74–87 (1987).

[3] W. O. Amrein and K. B. Sinha, *On pairs of projections in a Hilbert space,* Lin. Algebra Appl. **208/209**, 425–435 (1994).

[4] N. Aronszajn and W. F. Donoghue, *On exponential representations of analytic functions in the upper half-plane with positive imaginary part,* J. d'Anal. Math. **5**, 321–388 (1956-57).

[5] H. Baumgärtel and M. Wollenberg, *Mathematical Scattering Theory,* Birkhäuser, Basel, 1983.

[6] M. Sh. Birman, *Spectral shift function and double operator integrals,* in Linear and Complex Analysis Problem Book 3, V. P. Havin and N. K. Nikolski (eds.), Lecture Notes in Mathematics **1573**, Springer, Berlin, 1994, pp. 272–273.

[7] M. Sh. Birman and S. B. Entina, *A stationary approach in the abstract theory of scattering,* Sov. Math. Dokl. **5**, 432–435 (1964).

[8] M. Sh. Birman and S. B. Entina, *The stationary method in the abstract theory of scattering,* Math. USSR Izv. **1**, 391–420 (1967).

[9] M. Sh. Birman and M. G. Krein, *On the theory of wave operators and scattering operators,* Sov. Math. Dokl. **3**, 740–744 (1962).

[10] M. Sh. Birman and A. B. Pushnitski, *Spectral shift function, amazing and multifaceted,* Integr. Eqs. Operator Th. **30**, 191–199 (1998).

[11] M. Sh. Birman and M. Z. Solomyak, *Remarks on the spectral shift function,* J. Sov. Math. **3**, 408–419 (1975).

[12] M. Sh. Birman and D. R. Yafaev, *The spectral shift function. The work of M. G. Krein and its further development,* St. Petersburg Math. J. **4**, 833–870 (1993).

[13] M. Sh. Birman and D. R. Yafaev, *Spectral properties of the scattering matrix,* St. Petersburg Math. J. **4**, 1055–1079 (1993).

[14] D. Bollé, F. Gesztesy, H. Grosse, W. Schweiger, and B. Simon, *Witten index, axial anomaly, and Krein's spectral shift function in supersymmetric quantum mechanics,* J. Math. Phys. **28**, 1512–1525 (1987).

[15] N. Borisov, W. Müller, and R. Schrader, *Relative index theorems and supersymmetric scattering theory,* Commun. Math. Phys. **114**, 475–513 (1988).

[16] M. S. Brodskii, *Triangular and Jordan Representations of Linear Operators,* Amer. Math. Soc., Providence, RI, 1971.

[17] D. Buschmann and G. Stolz, *Two-parameter spectral averaging and localization for non-monotoneous random Schrödnder operators,* preprint, 1998.

[18] R. W. Carey, *A unitary invariant for pairs of self-adjoint operators,* J. reine angewandte Math. **283**, 294–312 (1976).

[19] R. W. Carey and W. D. Pepe, *The phase shift and singular measures,* Indiana Univ. Math. J. **22**, 1049–1064 (1973).

[20] R. Carmona, *One-dimensional Schrödinger operators with random or deterministic potentials: new spectral types,* J. Funct. Anal. **51**, 229–258 (1983).

[21] R. Carmona, *Absolute continuous spectrum of one-dimensional Schrödinger operators,* in Differential Equations, I. W. Knowles and R. T. Lewis (eds.), North-Holland, Amsterdam, 1984, pp. 77–86.

[22] R. Carmona and J. Lacroix, *Spectral Theory of Random Schrödinger Operators,* Birkhäuser, Boston, 1990.

[23] T. Christiansen, *Spectral asymptotics for compactly supported perturbations of the Laplacian on $\mathbb{R}^n$,* Commun. Part. Diff. Eq. **23**, 933–948 (1998).

[24] J.-M. Combes and P. D. Hislop, *Localization for continuous random Hamiltonians in d-dimensions,* J. Funct. Anal. **124**, 149–180 (1994).

[25] J. M. Combes, P. D. Hislop, and E. Mourre, *Spectral averaging, perturbation of singular spectrum, and localization,* Trans. Amer. Math. Soc. **348**, 4883–4894 (1996).

[26] L. de Branges, *Perturbations of self-adjoint transformations,* Amer. J. Math. **84**, 543–560 (1962).

[27] R. del Rio, B. Simon, and G. Stolz, *Stability of spectral types for Sturm-Liouville operators,* Math. Res. Lett. **1**, 437–450 (1994).

[28] J. Dixmier, *Les Algèbres d'Opérateurs dans l'Espace Hilbertien,* 2nd ed., Gauthier-Villars, Paris, 1969.

[29] J.-P. Eckmann and C.-A. Pillet, *Zeta functions with Dirichlet and Neumann boundary conditions for exterior domains,* Helv. Phys. Acta **70**, 44–65 (1997).

[30] R. Geisler, V. Kostrykin, and R. Schrader, *Concavity properties of Krein's spectral shift function,* Rev. Math. Phys. **7**, 161–181 (1995).

[31] F. Gesztesy and H. Holden, *On new trace formulae for Schrödinger operators,* Acta Appl. Math. **39**, 315–333 (1995).

[32] F. Gesztesy, H. Holden, and B. Simon, *Absolute summability of the trace relation for certain Schrödinger operators,* Commun. Math. Phys. **168**, 137–161 (1995).
[33] F. Gesztesy, H. Holden, B. Simon, and Z. Zhao, *Higher order trace relations for Schrödinger operators,* Rev. Math. Phys. **7**, 893–922 (1995).
[34] F. Gesztesy, N. J. Kalton, K. A. Makarov, and E. Tsekanovskii, *Some Applications of Operator-Valued Herglotz Functions*, preprint, 1998.
[35] F. Gesztesy, M. Krishna, and G. Teschl, *On isospectral sets of Jacobi operators,* Commun. Math. Phys. **181**, 631–645 (1996).
[36] F. Gesztesy and K. A. Makarov, in preparation.
[37] F. Gesztesy, K. A. Makarov, and E. Tsekanovskii, *An Addendum to Krein's Formula,* J. Math. Anal. Appl. **222**, 594–606 (1998).
[38] F. Gesztesy and B. Simon, *Topological invariance of the Witten index,* J. Funct. Anal. **79**, 91–102 (1988).
[39] F. Gesztesy and B. Simon, *The ξ function,* Acta Math. **176**, 49–71 (1996).
[40] F. Gesztesy and E. Tsekanovskii, *On matrix-valued Herglotz functions,* preprint, 1997.
[41] I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators,* Amer. Math. Soc., Providence, RI, 1969.
[42] V. A. Javrjan, *On the regularized trace of the difference between two singular Sturm-Liouville operators,* Sov. Math. Dokl. **7**, 888–891 (1966).
[43] V. A. Javrjan, *A certain inverse problem for Sturm-Liouville operators,* Izv. Akad. Nauk Armjan. SSR Ser. Math. **6**, 246–251 (1971). (Russian.)
[44] A. Jensen and T. Kato, *Asymptotic behavior of the scattering phase for exterior domains,* Commun. Part. Diff. Eqs. **3**, 1165–1195 (1978).
[45] I. S. Kac and M. G. Krein, *R-functions-analytic functions mapping the upper halfplane into itself,* Amer. Math. Soc. Transl. (2) **103**, 1–18 (1974).
[46] T. Kato, *Monotonicity theorems in scattering theory,* Hadronic J. **1**, 134–154 (1978).
[47] V. Kostrykin and R. Schrader, *Scattering theory approach to random Schrödinger operators in one dimension,* Rev. Math. Phys., to appear.
[48] S. Kotani, *Lyapunov exponents and spectra for one-dimensional random Schrödinger operators,* Contemp. Math. **50**, 277–286 (1986).
[49] S. Kotani and B. Simon, *Localization in general one-dimensional random systems,* Commun. Math. Phys. **112**, 103–119 (1987).
[50] M. G. Krein, *A contribution to the theory of linear non-selfadjoint operators,* Sov. Math. Dokl. **1**, 38–40 (1960).
[51] M. G. Krein, *Perturbation determinants and a formula for the traces of unitary and self-adjoint operators,* Sov. Math. Dokl. **3**, 707–710 (1962).
[52] M. G. Krein, *On certain new studies in the perturbation theory for self-adjoint operators,* in *M. G. Krein, Topics in Differential and Integral Equations and Operator Theory*, I. Gohberg (ed.), Birkhäuser, Basel, 1983.
[53] M. G. Krein, *On perturbation determinants and a trace formula for certain classes of pairs of operators,* Amer. Math. Soc. Trans. (2) **145**, 39–84 (1989).
[54] M. G. Krein and V. A. Javrjan, *On spectral shift functions arising in perturbations of a positive operator,* J. Operator Th. **6**, 155–191 (1981). (Russian.)
[55] H. Langer, *Eine Erweiterung der Spurformel der Störungstheorie,* Math. Nachr. **30**, 123–135 (1965).
[56] I. M. Lifshits, *On a problem of perturbation theory,* Uspekhi Mat. Nauk **7**, No. 1, 171–180 (1952).
[57] I. M. Lifshits, *Some problems of the dynamic theory of nonideal crystal lattices,* Nuovo Cimento Suppl. **3**, (Ser. X), 716–734 (1956).
[58] A. Mohapatra and K. B. Sinha, *Spectral shift function and trace formula for unitaries - a new proof,* Integr. Eqs. Operator Th. **24**, 285–297 (1996).

[59] W. Müller, *Relative zeta functions, relative determinants and scattering theory,* Commun. Math. Phys. **192**, 309–347 (1998).

[60] S. N. Naboko, *Uniqueness theorems for operator-valued functions with positive imaginary part, and the singular spectrum in the Friedrichs model,* Ark. Mat. **25**, 115–140 (1987).

[61] S. N. Naboko, *Boundary values of analytic operator functions with a positive imaginary part,* J. Sov. Math. **44**, 786–795 (1989).

[62] S. N. Naboko, *Nontangential boundary values of operator-valued R-functions in a half-plane,* Leningrad Math. J. **1**, 1255–1278 (1990).

[63] S. N. Naboko, *On the conditions for existence of wave operators in the nonselfadjoint case,* Amer. Math. Soc. Transl. (2) **157**, 127–149 (1993).

[64] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators,* Springer, Berlin, 1992.

[65] V. V. Peller, *Hankel operators in the perturbation theory of unitary and self-adjoint operators,* Funct. Anal. Appl. **19**, 111-123 (1985).

[66] A. Poltoratski, *The Krein spectral shift and rank one perturbations of spectra,* St. Petersburg Math. J., to appear.

[67] A. B. Pushnitski, *Representation for the spectral shift function for perturbations of a definite sign,* preprint, 1997.

[68] A. B. Pushnitski, *Spectral shift function of the Schrödinger operator in the large coupling constant limit,* preprint, 1998.

[69] A. B. Pushnitski, *Integral estimates for the spectral shift function,* preprint, 1998.

[70] D. Robert, *Semi-classical and high energy asymptotics of the scattering phase for perturbations of elliptic operators,* in Mathematical Physics X, K. Schmüdgen (ed.), Springer, Berlin, 1992, pp.442–446.

[71] D. Robert, *Semi-classical approximation in quantum mechanics. A survey of old and recent mathematical results,* Helv. Phys. Acta **71**, 44–116 (1998).

[72] D. Robert, *Semi-classical asymptotics for the spectral shift function,* preprint.

[73] D. Robert and V. Sordoni, *Generalized determinants for Sturm-Liouville problems on the real line,* CRM Proceedings and Lecture Notes **12**, 251–259 (1997).

[74] A. V. Rybkin, *The spectral shift function, the characteristic function of a contraction, and a generalized integral,* Russian Acad. Sci. Sb. Math. **83**, 237–281 (1995).

[75] A. V. Rybkin, *On A-integrability of the spectral shift function of unitary operators arising in the Lax-Phillips scattering theory,* Duke Math. J. **83**, 683–699 (1996).

[76] Yu. L. Shmul'yan, *On operator R-functions,* Siberian Math. J. **12**, 315–322 (1971).

[77] B. Simon, *Trace Ideals and their Applications,* Cambridge University Press, Cambridge, 1979.

[78] B. Simon, *Localization in general one dimensional random systems, I. Jacobi matrices,* Commun. Math. Phys. **102**, 327–336 (1985).

[79] B. Simon, *Spectral analysis of rank one perturbations and applications,* CRM Proceedings and Lecture Notes **8**, 109–149 (1995).

[80] B. Simon, *Spectral averaging and the Krein spectral shift,* Proc. Amer. Math. Soc. **126**, 1409–1413 (1998).

[81] B. Simon and T. Wolff, *Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians,* Commun. Pure Appl. Math. **39**, 75–90 (1986).

[82] K. B. Sinha, *On the theorem of M. G. Krein,* preprint, 1975, unpublished.

[83] K. B. Sinha and A. N. Mohapatra, *Spectral shift function and trace formula,* Proc. Indian Acad. Sci. (Math. Sci) **104**, 819–853 (1994).

[84] A. V. Sobolev, *Efficient bounds for the spectral shift function,* Ann. Inst. H. Poincaré **A58**, 55–83 (1993).

[85] B. Sz.-Nagy and C. Foias, *Sur les contractions de l'espace de Hilbert. IV,* Acta Sci. Math. **21**, 251–259 (1960).

[86] B. Sz.-Nagy and C. Foias, *Harmonic Analysis of Operators on Hilbert Space,* North-Holland, Amsterdam, 1970.
[87] D. Voiculescu, *On a trace formula of M. G. Krein,* in Operator Theory: Advances and Applications, Vol. 24, Birkhäuser, Basel, 1987, pp. 329–332.
[88] D. R. Yafaev, *Mathematical Scattering Theory,* Amer. Math. Soc., Providence, RI, 1992.

Department of Mathematics, University of Missouri, Columbia, MO 65211, USA
E-mail address: fritz@math.missouri.edu
URL: http://www.math.missouri.edu/people/fgesztesy.html

Department of Mathematics, University of Missouri, Columbia, MO 65211, USA
E-mail address: makarov@azure.math.missouri.edu

Department of Mathematical Physics, St. Petersburg State University, 198904 St. Petersburg, Russia
E-mail address: naboko@snoopy.phys.spbu.ru