ON THE KOPLIENKO SPECTRAL SHIFT FUNCTION, I. BASICS

FRITZ GESZTESY[1], ALEXANDER PUSHNITSKI[2], AND BARRY SIMON[3]

Abstract. We study the Koplienko Spectral Shift Function (KoSSF), which is distinct from the one of Krein (KrSSF). KoSSF is defined for pairs A, B with $(A - B) \in \mathcal{I}_2$, the Hilbert–Schmidt operators, while KrSSF is defined for pairs A, B with $(A-B) \in \mathcal{I}_1$, the trace class operators. We review various aspects of the construction of both KoSSF and KrSSF. Among our new results are: (i) that any positive Riemann integrable function of compact support occurs as a KoSSF; (ii) that there exist A, B with $(A-B) \in \mathcal{I}_2$ so $\det_2((A-z)(B-z)^{-1})$ does not have nontangential boundary values; (iii) an alternative definition of KoSSF in the unitary case; and (iv) a new proof of the invariance of the a.c. spectrum under $\mathcal{I}_1$-perturbations that uses the KrSSF.

1. Introduction

In 1941, Titchmarsh [63] (see also [20, pp. 1564–1566] for the result) proved that if

$$V \in L^1((0,\infty); dx), \quad V \text{ real-valued},$$

Date: May 16, 2007.

2000 *Mathematics Subject Classification.* Primary: 47A10, 81Q10. Secondary: 34B27, 47A40, 81Uxx.

Key words and phrases. Krein's spectral shift function, Koplienko's spectral shift function, self-adjoint operators, trace class and Hilbert–Schmidt perturbations, convexity properties, boundary values of (modified) Fredholm determinants.

[1] Department of Mathematics, University of Missouri, Columbia, MO 65211, USA. E-mail: fritz@math.mizzou.edu. Supported in part by NSF Grant DMS-0405526.

[2] Department of Mathematics, King's College London, Strand, London WC2R 2LS, England, UK. E-mail: alexander.pushnitski@kcl.ac.uk. Supported in part by the Leverhulme Trust.

[3] Mathematics 253-37, California Institute of Technology, Pasadena, CA 91125, USA. E-mail: bsimon@caltech.edu. Supported in part by NSF Grant DMS-0140592 and U.S.–Israel Binational Science Foundation (BSF) Grant No. 2002068.

Submitted to the Marchenko and Pastur birthday issue of Journal of Mathematical Physics, Analysis and Geometry.

and

$$
\begin{aligned}
&H_\theta = -\frac{d^2}{dx^2} + V, \qquad (1.1)\\
&\mathrm{dom}(H_\theta) = \{f \in L^2((0,\infty);dx) \,|\, f, f' \in AC([0,R]) \text{ for all } R>0;\\
&\qquad \sin(\theta)f'(0) + \cos(\theta)f(0) = 0;\ (-f'' + Vf) \in L^2((0,\infty);dx)\},
\end{aligned}
$$

for some $\theta \in [0,\pi)$, then

$$\sigma_{\mathrm{ac}}(H_\theta) = [0,\infty).$$

(Actually, he explicitly computed the spectral function in terms of the inverse square of the modulus of the Jost function for positive energies.) It was later realized that the a.c. invariance, that is,

$$\sigma_{\mathrm{ac}}(H_\theta) = \sigma_{\mathrm{ac}}(H_{0,\theta}) \qquad (1.2)$$

with

$$
\begin{aligned}
&H_{0,\theta} = -\frac{d^2}{dx^2}, \qquad (1.3)\\
&\mathrm{dom}(H_{0,\theta}) = \{f \in L^2((0,\infty);dx) \,|\, f, f' \in AC([0,R]) \text{ for all } R>0;\\
&\qquad \sin(\theta)f'(0) + \cos(\theta)f(0) = 0;\ f'' \in L^2((0,\infty);dx)\},
\end{aligned}
$$

is a special case of an invariance of the absolutely continuous spectrum, $\sigma_{\mathrm{ac}}(\cdot)$ for the passage from A to B if $(A-B) \in \mathcal{I}_1$, the trace class. In the present context of the pair $(H_\theta, H_{0,\theta})$ one has $[(H_\theta + E)^{-1} - (H_{0,\theta} + E)^{-1}] \in \mathcal{I}_1$ for $E>0$ sufficiently large. The abstract trace class result is associated with Birman [10, 11], Kato [31, 32], and Rosenblum [54].

Our original and continuing motivation is to find a suitable operator theoretic result connected with the remarkable discovery of Deift–Killip [18] that for the above (1.1)/(1.3) case, one has (1.2) if one only assumes $V \in L^2((0,\infty);dx)$. Note that $V \in L^2((0,\infty);dx)$ implies that

$$[(H_\theta + E)^{-1} - (H_{0,\theta} + E)^{-1}] \in \mathcal{I}_2,$$

the Hilbert–Schmidt class. However, there is no totally general invariance result for a.c. spectrum under non-trace class perturbations: It is a result of Weyl [67] and von Neumann [65] that given any self-adjoint A, there is a B with pure point spectrum and $(A-B) \in \mathcal{I}_2$. Kuroda [43] extends this to $\mathcal{I}_p$, $p \in (1,\infty)$, the trace ideals. Thus, we seek general operator criteria on when $(A-B) \in \mathcal{I}_2$ but (1.2) still holds.

We hope such a criterion will be found in the spectral shift function of Koplienko [36] (henceforth KoSSF), an object which we believe has not received the attention it deserves. One of our goals in the present paper is to make propaganda for this object.

Two references for trace ideals we quote extensively are Gohberg–Krein [27] and Simon [60]. We follow the notation of [60]. Throughout this paper all Hilbert spaces are assumed to be complex and separable.

The KoSSF, $\eta(\lambda; A, B)$, is defined when A and B are bounded self-adjoint operators satisfying $(A-B) \in \mathcal{I}_2$, and is given by

$$\int_{\mathbb{R}} f''(\lambda)\eta(\lambda; A, B)\, d\lambda = \operatorname{Tr}\biggl(f(A) - f(B) - \frac{d}{d\alpha} f(B + \alpha(A-B))\Big|_{\alpha=0}\biggr), \tag{1.4}$$

where the right-hand side is sometimes (certainly if $(A-B) \in \mathcal{I}_1$) the simpler-looking

$$\operatorname{Tr}(f(A) - f(B) - (A-B)f'(B)). \tag{1.5}$$

η has two critical properties: $\eta \in L^1(\mathbb{R})$ and $\eta \geq 0$. We mainly consider bounded A, B here, but see the remarks in Section 9.

Formula (1.4) requires some assumptions on f. In Koplienko's original paper [36] the case $f(x) = (x-z)^{-1}$ was considered and then (1.4) was extended to the class of rational functions with poles off the real axis. Later, Peller [52] extended the class of functions f and found sharp sufficient conditions on f which guarantee that (1.4) holds. These conditions were stated in terms of Besov spaces. Essentially, Peller's construction requires that (1.4) hold for some sufficiently wide class of functions, so that this class is dense in a certain Besov space, and then provides an extension onto the whole of this Besov space.

We will use this aspect of Peller's work and will not worry about the classes of f in this paper. For the most part we will work with $f \in C^\infty(\mathbb{R})$ and Peller's construction provides an extension to a wider function class.

The model for the KoSSF is, of course, the spectral shift function of Krein (henceforth KrSSF), denoted by $\xi(\lambda; A, B)$, and defined for A, B with $(A-B) \in \mathcal{I}_1$ by

$$\int_{\mathbb{R}} \xi(\lambda; A, B) f'(\lambda)\, d\lambda = \operatorname{Tr}(f(A) - f(B)). \tag{1.6}$$

In the appendix, we recall a quick way to define ξ, its main properties and, most importantly, present an argument that shows how it can be used to derive the invariance of a.c. spectrum without recourse to scattering theory.

As we will see in Section 2, it is easy to construct analogs of η for any $\mathcal{I}_n$, $n \in \mathbb{N}$, but they are only tempered distributions. What makes η different is its positivity, which also implies it lies in $L^1(\mathbb{R})$ (by taking f suitably). This positivity should be thought of as a general convexity result—something hidden in Koplienko's paper [36].

One of our goals here is to emphasize this convexity. Another is to present a "baby" finite-dimensional version of the double Stieltjes operator integral of Birman–Solomyak [12, 13, 15], essentially due to Löwner [46], whose contribution here seems to have been overlooked.

In Section 2, we define η when $(A-B)$ is trace class, and in Section 3, we discuss the convexity result that is equivalent to positivity of η. In Section 4, we prove a lovely bound of Birman–Solomyak [13, 15]:

$$\|f(A)-f(B)\|_{\mathcal{I}_2} \leq \|f'\|_\infty \|A-B\|_{\mathcal{I}_2}. \tag{1.7}$$

Here and in the remainder of this paper $\|\cdot\|_{\mathcal{I}_p}$ denotes the norm in the trace ideals $\mathcal{I}_p$, $p\in[1,\infty)$. In Section 5, we use (1.7) plus positivity of η to complete the construction of η.

We want to emphasize an important distinction between the KrSSF and the KoSSF. The former satisfies a chain rule

$$\xi(\,\cdot\,;A,C)=\xi(\,\cdot\,;A,B)+\xi(\,\cdot\,;B,C), \tag{1.8}$$

while η instead satisfies a corrected chain rule

$$\eta(\,\cdot\,;A,C)=\eta(\,\cdot\,;A,B)+\eta(\,\cdot\,;B,C)+\delta\eta(\,\cdot\,;A,B,C), \tag{1.9}$$

where $\delta\eta$ satisfies

$$\int_{\mathbb{R}} g'(\lambda)\delta\eta(\lambda;A,B)\,d\lambda = \operatorname{Tr}((A-B)(g(B)-g(C))). \tag{1.10}$$

(Here g corresponds to f' when comparing with (1.4)–(1.6).) It is in estimating (1.10) that (1.7) will be critical.

We view Sections 2–5 as a repackaging in a prettier ribbon of Koplienko's construction in [36]. Section 6 explores what η's can occur. In Sections 7 and 8, we discuss the connection to $\det_2(\cdot)$ and present a new result: an example of $(A-B)\in\mathcal{I}_2$ where $\det_2((A-z)(B-z)^{-1})$ does not have nontangential limits to the real axis a.e. This is in contradistinction to the KrSSF, where $(A-B)\in\mathcal{I}_1$ implies $\det((A-z)(B-z)^{-1})$ has a nontangential limit $z\to\lambda$ for a.e. $\lambda\in\mathbb{R}$. The latter is a consequence of the formula

$$\log(\det((A-z)(B-z)^{-1})) = \int_{\mathbb{R}} (\lambda-z)^{-1}\xi(\lambda;A,B)\,d\lambda, \quad z\in\mathbb{C}\backslash\mathbb{R}, \tag{1.11}$$

since the right-hand side of (1.11) represents a difference of two Herglotz functions.

Sections 9 and 10 discuss extensions of η to the case of unbounded operators with a trace class condition on the resolvents and to unitary operators. Here a key is that η is not determined until one makes a choice of interpolation. Section 11 discusses some conjectures.

In a future joint work, we will explore what one can learn about the KoSSF from Szegő's theorem [59], the work of Killip–Simon [34] and of Christ–Kiselev [17]. This will involve the study of η for suitable Schrödinger operators and Jacobi and CMV matrices for perturbations in L^p, respectively, ℓ^p, $p \in [1, 2)$.

We are indebted to E. Lieb, K. A. Makarov, V. V. Peller, and M. B. Ruskai for useful discussions. F. G. and A. P. wish to thank Gary Lorden and Tom Tombrello for the hospitality of Caltech where some of this work was done. F. G. gratefully acknowledges a research leave for the academic year 2005/06 granted by the Research Council and the Office of Research of the University of Missouri-Columbia. A. P. gratefully acknowledges financial support by the Leverhulme Trust.

It is a great pleasure to dedicate this paper to the birthdays of two giants of spectral theory: Vladimir A. Marchenko and Leonid A. Pastur.

2. The KoSSF $\eta(\,\cdot\,; A, B)$ in the Trace Class Case

We begin with what can be said of $\mathcal{I}_n$ perturbations, $n \in \mathbb{N}$, and then turn to what is special for $n = 1, 2$. We note that our approach has common elements to the one used by Dostanić [19].

Proposition 2.1. *Let A, B be bounded self-adjoint operators with*

$$A = B + X. \tag{2.1}$$

For $\alpha, t \in \mathbb{R}$, define

$$f_t(\alpha) = e^{it(B+\alpha X)}.$$

Then $f_t(\alpha)$ is C^∞ in α and

$$\frac{d^k f_t}{d\alpha^k} = i^k \int_{\substack{0<s_j<t \\ \sum_{j=1}^k s_j<t}} f_{s_1}(\alpha) X f_{s_2}(\alpha) \dots f_{s_k}(\alpha) X f_{t-s_1-\dots-s_k}(\alpha)\, ds_1 \dots ds_k. \tag{2.2}$$

If $X \in \mathcal{I}_p$ for $p \geq k$, $k \in \mathbb{N}$, then $d^k f_t / d\alpha^k \in \mathcal{I}_{p/k}$ and

$$\left\| \frac{d^k f_t}{d\alpha^k} \right\|_{\mathcal{I}_{p/k}} \leq \frac{t^k}{k!} \|X\|_{\mathcal{I}_p}^k. \tag{2.3}$$

In particular, if $n \in \mathbb{N}$ and $X \in \mathcal{I}_n$, then

$$g_t(A, B) \equiv \left(e^{itA} - e^{itB} - \sum_{k=1}^{n-1} \frac{1}{k!} \left(\frac{d}{d\alpha} \right)^k f_t(\alpha) \bigg|_{\alpha=0} \right) \in \mathcal{I}_1, \tag{2.4}$$

$$\|g_t(A, B)\|_{\mathcal{I}_1} \leq \frac{t^n}{n!} \|X\|_{\mathcal{I}_n}^n. \tag{2.5}$$

Proof. For $k = 1$, (2.2) comes from taking limits in DuHamel's formula

$$e^C - e^D = \int_0^1 e^{\beta C}(C - D)e^{(1-\beta)D}\, d\beta.$$

The general k case then follows by induction.

(2.2) implies (2.3) by Hölder's inequality for operators (see [60, p. 21]). (2.4) is then Taylor's theorem with remainder and (2.5) follows from (2.3). $\square$

Theorem 2.2. *Let A, B be bounded self-adjoint operators such that $X = (A - B) \in \mathcal{I}_n$ for some $n \in \mathbb{N}$. Let f be of compact support with $\widehat{f}$, its Fourier transform, satisfying*

$$\int_{\mathbb{R}} (1 + |k|)^n |\widehat{f}(k)|\, dk < \infty. \tag{2.6}$$

Then,

$$\biggl(f(A) - f(B) - \sum_{j=1}^{n-1} \frac{1}{k!}\biggl(\frac{d}{d\alpha}\biggr)^k f(B + \alpha X)\biggr|_{\alpha=0}\biggr) \in \mathcal{I}_1, \tag{2.7}$$

and there is a distribution T with

$$\mathrm{Tr}\biggl(f(A) - f(B) - \sum_{j=1}^{n-1} \frac{1}{k!}\biggl(\frac{d}{d\alpha}\biggr)^k f(B + \alpha X)\biggr|_{\alpha=0}\biggr) = \int_{\mathbb{R}} T(\lambda) f^{(n)}(\lambda)\, d\lambda. \tag{2.8}$$

Moreover, the distribution T is such that $\widehat{T} \in L^\infty(\mathbb{R}; dt)$.

Proof. This is immediate from the estimates in Proposition 2.1 and

$$f(A) = (2\pi)^{-1/2} \int_{\mathbb{R}} \widehat{f}(t)\, e^{itA}\, dt.$$

For, by (2.5), we have

$$\|\text{LHS of (2.7)}\|_{\mathcal{I}_1} \le C \int_{\mathbb{R}} |t|^n |\widehat{f}(t)|\, dt = C \int_{\mathbb{R}} |\widehat{f^{(n)}}(t)|\, dt. \tag{2.9}$$

Thus, (2.8) defines a distribution T with

$$|T(f)| \le C \int_{\mathbb{R}} |\widehat{f}(t)|\, dt,$$

so $\widehat{T}$ is a function in $L^\infty(\mathbb{R}; dt)$. $\square$

Notice that, as we have seen,

$$\frac{d}{d\alpha}\, e^{it(B+\alpha X)}\biggr|_{\alpha=0} = i \int_0^t e^{i\beta B} X e^{i(t-\beta)B}\, d\beta$$

so that formally,

$$\operatorname{Tr}\biggl(\frac{d}{d\alpha}\, f(B+\alpha X)\bigg|_{\alpha=0} - Xf'(B)\biggr) = 0,$$

and formally, (1.5) can replace the right-hand side of (1.4). This can be proven if the commutator $[B,X]=[B,A]$ is trace class.

Dostanić [19, Theorem 2.9] essentially proves that T is the derivative of an $L^2(\mathbb{R};d\lambda)$-function.

A key point for us is that in the case $n=2$, the distribution T is given by an $L^1(\mathbb{R};d\lambda)$-function. We start this construction here by considering the trace class case.

Lemma 2.3. *Let B be a self-adjoint operator and let $X=(A-B)\in\mathcal{I}_1$. Then there is a* (*complex*) *measure $d\mu_{B,X}$ on $\mathbb{R}$ such that for any bounded Borel function, f,*

$$\operatorname{Tr}(Xf(B)) = \int_{\mathbb{R}} f(\lambda)\, d\mu_{B,X}(\lambda). \tag{2.10}$$

Equation (2.10) *defines $d\mu_{B,X}$ uniquely.*

Proof. Equation (2.10) yields uniqueness of the measure $d\mu_{B,X}$ since it defines the integral for all continuous functions f. Regarding existence of $d\mu_{B,X}$, the spectral theorem asserts the existence of measures $d\mu_{B;\varphi,\psi}$, such that

$$\langle\varphi, f(B)\psi\rangle = \int_{\mathbb{R}} f(\lambda)\, d\mu_{B;\varphi,\psi}(\lambda) \tag{2.11}$$

and

$$\|\mu_{B;\varphi,\psi}\| \leq \|\varphi\|\,\|\psi\|. \tag{2.12}$$

The canonical decomposition for X (see [60, Sect. 1.2]) says (with N finite or infinite)

$$X = \sum_{j=1}^{N} \mu_j(X)\langle\varphi_j,\,\cdot\,\rangle\psi_j, \tag{2.13}$$

where $\{\psi_j\}_{j=1}^N$ and $\{\varphi_j\}_{j=1}^N$ are orthonormal sets, $\mu_j>0$, and

$$\sum_{j=1}^{N}\mu_j(X) = \|X\|_{\mathcal{I}_1}. \tag{2.14}$$

Define

$$d\mu_{B,X} = \sum_{j=1}^{N}\mu_j(X)\, d\mu_{B;\varphi_j,\psi_j} \tag{2.15}$$

which converges by (2.12) and (2.14). □

Theorem 2.4. *Let A, B be bounded operators and $X = (A - B) \in \mathcal{I}_1$. Let $\xi(\lambda; A, B)$ be the KrSSF. Let $d\mu_{B,X}$ be given by* (2.10). *Define*

$$\eta(\lambda; A, B) \equiv \mu_{B,X}((-\infty, \lambda)) - \int_{-\infty}^{\lambda} \xi(\lambda'; A, B)\, d\lambda', \quad \lambda \in \mathbb{R}. \tag{2.16}$$

Then $\eta(\,\cdot\,; A, B)$ has compact support and for any $f \in C^\infty(\mathbb{R})$, we have

$$\operatorname{Tr}\biggl(f(A) - f(B) - \frac{d}{d\alpha} f(B + \alpha X)\biggr|_{\alpha=0} \biggr) = \int_{\mathbb{R}} f''(\lambda)\eta(\lambda; A, B)\, d\lambda. \tag{2.17}$$

Remarks. 1. Since

$$\int_{\mathbb{R}} d\mu_{B,X}(\lambda) = \operatorname{Tr}(X) = \int_{\mathbb{R}} \xi(\lambda; A, B)\, d\lambda,$$

we can replace $(-\infty, \lambda)$ in both places in (2.16) by $[\lambda, \infty)$. This shows that η in (2.16) has compact support.

2. (2.17) determines η uniquely up to an affine term. The condition that η have compact support (as the η of (2.16) does) determines η uniquely.

Proof. We first claim that $\frac{d}{d\alpha} f(B + \alpha X)\bigr|_{\alpha=0}$ is trace class and that

$$\operatorname{Tr}\biggl(\frac{d}{d\alpha} f(B + \alpha X)\biggr|_{\alpha=0} \biggr) = \operatorname{Tr}(X f'(B)). \tag{2.18}$$

This is immediate for f nice enough (e.g., f such that (2.6) holds for $n = 1$) since

$$\operatorname{Tr}(e^{i\alpha\beta B} X e^{i\alpha(1-\beta)B}) = \operatorname{Tr}(X e^{i\alpha B}). \tag{2.19}$$

Thus, by (2.10),

$$\begin{aligned} \operatorname{Tr}\biggl(\frac{d}{d\alpha} f(B + \alpha X) \biggr) &= \int_{\mathbb{R}} f'(\lambda)\, d\mu_{B,X}(\lambda) \\ &= -\int_{\mathbb{R}} f''(\lambda)[\mu_{B,X}((-\infty, \lambda))]. \end{aligned}$$

Similarly, by (1.6),

$$\begin{aligned} \operatorname{Tr}(f(A) - f(B)) &= \int_{\mathbb{R}} f'(\lambda)\xi(\lambda; A, B)\, d\lambda \\ &= -\int_{\mathbb{R}} f''(\lambda) \biggl(\int_{-\infty}^{\lambda} \xi(\lambda'; A, B)\, d\lambda' \biggr) d\lambda. \end{aligned}$$ □

The next critical step will be to prove positivity of η.

3. Convexity of $\mathrm{Tr}(f(A))$

Positivity of η is essentially equivalent to the following result:

Theorem 3.1. *Let f be a convex function on $\mathbb{R}$. Then the mapping*

$$A \mapsto \mathrm{Tr}(f(A)) \tag{3.1}$$

is a convex function on the $m \times m$ self-adjoint matrices for every $m \in \mathbb{N}$.

Remarks. 1. More generally, if f is convex on (a,b), (3.1) is convex on matrices A with spectrum in (a,b). In fact, it is easy to see that any convex function f on (a,b) is a monotone limit on (a,b) of convex functions on $\mathbb{R}$. So this more general result is a consequence of Theorem 3.1.

2. We will discuss the infinite-dimensional situation below.

Two special cases of this are widely known and used:
(a) $A \mapsto \mathrm{Tr}(e^A)$ is convex.
(b) $A \mapsto \mathrm{Tr}(A\log(A))$ is convex on $A \geq 0$.
Both of these are rather special. In the first case, one has the stronger $A \mapsto \log(\mathrm{Tr}(e^A))$ is convex and the usual proof of it is via Hölder's inequality (cf., e.g., [28, p. 19–20] or [58, p. 57]) which proves the strong convexity of the $\log(\cdot)$, but does not prove Theorem 3.1. In the second case, by Kraus' theorem [37, 8] $A \mapsto A\log(A)$ is operator convex. (We also note that A^r, $r \in \mathbb{R}$, is operator convex for $A > 0$ if and only if $r \in [-1,0] \cup [1,2]$ (cf. [9, p. 147]).) We have found Theorem 3.1 stated in Alicki–Fannes [3, Sect. 9.1] and an equivalent statement in Ruelle [55, Sect. 2.5] (who attributes it to Klein [35] although Klein only has the special case $f(x) = x\log(x)$ and his proof is specific to that case; Ruelle's is not). We have also found it in Lieb–Pedersen [45] whose proof is closer to the one we label "Third Proof" below. The result is also mentioned in von Neumann [66], although the proof he gives earlier for a special f does not seem to establish the general case.

In any event, even though this result is not hard and is known to some experts, we provide several proofs because it is not widely known and is central to the theory of KoSSF. We provide several proofs because they illustrate different aspects of the theorem.

First Proof. This uses eigenvalue perturbation theory. By a limiting argument, it suffices to prove it for functions $f \in C^\infty(\mathbb{R})$. By approximating derivatives of f by polynomials, we see that matrix elements, and so the trace of $f(A)$, are C^∞-functions of A. By a limiting argument, we need only show $\lambda \to \mathrm{Tr}(A + \lambda X)$ has a nonnegative second derivative at $\lambda = 0$ in case A has distinct eigenvalues.

So by changing basis, we suppose A is diagonal with the eigenvalues $a_1 < a_2 < \cdots < a_m$. Let $e_j(\lambda)$ be the eigenvalue of $A + \lambda X$ near a_j for $|\lambda|$ sufficiently small. As is well known [33, Sect. II.2], [53, Sect. XII.1],

$$\frac{d^2 e_j}{d\lambda^2}\bigg|_{\lambda=0} = \sum_{\substack{k=1 \\ k\neq j}}^{m} \frac{|X_{k,j}^2|}{a_j - a_k}. \tag{3.2}$$

Clearly,

$$\begin{aligned} &\frac{d}{d\lambda}\,[f(e_j(\lambda))] = f'(e_j(\lambda))e_j'(\lambda), \\ &\frac{d^2}{d\lambda^2}\,f(e_j(\lambda)) = f''(e_j(\lambda))e_j'(\lambda)^2 + f'(e_j(\lambda))e_j''(\lambda), \end{aligned}$$

so

$$\frac{d^2}{d\lambda^2}\,\mathrm{Tr}(f(A(\lambda)))\bigg|_{\lambda=0} = \boxed{1} + \boxed{2}, \tag{3.3}$$

where

$$\boxed{1} = \sum_{j=1}^{m} f''(a_j)e_j'(0)^2 \geq 0$$

since $f'' \geq 0$ and, by (3.2),

$$\boxed{2} = \sum_{\substack{j,k=1 \\ k\neq j}}^{m} |X_{k,j}|^2 \left[\frac{f'(a_j) - f'(a_k)}{a_j - a_k}\right] \geq 0$$

since for $x < y$,

$$\frac{f'(y) - f'(x)}{y - x} = \frac{1}{y - x}\int_x^y f''(u)\,du \geq 0. \qquad \square$$

Second Proof. This one uses a variational principle. We consider first the case

$$f^+(x) = x_+ \equiv \begin{cases} x, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{3.4}$$

We claim first that

$$\mathrm{Tr}(f^+(A)) = \max\{\mathrm{Tr}(AB) \mid \|B\| \leq 1,\, B \geq 0\}, \tag{3.5}$$

where $\|\cdot\|$ is the matrix norm on $\mathbb{C}^m$ with the Euclidean norm. For in an orthonormal basis where A is a diagonal matrix,

$$\mathrm{Tr}(AB) = \sum_{j=1}^{m} a_j b_{j,j} \leq \sum_{j=1}^{m} (a_j)_+ = \mathrm{Tr}(f^+(A))$$

if $0 \leq b_{jj} \leq 1$. On the other hand, if B is the diagonal matrix with

$$b_{jj} = \begin{cases} 1, & a_j > 0, \\ 0, & a_j \leq 0, \end{cases}$$

then $B \geq 0$, $\|B\| \leq 1$, and $\operatorname{Tr}(AB) = \operatorname{Tr}(f^+(A))$. This proves (3.5).

Convexity is immediate for f^+ given by (3.4) once we have (3.5), since maxima of linear functionals are convex. Obviously, since $(x-\lambda)_+$ is just a translate of x_+, we get convexity for any function of

$$\int_{\lambda_0}^{\infty} (x-\lambda)_+ \, d\mu(\lambda)$$

for any Borel measure μ on (λ_0, ∞). But every convex function f with $f \equiv 0$ for $x \leq \lambda_0$ has this form.

Adding $ax + b$ to this, we get the result for any convex function f with $f''(x) = 0$ for $x \leq \lambda_0$. Taking $\lambda_0 \to -\infty$, we get the result for general convex functions f. □

Third Proof (*M. B. Ruskai, private communication*). If f is any convex function, C a self-adjoint $m \times m$ matrix with

$$Ce_j = \lambda_j e_j, \tag{3.6}$$

and $v \in \mathbb{C}^m$ a unit vector, then

$$\begin{aligned} \langle v, f(C)v \rangle &= \sum_{j=1}^{m} |\langle v, e_j \rangle|^2 f(\lambda_j) \\ &\geq f\bigg(\sum_{j=1}^{m} \lambda_j |\langle v, e_j \rangle|^2\bigg) \qquad (3.7) \\ &= f(\langle v, Cv \rangle), \qquad (3.8) \end{aligned}$$

where (3.7) employs Jensen's inequality.

Now suppose

$$C = \theta A + (1-\theta)B, \quad \theta \in [0,1].$$

Then,

$$\begin{aligned} \operatorname{Tr}(f(C)) &= \sum_{j=1}^{m} f(\langle e_j, Ce_j \rangle) \\ &= \sum_{j=1}^{m} f(\theta \langle e_j, Ae_j \rangle + (1-\theta)\langle e_j, Be_j \rangle) \end{aligned}$$

$$\leq \sum_{j=1}^{m} [\theta f(\langle e_j, Ae_j\rangle) + (1-\theta) f(\langle e_j, Be_j\rangle)] \tag{3.9}$$

$$\leq \theta \sum_{j=1}^{m} \langle e_j, f(A)e_j\rangle + (1-\theta)\sum_{j=1}^{m} \langle e_j, f(B)e_j\rangle \tag{3.10}$$

$$= \theta \mathrm{Tr}(f(A)) + (1-\theta)\mathrm{Tr}(f(B)), \tag{3.11}$$

proving convexity. In the above, (3.9) is direct convexity of f and (3.10) is (3.8) for $v = e_j$ and $C = A$ or B. $\square$

Corollary 3.2. *If $f \in C^1(\mathbb{R})$ is convex and B and X are $m \times m$ self-adjoint matrices, $m \in \mathbb{N}$, then*

$$\mathrm{Tr}\biggl(f(B+X) - f(B) - \frac{d}{d\alpha} f(B+\alpha X)\bigg|_{\alpha=0}\biggr) \geq 0. \tag{3.12}$$

Remarks. 1. It is not hard to see that (3.12) is equivalent to Theorem 3.1.

2. It is in this form that the result appears in Ruelle [55, Sect. 2.5], and for the case $f(x) = x\log(x)$, $x > 0$, in Klein [35].

Proof. If $g \in C^1(\mathbb{R})$ is a convex function,

$$g(x+y) - g(x) - g'(x)y \geq 0, \tag{3.13}$$

since convexity says that g lies above the tangent line at any point. (3.12) is (3.13) for $g(\alpha) = \mathrm{Tr}(f(B+\alpha X))$, $x = 0$, $y = 1$. $\square$

Corollary 3.3. *For finite-dimensional matrices A and B, the KoSSF, $\eta(\,\cdot\,; A, B)$, satisfies $\eta(\lambda; A, B) \geq 0$ for a.e. $\lambda \in \mathbb{R}$.*

Proof. Let $h : \mathbb{R} \to [0,\infty)$ be a measurable function bounded and supported on an interval (a,b) with $\sigma(A) \cup \sigma(B) \subset (a,b)$ (so, by (2.16), η is supported on (a,b)). Let f be the unique convex function with $f = 0$ near $-\infty$ and $f'' = h$. By (3.12) and (2.17),

$$0 \leq \int_{\mathbb{R}} h(\lambda)\eta(\lambda; A, B)\, d\lambda. \tag{3.14}$$

Since h is arbitrary, $\eta \geq 0$ a.e. $\square$

Theorem 3.4. *For any finite self-adjoint matrices A, B (of the same size),*

$$\int_{\mathbb{R}} |\eta(\lambda; A, B)|\, d\lambda = \tfrac{1}{2}\, \|A - B\|_{\mathcal{I}_2}^2. \tag{3.15}$$

Remarks. 1. It is remarkable that we always have equality in (3.15). The analog for the KrSSF is

$$\int_{\mathbb{R}} |\xi(\lambda; A, B)|\, d\lambda \leq \|A - B\|_{\mathcal{I}_1}, \tag{3.16}$$

where equality, in general, holds if $A - B$ is either positive or negative.

2. (3.15) emphasizes again the lack of a chain rule for η; η is nonlinear in $(A - B)$.

Proof. Take $f(x) = \frac{1}{2}x^2$ such that $f''(x) = 1$ and

$$\begin{aligned} f(B+X) - f(B) - \frac{d}{d\alpha} f(B+\alpha X)\bigg|_{\alpha=0} \\ = \tfrac{1}{2}\big[(B+X)^2 - B^2 - XB - BX\big] = \tfrac{1}{2} X^2. \end{aligned}$$

Since $\eta \geq 0$, $\int_{\mathbb{R}} f''(\lambda)\eta(\lambda; A, B)\, d\lambda = \int_{\mathbb{R}} |\eta(\lambda; A, B)|\, d\lambda$ and (3.15) holds. $\square$

In Section 5 we take limits from the finite-dimensional situation, but one can easily extend Theorem 3.1 in two ways and from there directly prove $\eta \geq 0$ and (3.15) in case $(A - B) \in \mathcal{I}_1$. Without proof, we state the extensions (the results are simple limiting arguments from finite dimensions):

Theorem 3.5. *If f is convex on $\mathbb{R}$ and $f(0) = 0$, then $f(A)$ is trace class for any self-adjoint trace class operator A, and for such A's, the mapping $A \mapsto \operatorname{Tr}(f(A))$ is convex.*

In this context we note that convex functions are Lipschitz continuous. (For this and additional regularity results of convex functions, see, e.g., [9, p. 145–146].)

Theorem 3.6. *For any convex function $f \in C^\infty(\mathbb{R})$, any bounded self-adjoint operator B, and any self-adjoint operator $X \in \mathcal{I}_1$,*

$$[f(B+X) - f(B)] \in \mathcal{I}_1$$

and the mapping $X \mapsto \operatorname{Tr}(f(B+X) - f(B))$ is convex.

Convexity of maps of the type $s \mapsto \operatorname{Tr}(f(B + X(s)) - f(B))$, $s \in (s_1, s_2)$, for convex f and certain classes of $X(\cdot) \in \mathcal{I}_1$ was also studied in [24].

4. Löwner's Formula and the Finite-Dimensional Birman–Solomyak Bound

The final element needed to construct the KoSSF is the following lovely theorem of Birman–Solomyak [13] (see also [15]):

Theorem 4.1. *Let A, B be bounded self-adjoint operators with $(A-B)$ Hilbert–Schmidt. Let f be a function defined on an interval $[a,b] \supset \sigma(A) \cup \sigma(B)$. Suppose f is uniformly Lipschitz, that is,*

$$\|f\|_L = \sup_{\substack{x,y\in[a,b]\\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|} < \infty. \tag{4.1}$$

Then $[f(A)-f(B)]$ is also Hilbert–Schmidt and

$$\|f(A)-f(B)\|_{\mathcal{I}_2} \leq \|f\|_L \|A-B\|_{\mathcal{I}_2}. \tag{4.2}$$

The proof in [13] depends on the deep machinery of double Stieltjes operator integrals. Our two points in this section are:

(1) The inequality for finite matrices is quite elementary and, by limits, extends to (4.2).

(2) The key to our proof, a kind of "Double Stieltjes Operator Integral for Dummies," goes back to Löwner [46] in 1934 whose contributions to this theme seem not to have been appreciated in the literature on double Stieltjes operator integrals.

Given two finite $m \times m$ self-adjoint matrices A, B with respective eigenvectors $\{\varphi_j\}_{j=1}^m$ and $\{\psi_j\}_{j=1}^m$ and eigenvalues $\{x_j\}_{j=1}^m$ and $\{y_j\}_{j=1}^m$ such that

$$A\varphi_j = x_j\varphi_j, \qquad B\psi_j = y_j\psi_j, \tag{4.3}$$

we introduce the (modified) Löwner matrix of a function f by

$$L_{k,\ell} = \begin{cases} \frac{f(y_k)-f(x_\ell)}{y_k-x_\ell}, & y_k \neq x_\ell, \\ 0, & y_k = x_\ell, \end{cases} \quad 1 \leq k,\ell \leq m. \tag{4.4}$$

(Löwner [46] originally supposed $y_k \neq x_\ell$ for all $1 \leq k,\ell \leq m$.) Clearly, if f is Lipschitz,

$$\sup_{1\leq k,\ell\leq m} |L_{k,\ell}| \leq \|f\|_L. \tag{4.5}$$

Löwner noted that since

$$f(A)\varphi_j = f(x_j)\varphi_j, \qquad f(B)\psi_j = f(y_j)\psi_j, \tag{4.6}$$

we have Löwner's formula:

$$\langle \psi_k, [f(B)-f(A)]\varphi_\ell\rangle = L_{k\ell}\langle\psi_k, (B-A)\varphi_\ell\rangle, \tag{4.7}$$

and this holds even if $y_k = x_\ell$ (since then both matrix elements vanish). This is the "baby" version of the double Stieltjes operator integral formula

$$f(B)-f(A) = \int_{\sigma(A)}\int_{\sigma(B)} \frac{f(y)-f(x)}{y-x}\, dE_B(x)(B-A)dE_A(y)$$

due to Birman and Solomyak [13, 15]. Here the integration is with respect to the spectral measures of A and B.

Löwner's formula immediately implies:

Proposition 4.2. (4.2) *holds for finite self-adjoint matrices.*

Proof. Hilbert–Schmidt norms can be computed in any basis, even two different ones, that is,

$$\|C\|_{\mathcal{I}_2}^2 = \sum_{\ell=1}^{m} \|C\varphi_\ell\|^2 = \sum_{k,\ell=1}^{m} |\langle \psi_k, C\varphi_\ell\rangle|^2.$$

Thus,

$$\begin{aligned}
\|f(A)-f(B)\|_{\mathcal{I}_2}^2 &= \sum_{k,\ell=1}^{m} |\langle \psi_k, (f(B)-f(A))\varphi_\ell\rangle|^2 \\
&= \sum_{k,\ell=1}^{m} L_{k\ell}^2 |\langle \psi_k, (B-A)\varphi_\ell\rangle|^2 && \text{by (4.7)} \\
&\le \|f\|_L^2 \sum_{k,\ell=1}^{m} |\langle \psi_k, (B-A)\varphi_\ell\rangle|^2 && \text{by (4.5)} \\
&= \|f\|_L^2 \|A-B\|_{\mathcal{I}_2}^2.
\end{aligned}$$

□

Proof of Theorem 4.1. Let $\{\zeta_j\}_{j=1}^\infty$ be an orthonormal basis for $\mathcal{H}$ and P_N the orthogonal projections onto the linear span of $\{\zeta_j\}_{j=1}^N$. For any A and B and Lipschitz f, by Proposition 4.2,

$$\begin{aligned}
\|f(P_NBP_N)-f(P_NAP_N)\|_{\mathcal{I}_2} &\le \|f\|_L \|P_N(B-A)P_N\|_{\mathcal{I}_2} && (4.8) \\
&\le \|f\|_L \|B-A\|_{\mathcal{I}_2} && (4.9)
\end{aligned}$$

if $B-A$ is Hilbert–Schmidt, since

$$\|P_N(B-A)P_N\|_{\mathcal{I}_2}^2 = \sum_{j=1}^{n} \|P_N(B-A)\zeta_j\|^2 \le \sum_{j=1}^{n} \|(B-A)\zeta_j\|^2 \tag{4.10}$$

$$\le \|B-A\|_{\mathcal{I}_2}^2. \tag{4.11}$$

Thus, for any $k \in \mathbb{N}$,

$$\sum_{j=1}^{k} \|[f(P_NBP_N)-f(P_NAP_N)]\zeta_j\|^2 \le \|f\|_L^2 \|B-A\|_{\mathcal{I}_2}^2. \tag{4.12}$$

As $P_N B P_N \underset{N\to\infty}{\longrightarrow} B$ strongly, one infers by continuity of the functional calculus that $f(P_N B P_N) \underset{N\to\infty}{\longrightarrow} f(B)$ strongly. Since the sum in (4.12) is finite, one concludes that

$$\sum_{j=1}^{k} \|(f(B)-f(A))\zeta_j\|^2 \le \|f\|_L^2 \|B-A\|_{\mathcal{I}_2}^2.$$

Taking $k\to\infty$, we see that $[f(B)-f(A)] \in \mathcal{I}_2$ and that (4.2) holds. □

5. General Construction of the KoSSF $\eta(\,\cdot\,; A, B)$

The general construction and proof of properties of η depends first on an approximation of trace class operators by finite rank ones and then on an approximation of Hilbert–Schmidt operators by trace class operators. In this section, we mostly follow the approach of [36, Lemma 3.3].

Theorem 5.1. *Let B_n, B, $n\in\mathbb{N}$, be uniformly bounded self-adjoint operators such that $B_n \underset{n\to\infty}{\longrightarrow} B$ strongly. Let X_n, X, $n\in\mathbb{N}$, be a sequence of self-adjoint trace class operators such that $\|X-X_n\|_{\mathcal{I}_1} \underset{n\to\infty}{\longrightarrow} 0$. Then for any continuous function, g, of compact support, we conclude that*

$$\int_{\mathbb{R}} g(\lambda)\eta(\lambda; B_n+X_n, B_n)\, d\lambda \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} g(\lambda)\eta(\lambda; B+X, B)\, d\lambda. \tag{5.1}$$

In particular, $\eta(\,\cdot\,; A, B)\ge 0$ a.e. on $\mathbb{R}$ if $(A-B)\in\mathcal{I}_1$ and, in that case, (3.15) *holds.*

Proof. By Theorem A.7 and

$$\int_{\mathbb{R}} g(\lambda)\bigg(\int_{-\infty}^{\lambda} \xi(\lambda'; A, B)\, d\lambda'\bigg)\, d\lambda = \int_{-\infty}^{\infty} \xi(\lambda'; A, B)\bigg(\int_{\lambda'}^{\infty} g(\lambda)\, d\lambda\bigg)\, d\lambda'$$

we get convergence of the second term in (2.16). By (2.10),

$$d\mu_{B_n,X_n} \underset{n\to\infty}{\longrightarrow} d\mu_{B,X}$$

weakly by the strong continuity of the functional calculus since

$$\begin{aligned}
&|\mathrm{Tr}(X_n f(B_n)) - \mathrm{Tr}(X f(B))| \\
&\quad \le |\mathrm{Tr}(X[f(B_n)-f(B)])| + \|f\|_\infty \|X-X_n\|_{\mathcal{I}_1} \underset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

as f is continuous.

Since weak limits of positive measures are positive, the positivity follows from positivity in the finite-dimensional case taking $B_n = P_n B P_n$ and $X_n = P_n X P_n$ for finite-dimensional P_n converging strongly to I, the identity operator.

Once we have positivity, we obtain (3.15) directly by following the proof of Theorem 3.4. □

Theorem 5.2. *Let A, B, C be bounded self-adjoint operators such that $(A - C) \in \mathcal{I}_1$ and $(B - C) \in \mathcal{I}_1$. Then*

$$\int_{\mathbb{R}} |\eta(\lambda; A, C) - \eta(\lambda; B, C)| \, d\lambda \leq \|A - B\|_{\mathcal{I}_2} \big[\tfrac{1}{2} \|A - B\|_{\mathcal{I}_2} + \|B - C\|_{\mathcal{I}_2}\big]. \tag{5.2}$$

Proof. We begin with (1.9) which follows from the fact that (2.17) holds when $(A - B) \in \mathcal{I}_1$. Here (1.10) holds for nice functions g, say, $g \in C^\infty(\mathbb{R})$. Thus,

$$\text{LHS of (5.2)} \leq \int_{\mathbb{R}} |\eta(\lambda; A, B)| \, d\lambda + \int_{\mathbb{R}} |\delta\eta(\lambda; A, B, C)| \, d\lambda. \tag{5.3}$$

By (3.15),

$$\text{First term on RHS of (5.3)} \leq \tfrac{1}{2} \|A - B\|_{\mathcal{I}_2} \|A - B\|_{\mathcal{I}_2}. \tag{5.4}$$

As for $\delta\eta$, by (1.10),

$$\begin{aligned} \left| \int_{\mathbb{R}} g'(\lambda) \delta\eta(\lambda) \, d\lambda \right| &\leq \|A - B\|_{\mathcal{I}_2} \|g(B) - g(C)\|_{\mathcal{I}_2} \\ &\leq \|g'\|_\infty \|A - B\|_{\mathcal{I}_2} \|B - C\|_{\mathcal{I}_2} \end{aligned}$$

by Theorem 4.1. Since $\delta\eta \in L^1(\mathbb{R})$ and the bounded $C^\infty(\mathbb{R})$-functions are $\|\cdot\|_\infty$-dense in the bounded continuous functions, and for $h \in L^1(\mathbb{R})$,

$$\|h\|_1 = \sup_{\substack{f \in C(\mathbb{R}) \\ \|f\|_\infty = 1}} \left| \int_{\mathbb{R}} f(x) h(x) \, dx \right|,$$

we conclude

$$\|\delta\eta\|_1 \leq \|A - B\|_{\mathcal{I}_2} \|B - C\|_{\mathcal{I}_2}. \tag{5.5}$$

Relations (5.3)–(5.5) imply (5.2). □

Here is the main theorem on the existence of the KoSSF:

Theorem 5.3. *Let A, B be two bounded self-adjoint operators with $(B - A) \in \mathcal{I}_2$. Then there exists a unique $L^1(\mathbb{R}; d\lambda)$-function $\eta(\,\cdot\,; A, B)$ supported on $(-\max(\|A\|, \|B\|), \max(\|A\|, \|B\|))$ such that for any $g \in C^\infty(\mathbb{R})$,*

$$\left(g(A) - g(B) - \frac{d}{d\alpha} g(B + \alpha(A - B)) \bigg|_{\alpha = 0} \right) \in \mathcal{I}_1 \tag{5.6}$$

and

$$\operatorname{Tr}\bigg(g(A)-g(B)-\frac{d}{d\alpha}\,g(B+\alpha(A-B))\bigg|_{\alpha=0}\bigg) = \int_{\mathbb{R}} g''(\lambda)\eta(\lambda;A,B)\,d\lambda. \tag{5.7}$$

Moreover,

$$\eta(\,\cdot\,;A,B)\geq 0 \ \text{ a.e. on } \mathbb{R}, \tag{5.8}$$

$$\int_{\mathbb{R}} |\eta(\lambda;A,B)|\,d\lambda = \tfrac{1}{2}\,\|A-B\|^2_{\mathcal{I}_2}, \tag{5.9}$$

and for any bounded self-adjoint operators A, B, C with $(A-C)\in\mathcal{I}_2$ and $(B-C)\in\mathcal{I}_2$,

$$\int_{\mathbb{R}} |\eta(\lambda;A,C)-\eta(\lambda;B,C)|\,d\lambda \leq \|A-B\|_{\mathcal{I}_2}\big[\tfrac{1}{2}\,\|A-B\|_{\mathcal{I}_2}+\|B-C\|_{\mathcal{I}_2}\big]. \tag{5.10}$$

Remark. For a sharp condition on the class of functions η for which Koplienko's trace formula holds, we refer to Peller [52].

Proof. Let $X=A-B$ and pick $X_n\in\mathcal{I}_1$, $n\in\mathbb{N}$, such that $\|X_n - X\|_{\mathcal{I}_2} \underset{n\to\infty}{\longrightarrow} 0$. By (5.10), which we have proven for

$$\eta(\lambda;B+X_n,B)-\eta(\lambda;B+X_m,B),$$

we see that $\eta(\,\cdot\,;B+X_n,B)$ is Cauchy in $L^1(\mathbb{R};d\lambda)$ and so converges a.e. to what we will define as $\eta(\,\cdot\,;A,B)$. (5.7) holds by taking limits; $\eta\geq 0$ as a limit of positive functions η. (5.9) and (5.10) hold by taking limits. □

Remark. Here is an alternative method of proving the estimate (5.2), bypassing Theorem 5.1: In the same way as in Corollary 3.3, one can deduce positivity of $\eta(\,\cdot\,;A,B)$ for $(A-B)\in\mathcal{I}_1$ from Theorem 3.6. The only place where Theorem 5.1 is used in the proof of Theorem 5.2 is in the estimate (5.4). This estimate (see Theorem 3.4) follows directly from the positivity of η and the trace formula (2.15).

6. What Functions η are Possible?

We introduce the classes of functions

$$\eta(\mathcal{I}_2)=\{\eta(\,\cdot\,;A,B)\mid A,B \text{ bounded and self-adjoint}, (A-B)\in\mathcal{I}_2\},$$
$$\eta(\mathcal{I}_1)=\{\eta(\,\cdot\,;A,B)\mid A,B \text{ bounded and self-adjoint}, (A-B)\in\mathcal{I}_1\}.$$

In this section we would like to raise the question of the description of the classes $\eta(\mathcal{I}_2)$ and $\eta(\mathcal{I}_1)$.

Since, for now, we are considering A, B bounded, η has compact support.

First we discuss the class $\eta(\mathcal{I}_2)$. By Theorem 5.3, all functions of this class are nonnegative and Lebesgue integrable. It would be interesting to see if the class $\eta(\mathcal{I}_2)$ contains all nonnegative Lebesgue integrable functions.

As a step towards answering this question, we give the following elementary result:

Theorem 6.1. *The class $\eta(\mathcal{I}_2)$ contains all nonnegative Riemann integrable functions of compact support.*

Remark. A contemporary account of the theory of Riemann integrable functions can be found, for instance, in Stein–Shakarchi [62].

Proof. First we consider a simple example. Let $a \in \mathbb{R}$ and $\varepsilon > 0$; consider the operators in $\mathbb{C}^2$ given by the diagonal 2×2 matrices $B = \operatorname{diag}(a - \varepsilon, a + \varepsilon)$ and $A = \operatorname{diag}(a + \varepsilon, a - \varepsilon)$. Then the KoSSF for this pair is given by (cf. (2.16)) $\eta(\lambda) = 2\varepsilon \chi_{(a-\epsilon, a+\epsilon)}(\lambda)$. We note that

$$\int_{\mathbb{R}} \eta(\lambda)\, d\lambda = 4\varepsilon^2 = \tfrac{1}{2}\mathrm{Tr}((A - B)^2),$$

in agreement with (5.9).

Next, suppose that $0 \leq \eta \in L^1(\mathbb{R}; d\lambda)$ is represented by the $L^1(\mathbb{R}; d\lambda)$-convergent series

$$\eta(\lambda) = \sum_{n=1}^{\infty} |I_n| \chi_{I_n}(\lambda), \tag{6.1}$$

where $I_n \subset \mathbb{R}$ are (not necessarily disjoint) finite intervals and $|I_n|$ is the length of I_n. Denote by a_n the midpoint of I_n and let $\varepsilon_n = \frac{1}{2}|I_n|$. We introduce

$$B = \oplus_{n=1}^{\infty} \operatorname{diag}(a_n - \varepsilon_n, a_n + \varepsilon_n) \ \text{ and } \ A = \oplus_{n=1}^{\infty} \operatorname{diag}(a_n + \varepsilon_n, a_n - \varepsilon_n)$$

in the Hilbert space $\oplus_{n=1}^{\infty} \mathbb{C}^2$. Note that the $L^1(\mathbb{R}; d\lambda)$-convergence of the series (6.1) is equivalent to the condition $\sum_{n=1}^{\infty} \varepsilon_n^2 < \infty$ and so $A - B = \oplus_{n=1}^{\infty} \operatorname{diag}(2\varepsilon_n, -2\varepsilon_n)$ is a Hilbert–Schmidt operator. It is clear that the KoSSF for the pair A, B coincides with η.

Thus, it suffices to prove that any Riemann integrable function $0 \leq \eta \in L^1(\mathbb{R}; d\lambda)$ can be represented as an $L^1(\mathbb{R}; d\lambda)$-convergent series (6.1).

Let $0 \leq \eta \in L^1(\mathbb{R}; d\lambda)$ be Riemann integrable. According to the definition of the Riemann integral, there exists a finite set of disjoint open squares Q_n, $n \in \{1, \dots, M\}$, which fit under the graph of η and

$$\sum_{n=1}^{M} \operatorname{area}(Q_n) \geq \tfrac{1}{2} \int_{\mathbb{R}} \eta(\lambda)\, d\lambda.$$

In other words, there exists a finite set of (not necessarily disjoint) open intervals $I_n \subset \mathbb{R}$, $n \in \{1, \dots, N\}$, such that

$$\begin{aligned}
&\sum_{n=1}^{N} |I_n| \chi_{I_n}(\lambda) \leq \eta(\lambda), \quad \lambda \in \mathbb{R}, \\
&\int_{\mathbb{R}} \bigg(\sum_{n=1}^{N} |I_n| \chi_{I_n}(\lambda) \bigg) d\lambda = \sum_{n=1}^{N} |I_n|^2 \geq \tfrac{1}{2} \int_{\mathbb{R}} \eta(\lambda)\, d\lambda.
\end{aligned}$$

Thus, we can represent η as

$$\begin{aligned}
&\eta(\lambda) = \sum_{n=1}^{N_1} |I_n| \chi_{I_n}(\lambda) + \eta_1(\lambda), \\
&\eta_1(\lambda) \geq 0, \quad \int_{\mathbb{R}} \eta_1(\lambda)\, d\lambda \leq \tfrac{1}{2} \int_{\mathbb{R}} \eta(\lambda)\, d\lambda,
\end{aligned}$$

and the sum is taken over a finite set of indices $n \in \{1, \dots, N_1\}$. Iterating this procedure, we see that for any $m \in \mathbb{N}$ we can represent η as

$$\begin{aligned}
&\eta(\lambda) = \sum_{n=1}^{N_m} |I_n| \chi_{I_n}(\lambda) + \eta_m(\lambda), \\
&\eta_m(\lambda) \geq 0, \quad \int_{\mathbb{R}} \eta_m(\lambda)\, d\lambda \leq 2^{-m} \int_{\mathbb{R}} \eta(\lambda)\, d\lambda,
\end{aligned}$$

where $\varepsilon_n > 0$, $a_n \in \mathbb{R}$, and the sum is taken over a finite set of indices n. Taking $m \to \infty$, it follows that η can be represented as an $L^1(\mathbb{R}; d\lambda)$-convergent series (6.1). $\square$

Regarding the class $\eta(\mathcal{I}_1)$, we note only that every function of this class is of bounded variation. This follows from (2.16), since both terms on the right-hand side of (2.16) are of bounded variation. We also note that it follows from the proof of Theorem 6.1 that the class $\eta(\mathcal{I}_1)$ contains all functions of the type

$$\eta(\lambda) = \sum_{n=1}^{\infty} |I_n| \chi_{I_n}(\lambda), \quad \sum_{n=1}^{\infty} |I_n| < \infty.$$

7. Modified Determinants and the KoSSF

In this section, as a preliminary to the next, we want to use our viewpoint to prove a formula for modified perturbation determinants in terms of the KoSSF originally derived by Koplienko [36]. We recall that one of Krein's motivating formulas for the KrSSF is (see (A.32)):

$$\det((A-z)(B-z)^{-1}) = \exp\biggl(\int_{\mathbb{R}} (\lambda-z)^{-1}\xi(\lambda)\,d\lambda\biggr), \quad z\in\mathbb{C}\backslash\mathbb{R}. \tag{7.1}$$

Here $\det(\cdot)$ is the Fredholm determinant defined on $I+\mathcal{I}_1$ (since $A-B=X\in\mathcal{I}_1$ implies $(A-z)(B-z)^{-1}-I = X(B-z)^{-1}\in\mathcal{I}_1$); see [27, 60].

We recall that for $C\in\mathcal{I}_1$, one can define $\det_2(\cdot)$ by

$$\det{}_2(I+C) = \det(I+C)e^{-\operatorname{Tr}(C)} \tag{7.2}$$

and that $C\mapsto \det_2(I+C)$ extends uniquely and continuously to $\mathcal{I}_2$, the Hilbert–Schmidt operators, although the right-hand side of (7.2) no longer makes sense (see [27, Ch. IV], [60, Ch. 9]). Our goal in this section is to prove the following formula first derived by Koplienko [36]:

Theorem 7.1. *Let B and X be bounded self-adjoint operators and $X\in\mathcal{I}_2$. Let $A=B+X$. Then for any $z\in\mathbb{C}\backslash\mathbb{R}$, $(A-z)(B-z)^{-1}\in I+\mathcal{I}_2$ and*

$$\det{}_2((A-z)(B-z)^{-1}) = \exp\biggl(-\int_{\mathbb{R}} \frac{\eta(\lambda;A,B)}{(\lambda-z)^2}\,d\lambda\biggr). \tag{7.3}$$

Proof. It suffices to prove (7.3) for $X\in\mathcal{I}_1$ since both sides are continuous in $\mathcal{I}_2$ norm and $\mathcal{I}_1$ is dense in $\mathcal{I}_2$. Continuity of the left-hand side follows from Theorem 9.2(c) of [60] and of the right-hand side by Theorem 5.2 above.

When $X\in\mathcal{I}_1$, we can use (7.2). Let

$$g_1(\lambda) = \mu_{B,X}((-\infty,\lambda)), \quad g_2(\lambda) = \int_{-\infty}^{\lambda} \xi(\lambda';A,B)\,d\lambda'. \tag{7.4}$$

By an integration by parts argument (using $g_2'=\xi$),

$$\int_{\mathbb{R}} \frac{\xi(\lambda)}{\lambda-z}\,d\lambda = \int_{\mathbb{R}} \frac{g_2(\lambda)}{(\lambda-z)^2}\,d\lambda. \tag{7.5}$$

By an integration by parts in a Stieltjes integral and by (2.10),

$$\operatorname{Tr}(X(B-z)^{-1}) = \int_{\mathbb{R}} \frac{1}{\lambda-z}\,d\mu_{B,X}(\lambda) \tag{7.6}$$

$$= \int_{\mathbb{R}} g_1(\lambda)\,\frac{1}{(\lambda-z)^2}\,d\lambda. \tag{7.7}$$

Thus, by (7.1) and (7.2),

$$\det{}_2(1+X(B-z)^{-1}) = \exp\biggl(-\int_{\mathbb{R}} (g_1(\lambda)-g_2(\lambda))(\lambda-z)^{-2}\,d\lambda\biggr),$$

which, given (2.16) is (7.3). □

8. On Boundary Values of Modified Perturbation Determinants $\det_2((A-z)(B-z)^{-1})$

By (7.1), if $(A-B)\in\mathcal{I}_1$, $\det((A-z)(B-z)^{-1})$ has a limit as $z\to\lambda+i0$ for a.e. $\lambda\in\mathbb{R}$ since $\int_{\mathbb{R}}(\nu-z)^{-1}\xi(\nu)\,d\nu$ is a difference of Herglotz functions. In this section, we will consider nontangential boundary values to the real axis of modified perturbation determinants

$$\det{}_2((A-z)(B-z)^{-1}), \quad z\in\mathbb{C}_+,$$

where $X=(A-B)\in\mathcal{I}_2$. Unlike the trace class, we will see nontangential boundary values may not exist a.e. on $\mathbb{R}$.

For notational simplicity in the remainder of this section, we now abbreviate KoSSF simply by η, that is, $\eta\equiv\eta(\,\cdot\,;A,B)$.

In contrast to the usual (trace class) SSF theory, we have the following nonexistence result for boundary values of modified perturbation determinants:

Theorem 8.1. *There exists a pair of self-adjoint operators A,B (in a complex, separable Hilbert space) such that $X=(A-B)\in\mathcal{I}_2$, $\sigma(B)$ is an interval, and for a.e. $\lambda\in\sigma(B)$, the nontangential limit $\lim_{z\to\lambda,\, z\in\mathbb{C}_+}\det_2(I+X(B-zI)^{-1})$ does not exist.*

Proof. By Theorems 6.1 and 7.1, the proof reduces to the following statement: There exists a Riemann integrable $0\le\eta\in L^1(\mathbb{R};d\lambda)$ with support being an interval such that for a.e. $\lambda\in\operatorname{supp}(\eta)$, the nontangential limit

$$\lim_{\substack{z\to\lambda\\ z\in\mathbb{C}_+}} \int_{\mathbb{R}} \frac{\eta(\lambda)\,d\lambda}{(\lambda-z)^2} \tag{8.1}$$

does not exist.

First we note that the existence of the limit in (8.1) at the point λ depends only on the behavior of $\eta(t)$ when t varies in a small neighborhood of λ. Thus, it suffices to construct $0\le\eta\in L^1(\mathbb{R};d\lambda)$ such that the limits (8.1) do not exist for a.e. $\lambda\in(-1,1)$; by shifting and scaling such a function η, one obtains the required statement for a.e. $\lambda\in\sigma(B)$.

Let us first obtain the required example of η defined on the unit circle $\partial\mathbb{D}$, and then transplant it onto the real line. By a well-known construction employing either lacunary series or Rademacher functions (see [21], [22, App. A], [70, I, p. 6]), there exists a power series $f(z) = \sum_{n=1}^{\infty} c_n z^n$, $|z| \leq 1$, such that $\sum_{n=1}^{\infty} |c_n| < \infty$ and for a.e. $z \in \partial\mathbb{D}$, the limit $\lim_{\zeta \to z} f'(\zeta)$ does not exist as ζ approaches z from inside of the unit disc along any nontangential trajectory. By construction, $\mathrm{Im}(f)$ is continuous on $\partial\mathbb{D}$ and

$$f(z) = \frac{1}{\pi} \int_{\partial\mathbb{D}} \frac{\mathrm{Im}(f(\zeta))}{(\zeta - z)}\, d\zeta, \quad f'(z) = \frac{1}{\pi} \int_{\partial\mathbb{D}} \frac{\mathrm{Im}(f(\zeta))}{(\zeta - z)^2}\, d\zeta, \quad |z| < 1.$$

Let $a > -\min_{\zeta \in \partial\mathbb{D}} \mathrm{Im}(f(\zeta))$ and set $v(\zeta) = \mathrm{Im}(f(\zeta)) + a$ if $|\arg \zeta| < \pi/2$ and $v(\zeta) = 0$ otherwise. Then $v \geq 0$ and v is piecewise continuous (with the possible discontinuities only for $\arg(\zeta) = \pm\pi/2$); in particular, v is Riemann integrable. Again by a localization argument, for a.e. $\theta \in (-\pi/2, \pi/2)$, the limit

$$\lim_{z \to e^{i\theta}} \int_{\partial\mathbb{D}} \frac{\mathrm{Im}(f(\zeta))}{(\zeta - z)^2}\, d\zeta$$

does not exist as z approaches $e^{i\theta}$ from inside of the unit disc along any nontangential trajectory.

It remains to transplant v from the unit circle onto the real line. Let $t = i\frac{1-\zeta}{1+\zeta}$, $w = i\frac{1-z}{1+z}$, and $\eta(t) = v(\zeta(t))$. Then $0 \leq \eta \in L^1(\mathbb{R}; d\lambda)$, $\mathrm{supp}\,(\eta) \subset (-1, 1)$, η is Riemann integrable, and

$$\int_{\partial\mathbb{D}} \frac{\mathrm{Im}(f(\zeta))}{(\zeta - z)^2}\, d\zeta = -\frac{(w+i)^2}{2i} \int_{-1}^{1} \frac{\eta(\lambda)\, d\lambda}{(\lambda - w)^2}.$$

Thus, the limit (8.1) does not exist for a.e. $\lambda \in (-1, 1)$. □

9. KoSSF for Unbounded Operators

In this section we briefly discuss the question of existence of KoSSF under the assumption

$$[(A - z)^{-1} - (B - z)^{-1}] \in \mathcal{I}_2 \tag{9.1}$$

instead of $(A - B) \in \mathcal{I}_2$. This question was studied in [49] and [52] (see also [36] for related issues).

First recall the invariance principle for the KrSSF. Assume that A, B are bounded self-adjoint operators and $(A - B) \in \mathcal{I}_1$. Let $\varphi = \overline{\varphi} \in C^\infty(\mathbb{R})$, $\varphi' \neq 0$ on $\mathbb{R}$. Then we have

$$\xi(\lambda; A, B) = \mathrm{sign}\,(\varphi')\, \xi(\varphi(\lambda); \varphi(A), \varphi(B)) + \mathrm{const} \ \text{ for a.e. } \lambda \in \mathbb{R}. \tag{9.2}$$

This is a consequence of Krein's trace formula (1.6). With an appropriate choice of normalization of KrSSF, the constant in the right-hand side of (9.2) vanishes.

When both $(A-B) \in \mathcal{I}_1$ and $[\varphi(A)-\varphi(B)] \in \mathcal{I}_1$, formula (9.2) is an easily verifiable *identity*. But when $[\varphi(A)-\varphi(B)] \in \mathcal{I}_1$ yet $(A-B) \notin \mathcal{I}_1$, this formula can be regarded as a *definition* of $\xi(\,\cdot\,; A, B)$.

In contrast to this, no explicit formula relating $\eta(\varphi(\cdot); \varphi(A), \varphi(B))$ to $\eta(\,\cdot\,; A, B)$ is known. The reason is simple: The definition of η involves not only a trace formula but a choice of interpolation $A(\theta)$ between B and A. For bounded self-adjoint operators, the choice $A(\theta) = (1-\theta)B + \theta A$, $\theta \in [0,1]$, is natural. But when one only has (9.1), what choice does one make? It is natural to define $A(\theta)$ by

$$(A(\theta)-z)^{-1} = (1-\theta)(B-z)^{-1} + \theta(A-z)^{-1}, \quad \theta \in [0,1]. \tag{9.3}$$

For this to be self-adjoint, we need $z \in \mathbb{R}$, which means we should have some real point in the intersection of the resolvent sets for A and B. Even if there were such a z, it is not unique and the interpolation will not be unique. Moreover, the convexity that led to $\eta \geq 0$ may be lost. The net result is that the situation, both after the work of others and our work, is less than totally satisfactory.

Let us discuss a certain surrogate of (9.2) for the KoSSF. The formulas below are a slight variation on the theme of the construction of [49].

First assume that A and B are bounded operators and $X = (A-B) \in \mathcal{I}_2$. Let $\delta \subset \mathbb{R}$ be an interval which contains the spectra of A and B and $\varphi \in C^\infty(\delta)$, $\varphi' \neq 0$. Denote $a = \varphi(A)$, $b = \varphi(B)$, $x = a-b$. By the Birman–Solomyak bound (1.7), we have $x \in \mathcal{I}_2$ and so both $\eta(\,\cdot\,; A, B)$ and $\eta(\,\cdot\,; a, b)$ are well defined. Let us display the corresponding trace formulas:

$$\operatorname{Tr}\biggl(f(A) - f(B) - \frac{d}{d\alpha} f(B+\alpha X)\bigg|_{\alpha=0}\biggr) = \int_{\mathbb{R}} \eta(\lambda; A, B) f''(\lambda)\, d\lambda, \tag{9.4}$$

$$\operatorname{Tr}\biggl(g(a) - g(b) - \frac{d}{d\alpha} g(b+\alpha x)\bigg|_{\alpha=0}\biggr) = \int_{\mathbb{R}} \eta(\mu; a, b) g''(\mu)\, d\mu. \tag{9.5}$$

Now suppose $f = g \circ \varphi$. In contrast to the corresponding calculation for the KrSSF, the left-hand sides of (9.4) and (9.5) are, in general, distinct. However, we can make the right-hand sides look similar if we introduce the following modified KoSSF:

$$\widetilde{\eta}(\lambda; A, B) = \eta(\varphi(\lambda); a, b)\, \frac{1}{\varphi'(\lambda)} - \int_{\lambda_0}^{\lambda} \eta(\varphi(t); a, b) \left(\frac{1}{\varphi'(t)}\right)' dt. \tag{9.6}$$

The choice of λ_0 above is arbitrary; it affects only the constant term in the definition of $\widetilde{\eta}$.

By a simple calculation involving integration by parts, we get

$$\int_{\mathbb{R}} \eta(\mu; a, b) g''(\mu)\, d\mu = \int_{\mathbb{R}} \widetilde{\eta}(\lambda; A, B) f''(\lambda)\, d\lambda, \quad f = g \circ \varphi \in C_0^\infty(\mathbb{R}). \tag{9.7}$$

Combining (9.5) and (9.7), we get the modified trace formula

$$\operatorname{Tr}\bigg(f(A) - f(B) - \frac{d}{d\alpha} f \circ \varphi^{-1}(b + \alpha x)\bigg|_{\alpha=0} \bigg) = \int_{\mathbb{R}} \widetilde{\eta}(\lambda; A, B) f''(\lambda)\, d\lambda \tag{9.8}$$

for all $f \in C_0^\infty(\mathbb{R})$. Precisely as for the KrSSF, one can treat (9.6) and (9.8) as the definition of a modified KoSSF $\widetilde{\eta}(\,\cdot\,; A, B)$.

We consider an example of this construction which might be useful in applications. Suppose that A and B are lower semibounded self-adjoint operators such that for some (and thus for all) $z \in \mathbb{C}\backslash(\sigma(A) \cup \sigma(B))$ the inclusion (9.1) holds. Choose $E \in \mathbb{R}$ such that $\inf \sigma(A + E) > 0$ and $\inf \sigma(B + E) > 0$. Take $\varphi(\lambda) = \frac{1}{\lambda+E}$ and let $a = (A + E)^{-1}$, $b = (B + E)^{-1}$, $x = a - b$. For $\lambda > -E$, define

$$\begin{aligned} \widetilde{\eta}(\lambda; A, B) = &-\eta((\lambda + E)^{-1}; a, b)(\lambda + E)^2 \\ &+ 2 \int_{-E}^{\lambda} \eta((t + E)^{-1}; a, b)(t + E)\, dt. \end{aligned} \tag{9.9}$$

Note that $\eta(\,\cdot\,; a, b)$ is integrable and $\eta(\lambda; a, b)$ vanishes for large λ and therefore the integral in (9.9) converges. Moreover, this definition ensures that $\widetilde{\eta}(\lambda; A, B) = 0$ for $\lambda < \inf(\sigma(A) \cup \sigma(B))$. Thus, it is natural to define

$$\widetilde{\eta}(\lambda; A, B) = 0 \ \text{ for } \ \lambda \leq -E. \tag{9.10}$$

The above calculations prove the following result:

Theorem 9.1. *Let A, B, a, b, x be as above. Then there exists a function $\widetilde{\eta}(\,\cdot\,; A, B)$ such that*

$$\int_{\mathbb{R}} \widetilde{\eta}(\lambda; A, B)(\lambda + E)^{-4} d\lambda < \infty \tag{9.11}$$

and $\widetilde{\eta}(\lambda; A, B) = 0$ for $\lambda < \inf(\sigma(A) \cup \sigma(B))$ and for all $f \in C_0^\infty(\mathbb{R})$ the following trace formula holds:

$$\operatorname{Tr}\bigg(f(A) - f(B) - \frac{d}{d\alpha} f((b + \alpha x)^{-1} - E)\bigg|_{\alpha=0} \bigg) = \int_{\mathbb{R}} \widetilde{\eta}(\lambda; A, B) f''(\lambda)\, d\lambda. \tag{9.12}$$

We note that condition (9.11) does not fix the linear term in the definition of $\widetilde{\eta}$ but (9.10) does.

In [49], a pair of self-adjoint operators A, B was considered under the assumption (9.1) alone (without the lower semiboundedness assumption). Another regularization of $\eta(\,\cdot\,; A, B)$ was suggested in this case. The construction of [49] is more intricate than the above calculation and uses KoSSF for unitary operators.

In [36], the assumption

$$(A - B)|A - iI|^{-1/2} \in \mathcal{I}_2$$

was used. This assumption is intermediate between $(A - B) \in \mathcal{I}_2$ and (9.1). Under this assumption, the trace formula (5.7) was proven with $0 \le \eta \in L^1(\mathbb{R}; (1 + \lambda)^{-\gamma} d\lambda)$ for any $\gamma > \frac{1}{2}$.

Finally, in [49], the assumption

$$(A - B)(A - iI)^{-1} \in \mathcal{I}_2$$

was used and formula (5.7) was proven with $\eta \in L^1(\mathbb{R}; (1 + \lambda^2)^{-2} d\lambda)$.

Note that the difference between the last two results and Theorem 9.1 is that in Theorem 9.1, a modified trace formula (9.12) is proven rather than the original formula (5.7). Theorem 9.1 is nothing but a change of variables in the trace formula for resolvents, whereas the abovementioned results of [36] and [49] require some work.

10. The Case of Unitary Operators

In this section, we want to briefly discuss a definition of η for a pair of unitaries. Once again, there is an issue of interpolation. If A and B are the unitaries,

$$A(\theta) = (1 - \theta)B + \theta A, \quad \theta \in [0, 1], \tag{10.1}$$

is not unitary, so we cannot define $f(A(\theta))$ for arbitrary C^∞-functions on $\partial\mathbb{D} = \{z \in \mathbb{C} \,|\, |z| = 1\}$. Neidhardt [49] (see also [52]) discussed one way of interpolating by writing $A = e^C$, $B = e^D$ for suitable C and D and interpolating, but there is considerable ambiguity in how to choose C, D as well as whether to look at $e^{\theta C + (1-\theta)D}$ or $e^{(1-\theta)D} e^{\theta C}$, etc.

Here, with Szegő's theorem as background [25], we want to discuss an alternative to Neidhardt's approach.

Lemma 10.1. *Let A, B be unitary with $(A - B) \in \mathcal{I}_2$. Then for any $n = 0, 1, 2, \dots$,*

$$\bigg(A^n - B^n - \frac{d}{d\theta} A(\theta)^n \bigg|_{\theta=0}\bigg) \in \mathcal{I}_1. \tag{10.2}$$

We have

$$\bigg\| A^n - B^n - \frac{d}{d\theta} A(\theta)^n \bigg|_{\theta=0} \bigg\|_{\mathcal{I}_1} \le \frac{n(n-1)}{2} \, \|A - B\|_{\mathcal{I}_2}^2. \tag{10.3}$$

In fact,

$$\textit{LHS of } (10.3) = o(n^2). \tag{10.4}$$

Proof. Let $X = A - B$. Then, by telescoping,

$$A(\theta)^n - B^n = \sum_{j=0}^{n-1} A(\theta)^j (\theta X) B^{n-1-j}. \tag{10.5}$$

Thus, since $\|A(\theta)\| \le 1$, $\|B\| = 1$,

$$\|A(\theta)^n - B^n\|_{\mathcal{I}_2} \le n|\theta| \, \|X\|_{\mathcal{I}_2} \tag{10.6}$$

and, of course,

$$\|A(\theta)^n - B^n\| \le 2. \tag{10.7}$$

Dividing (10.5) by θ and taking θ to zero yields

$$\frac{d}{d\theta} A(\theta)^n = \sum_{j=0}^{n-1} B^j \, X \, B^{n-1-j},$$

so

$$\text{LHS of } (10.2) = \sum_{j=0}^{n-1} (A^j - B^j) X \, B^{n-1-j}. \tag{10.8}$$

(10.3) is immediate since (10.8) and (10.6) implies

$$\text{LHS of } (10.3) \le \|X\|_2^2 \biggl(\sum_{j=0}^{n-1} j \biggr). \tag{10.9}$$

To get (10.4), we write $X = X_\varepsilon^{(1)} + X_\varepsilon^{(2)}$ where $\|X_\varepsilon^{(2)}\|_{\mathcal{I}_2} \le \varepsilon$ and $\|X_\varepsilon^{(1)}\|_{\mathcal{I}_1} < \infty$. Thus, (10.8) implies

$$\text{LHS of } (10.3) \le \varepsilon \|X\|_{\mathcal{I}_1} \frac{n(n-1)}{2} + 2n \|X_\varepsilon^{(1)}\|_{\mathcal{I}_1}$$

using (10.7) instead of (10.6). Dividing by n^2, taking $n \to \infty$, and then $\varepsilon \downarrow 0$, show

$$\limsup_{n\to\infty} n^{-2} \text{ LHS of } (10.3) = 0. \qquad \square$$

Theorem 10.2. *Let A and B be unitary so $(A - B) \in \mathcal{I}_2$. Then there exists a real distribution $\eta(\lambda; A, B)$ on $\partial\mathbb{D}$ so that for any polynomial $P(z)$, $\bigl[P(A) - P(B) - \frac{d}{d\theta} P(A(\theta))\bigr]\big|_{\theta=0} \in \mathcal{I}_2$ and*

$$\operatorname{Tr}\biggl(P(A) - P(B) - \frac{d}{d\theta} P(A(\theta)) \bigg|_{\theta=0} \biggr) = \int_0^{2\pi} P''(e^{i\theta}) \eta(e^{i\theta}; A, B) \, \frac{d\theta}{2\pi}. \tag{10.10}$$

Moreover, the moments of η satisfy

$$\int_0^{2\pi} e^{in\theta}\eta(e^{i\theta})\,\frac{d\theta}{2\pi} \underset{|n|\to\infty}{=} o(1). \tag{10.11}$$

Remarks. 1. As usual, we use $\int_0^{2\pi} f(e^{i\theta})\eta(e^{i\theta})\frac{d\theta}{2\pi}$ as shorthand for the distribution η acting on the function f.

2. As we will discuss, η is determined by (10.10) up to three real constants in an affine term.

3. For a sharp condition on the class of functions for which Neidhardt's version of Koplienko's trace formula for unitary operators holds, we refer to Peller [52].

Proof. Let c_n, $n \in \mathbb{Z}$, be defined by

$$c_n = \begin{cases} 0, & n = 0,1, \\ [n(n-1)]^{-1}\operatorname{Tr}\big(A^n - B^n - \frac{d}{d\theta}\,A(\theta)^n\big|_{\theta=0}\big), & n \geq 2, \\ \overline{c_{-n}}, & n \leq -1. \end{cases} \tag{10.12}$$

By Lemma 10.1, $c_n = o(1)$ as $n \to \infty$, so there is a distribution $\eta = \eta(\,\cdot\,; A, B)$ satisfying

$$c_n = \int_0^{2\pi} e^{i(n-2)\theta}\eta(e^{i\theta})\,\frac{d\theta}{2\pi}, \quad n \geq 2. \tag{10.13}$$

By (10.12), we have (10.10) for $P(z) = z^n$ for $n \geq 2$ and both sides are zero for $P(z) = z^m$, $m = 0, 1$. Thus, (10.10) holds for all polynomials. $\square$

For any $c_0 \in \mathbb{R}$, $c_1 \in \mathbb{C}$, we can add $c_0 + c_1 e^{i\theta} + \bar{c}_1 e^{-i\theta}$ to η without changing the right-hand side of (10.10). We wonder if η is always in $L^1(\partial\mathbb{D})$ with $\eta \geq 0$ for some choice of c_0 and c_1. The condition $c_n \to 0$ is, of course, consistent with $\eta \in L^1(\partial\mathbb{D})$.

11. Open Problems and Conjectures

While we have found some new aspects of η here and summarized much of the prior literature, there are many open issues. The most important one concerns properties of η and the invariance of the a.c. spectrum:

Conjecture 11.1. Suppose A, B are self-adjoint with $(A - B) \in \mathcal{I}_2$ and that on some interval $(a, b) \subset \sigma(A) \cap \sigma(B)$, we have $\eta(\,\cdot\,; A, B)$ and $\eta(\,\cdot\,; B, A)$ are of bounded variation with distributional derivatives in $L^p((a, b); d\lambda)$ on (a, b) for some $p > 1$. Then $\sigma_{\mathrm{ac}}(A) \cap (a, b) = \sigma_{\mathrm{ac}}(B) \cap (a, b)$.

In the appendix, we prove the invariance for $\mathcal{I}_1$-perturbations using boundary values of $\det((A-z)(B-z)^{-1})$. When η has the properties in the conjecture, $\det_2((A-z)(B-z)^{-1})$ has boundary values and we hope those can be used to get the invariance of a.c. spectrum. While we made the conjecture assuming control of $\eta(\,\cdot\,;A,B)$ and $\eta(\,\cdot\,;B,A)$, we wonder if only one suffices. Similarly, we wonder if L^p, $p>1$, can be replaced by the weaker condition that the derivative is a sum of an L^1-piece and the Hilbert transform of an L^1-piece.

Open Question 11.2. *Is the η we constructed in Section 10 for the unitary case an $L^1(\partial\mathbb{D})$ function?*

Open Question 11.3. *Is the class $\eta(\mathcal{I}_2)$ introduced in Section 6 all of $L^1(\mathbb{R};d\lambda)$ (of compact support), or only the Riemann integrable functions, or something in between?*

Open Question 11.4. *Is the class $\eta(\mathcal{I}_1)$ all functions of bounded variation or a subset, and if so, what subset?*

Appendix: On the KrSSF $\xi(\,\cdot\,;A,B)$

Both for comparison and because the Krein spectral shift (KrSSF) is needed in our construction of the KoSSF, we present the basics of the KrSSF here. Most of the results in this appendix are known (see, e.g., [7, Sect. 19.1.4]), [14], [16], [38], [39], [40], [61], [64], [68, Ch. 8], [69] and the references therein) so this appendix is largely pedagogical, but our argument proving the invariance of a.c. spectrum under trace class perturbations at the end of this appendix is new. Moreover, we fill in the details of an approach sketched in [60, Ch. 11] exploiting the method Gesztesy–Simon [26] used to construct the rank-one KrSSF. Most approaches define ξ via perturbation determinants.

We will need the following strengthening of Theorem 2.2:

Theorem A.1. *Let f be a function of compact support whose Fourier transform $\widehat{f}$ satisfies* (2.6) *for $n=1$ (in particular, f can be $C^{2+\varepsilon}(\mathbb{R})$). Then,*

(a) *For any bounded self-adjoint operators A,B with $(A-B)\in\mathcal{I}_1$, $(f(A)-f(B))\in\mathcal{I}_1$. Moreover,*

$$\|f(A)-f(B)\|_{\mathcal{I}_1}\leq\|k\widehat{f}\|_1\|A-B\|_{\mathcal{I}_1}, \tag{A.1}$$

where

$$\|k\widehat{f}\|_1\leq\int_{\mathbb{R}}k|\widehat{f}(k)|\,dk. \tag{A.2}$$

(b) *Let B_n, B, $n \in \mathbb{N}$, be uniformly bounded self-adjoint operators such that $B_n \underset{n\to\infty}{\longrightarrow} B$ strongly. Let X_n, X, $n \in \mathbb{N}$, be a sequence of self-adjoint trace class operators such that $\|X - X_n\|_{\mathcal{I}_1} \underset{n\to\infty}{\longrightarrow} 0$. Then,*

$$\operatorname{Tr}(f(B_n + X_n) - f(B_n)) \underset{n\to\infty}{\longrightarrow} \operatorname{Tr}(f(B + X) - f(B)). \tag{A.3}$$

Proof. (a) is immediate from Proposition 2.1 which implies

$$\begin{aligned} &f(A) - f(B) \\ &\quad = (2\pi)^{-1/2} \int_{\mathbb{R}} ik\widehat{f}(k) \bigg[\int_0^1 e^{i\beta kA}(A-B)e^{i(1-\beta)kB}\, d\beta\bigg]\, dk. \end{aligned} \tag{A.4}$$

This also implies (b) via the dominated convergence theorem, continuity of the functional calculus (so $C_n \underset{n\to\infty}{\longrightarrow} C$ strongly implies $e^{itC_n} \underset{n\to\infty}{\longrightarrow} e^{itC}$ strongly), and the fact that if $X_n \underset{n\to\infty}{\longrightarrow} X$ in $\mathcal{I}_1$ and $C_n \underset{n\to\infty}{\longrightarrow} C$ strongly (with C_n, C uniformly bounded), then $\operatorname{Tr}(C_n X_n) \underset{n\to\infty}{\longrightarrow} \operatorname{Tr}(CX)$. This latter fact comes from

$$\begin{aligned} |\operatorname{Tr}(C_n X_n - CX)| &\le |\operatorname{Tr}(C_n(X_n - X) - (C - C_n)X)| \\ &\le \|C_n\|\, \|X_n - X_1\|_{\mathcal{I}_1} + |\operatorname{Tr}((C - C_n)X)| \end{aligned}$$

and if $X = \sum_{m\in\mathbb{N}} \mu_m(X)\langle\varphi_m, \cdot\,\rangle\psi_m$, then

$$|\operatorname{Tr}((C_n - C)X)| \le \sum_{m\in\mathbb{N}} \mu_m(X)|\langle\varphi_m, (C_n - C)\psi_m\rangle| \underset{n\to\infty}{\longrightarrow} 0$$

by the dominated convergence theorem. ☐

Part (a) in Theorem A.1, in a slightly more general form, is stated and proved in [40, p. 141].

Now let B be a bounded self-adjoint operator and φ a unit vector. For $\alpha \in \mathbb{R}$, define

$$A_\alpha = B + \alpha(\varphi, \cdot\,)\varphi \tag{A.5}$$

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

$$F_\alpha(z) = (\varphi, (A_\alpha - z)^{-1}\varphi), \tag{A.6}$$

$$G_\alpha(z) = 1 + \alpha F_0(z). \tag{A.7}$$

The resolvent formula implies (see [60, Sect. 11.2])

$$F_\alpha(z) = \frac{F_0(z)}{1 + \alpha F_0(z)}, \quad z \in \mathbb{C}\backslash\mathbb{R}, \tag{A.8}$$

and that

$$\begin{aligned}(A_\alpha - z)^{-1} &- (B-z)^{-1} \\ &= -\frac{\alpha}{1+\alpha F_0(z)}\,((B-\bar z)^{-1}\varphi,\,\cdot\,)(B-z)^{-1}\varphi, \quad z\in\mathbb{C}\backslash\mathbb{R},\end{aligned} \tag{A.9}$$

implying

$$\begin{aligned}\operatorname{Tr}((B-z)^{-1} - (A_\alpha - z)^{-1}) &= \frac{\alpha}{1+\alpha F_0(z)}\,(\varphi,(B-z)^{-2}\varphi) \\ &= \frac{d}{dz}\log(G_\alpha(z)), \quad z\in\mathbb{C}\backslash\mathbb{R}.\end{aligned} \tag{A.10}$$

Theorem A.2. *Let B be a bounded self-adjoint operator and A_α given by* (A.5) *for $\alpha\in\mathbb{R}$ and φ with $\|\varphi\|=1$. Then for a.e. $\lambda\in\mathbb{R}$,*

$$\xi_\alpha(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\downarrow 0}\arg(G_\alpha(\lambda+i\varepsilon)) \tag{A.11}$$

exists and satisfies

(i)
$$0 \le \pm\xi_\alpha(\,\cdot\,) \le 1 \ \ \textit{if } 0 < \pm\alpha. \tag{A.12}$$

(ii) $\xi_\alpha(\lambda)=0$ *if* $\lambda\le\min(\sigma(A_\alpha)\cup\sigma(B))$ *or* $\lambda\ge\max(\sigma(A_\alpha)\cup\sigma(B))$.

(iii)
$$\int |\xi_\alpha(\lambda)|\,d\lambda = |\alpha| \tag{A.13}$$

(iv) *For any $z\in\mathbb{C}\backslash\mathbb{R}$,*

$$G_\alpha(z) = \exp\biggl(\int_{\mathbb{R}} (\lambda-z)^{-1}\xi_\alpha(\lambda)\,d\lambda\biggr). \tag{A.14}$$

(v)
$$\det((A_\alpha - z)(B-z)^{-1}) = G_\alpha(z). \tag{A.15}$$

(vi) *For any $z\in\mathbb{C}\backslash\mathbb{R}$,*

$$\operatorname{Tr}((B-z)^{-1} - (A_\alpha - z)^{-1}) = \int_{\mathbb{R}} (\lambda-z)^{-2}\xi_\alpha(\lambda)\,d\lambda. \tag{A.16}$$

(vii) *For any f satisfying the hypotheses of Theorem A.1,*

$$\operatorname{Tr}(f(A_\alpha) - f(B)) = \int_{\mathbb{R}} f'(\lambda)\xi_\alpha(\lambda)\,d\lambda. \tag{A.17}$$

Remarks. 1. This theorem and its proof are essentially the same as the starting point of Krein's construction in [38] (see also [40, p. 134–136] or [16, Sect. 3]).

2. In (A.11), $\arg(G_\alpha(z))$ is defined uniquely for $\operatorname{Im}(z)>0$ by demanding continuity in z and

$$\lim_{y\uparrow\infty}\arg(G_\alpha(iy)) = 0. \tag{A.18}$$

For $\mathrm{Im}(z) < 0$ one has $\overline{G_\alpha(\overline{z})} = G_\alpha(z)$.

3. By (A.9), $(A_\alpha - z)(B - z)^{-1}$ is of the form $I+$ rank one, and so lies in $I + \mathcal{I}_1$. The $\det(\cdot)$ in (A.15) is the Fredholm determinant (see [60, Ch. 3]). This is the same as the finite-dimensional determinant $\det(C)$ for $I + D$ with D finite rank and $C = (I + D) \restriction \mathcal{K}$ where $\mathcal{K}$ is any finite-dimensional space containing $\mathrm{ran}(D)$ and $(\ker(D))^\perp$.

4. The exponential Herglotz representation basic to this proof goes back to Aronszajn and Donoghue [5].

5. Comparing (A.17) and (1.6), one concludes

$$\xi_\alpha(\,\cdot\,) = \xi(\,\cdot\,; A, B).$$

Proof. By the spectral theorem, there is a probability measure $d\mu_\alpha(\lambda)$ such that

$$F_\alpha(z) = \int_{\mathbb{R}} \frac{d\mu_\alpha(\lambda)}{\lambda - z}. \tag{A.19}$$

In particular,

$$\mathrm{Im}(F_0(z)) > 0 \ \text{ if } \ \mathrm{Im}(z) > 0, \tag{A.20}$$

so on $\mathbb{C}_+ = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$,

$$\pm \mathrm{Im}(G_\alpha(z)) > 0 \ \text{ if } \ \pm \alpha > 0. \tag{A.21}$$

Since $G_\alpha(iy) \to 1$, as $y \uparrow \infty$, we can define

$$\log(G_\alpha(z)) = H_\alpha(z)$$

on $\mathbb{C}_+$ uniquely if we require $H_\alpha(iy) \underset{y\uparrow\infty}{\longrightarrow} 0$. By (A.21),

$$0 < \pm \mathrm{Im}(H_\alpha(\,\cdot\,)) \leq \pi. \tag{A.22}$$

By the general theory of Herglotz functions (see, e.g., [4], [5]), the limit in (A.11) exists and (A.12) holds by (A.22). (A.22) also implies that the limiting measure $\text{w-lim}_{\varepsilon\downarrow 0} \pm\frac{1}{\pi} \mathrm{Im}(H_\alpha(\lambda+i\varepsilon))\, d\lambda$ in the Herglotz representation theorem is purely absolutely continuous, hence (A.14) holds.

(A.16) then follows from (A.14) and (A.10).

Since

$$F_\alpha(z) \underset{z\to\infty}{=} -z^{-1} + O(z^{-2}), \tag{A.23}$$

(A.17) implies

$$G_\alpha(z) \underset{z\to\infty}{=} 1 - \alpha z^{-1} + O(z^{-2}), \tag{A.24}$$

and thus (A.14) implies

$$\int_{\mathbb{R}} \xi_\alpha(\lambda)\, d\lambda = \alpha, \tag{A.25}$$

which, given (A.12), implies (A.13). This proves everything except the parts (ii), (v), and (vii).

To prove (A.15), we note that with $P_\varphi = (\varphi, \cdot\,)\varphi$, we have

$$(A_\alpha - z)(B-z)^{-1} = I + \alpha P_\varphi (B-z)^{-1},$$

which, since P_φ is rank one, implies

$$\begin{aligned}\det((A_\alpha - z)(B-z)^{-1}) &= 1 + \operatorname{Tr}(\alpha P_\varphi (B-z)) \\ &= 1 + \alpha F_0(z) \\ &= G_\alpha(z).\end{aligned}$$

Let us prove (ii) for $\alpha > 0$. The proof of $\alpha < 0$ is similar. Let $a = \min(\sigma(B))$, $b = \max(\sigma(B))$. Then, by (A.19),

$$F_0'(x) = \int_{\mathbb{R}} \frac{d\mu_0(\lambda)}{(\lambda - x)^2} > 0$$

on $(-\infty, a) \cup (b, \infty)$ and $F_0 \underset{x\to\pm\infty}{\longrightarrow} 0$. Thus, $F > 0$ on $(-\infty, a)$ and $F < 0$ on (b, ∞). Let $f = \lim_{x\downarrow b} F(x)$ which may be $-\infty$. If $1 + \alpha f < 0$, there is a unique c with $1 + \alpha F_0(c) = 0$, and then G_α is positive on (c, ∞). By (A.8), $F_\alpha(z)$ is analytic away from $(a,b) \cup \{c\}$. Thus, $\sigma(A_\alpha) \in (a,b) \cup \{c\}$ and $c = \max(\sigma(A_\alpha), \sigma(B))$, so (ii) says that $\xi_\alpha(\lambda) = 0$ on $(-\infty, b)$ and (c, ∞). Since $G_\alpha(x) > 0$ there and $0 < \arg(G_\alpha(z + i\varepsilon)) < \pi$, we see that $\xi_\alpha(x) = 0$ on these intervals.

Finally, we turn to (vii). Since $B^n - A_\alpha^n$ can be written as a telescoping series, it is trace class and

$$\|B^n - A_\alpha^n\|_{\mathcal{I}_1} \leq n\,[\sup(\|A_\alpha\|, \|B\|)]^{n-1} \|B - A\|_{\mathcal{I}_1}. \tag{A.26}$$

Thus, both sides of (A.16) are analytic about $z = \infty$, so identifying Taylor coefficients,

$$\operatorname{Tr}(B^n - A_\alpha^n) = \int_{\mathbb{R}} n\lambda^{n-1} \xi_\alpha(\lambda)\, d\lambda. \tag{A.27}$$

Summing Taylor series for $e^{z\lambda}$, using (A.26) and (A.27) proves $(e^{zB} - e^{zA_\alpha}) \in \mathcal{I}_1$ and

$$\operatorname{Tr}(e^{zB} - e^{zA_\alpha}) = z \int_{\mathbb{R}} e^{z\lambda} \xi_\alpha(\lambda)\, d\lambda. \tag{A.28}$$

This leads to (A.17) by using (A.4). $\square$

In extending this, the following uniqueness result will be useful:

Proposition A.3. *Suppose A and B are bounded self-adjoint operators and $(A-B)\in\mathcal{I}_1$. Suppose $\xi_j\in L^1(\mathbb{R};d\lambda)$ for $j=1,2$, and for all $f\in C_0^\infty(\mathbb{R})$,*

$$\operatorname{Tr}(f(A)-f(B))=\int_{\mathbb{R}} f'(\lambda)\xi_j(\lambda)\,d\lambda. \tag{A.29}$$

Then $\xi_1=\xi_2$. Moreover, if $(a,b)\subset\mathbb{R}\backslash\sigma(A)\cup\sigma(B)$, $\xi_j(\cdot)$ is an integer on (a,b), and if $a=-\infty$ or $b=\infty$, it is zero on (a,b), and so ξ_j has compact support.

Proof. By (A.29), the distribution $\xi_1-\xi_2$ has vanishing distributional derivative, so is constant. Since it lies in $L^1(\mathbb{R};d\lambda)$, it must be zero.

If $f\in C_0^\infty((a,b))$, $f(A)=f(B)=0$, so ξ_j' has zero derivative on (a,b) and so is constant. If $a=-\infty$ or $b=\infty$, the constant must be zero since $\xi_j\in L^1(\mathbb{R};d\lambda)$. Now pick f which is supported on $(c,(a+b)/2)$ for some $c<d<\min(\sigma(A)\cup\sigma(B))$ with $f=1$ on $(d,(3a+b)/4)$. Thus, the right-hand side of (A.29) is the negative of the constant value of ξ_j on (a,b), while the left-hand side is the trace of a trace class difference of projections which is always an integer (see [6, 23]). ☐

Theorem A.4. *For any pair of bounded self-adjoint operators A,B with $(A-B)$ of finite rank, there exists a function, $\xi(\,\cdot\,;A,B)$ such that the following hold:*

(i) (A.17) *holds for any f satisfying the hypotheses of Theorem A.1.*

(ii)

$$|\xi(\,\cdot\,;A,B)|\leq\operatorname{rank}(A-B). \tag{A.30}$$

(iii)

$$\int_{\mathbb{R}}|\xi(\lambda;A,B)|\,d\lambda\leq\|A-B\|_{\mathcal{I}_1}. \tag{A.31}$$

(iv) *For $z\in\mathbb{C}\backslash\mathbb{R}$, one has*

$$\det((A-z)(B-z)^{-1})=\exp\bigg(\int_{\mathbb{R}}(\lambda-z)^{-1}\xi(\lambda)\,d\lambda\bigg). \tag{A.32}$$

(v) *$\xi(\lambda)=0$ for $\lambda\leq\min(\sigma(A)\cup\sigma(B))$ or $\lambda\geq\max(\sigma(A)\cup\sigma(B))$.*

(vi) *If $(A-B)$ and $(B-C)$ are both finite rank,*

$$\xi(\,\cdot\,;A,C)=\xi(\,\cdot\,;A,B)+\xi(\,\cdot\,;B,C). \tag{A.33}$$

Proof. If $(A-B)$ has rank n, we can find $A_0=A,A_1,\dots,A_n=B$ so $(A_{j+1}-A_j)$ has rank one, and

$$\sum_{j=0}^{n-1}\|A_{j+1}-A_j\|_{\mathcal{I}_1}=\|B-A\|_{\mathcal{I}_1}. \tag{A.34}$$

We define

$$\xi(\,\cdot\,; A, B) = \sum_{j=0}^{n-1} \xi(\,\cdot\,; A_j, A_{j+1}), \tag{A.35}$$

where $\xi(\,\cdot\,; A_j, A_{j+1})$ is constructed via Theorem A.2. (A.17) holds by telescoping and the rank-one case. (A.30) and (A.31) follow from (A.12), (A.13), and (A.34).

(A.32) follows from

$$(A - z)(B - z)^{-1} = [(A_0 - z)(A_1 - z)^{-1}][(A_1 - z)(A_2 - z)^{-1}] \dots$$

using

$$\det((1 + X_1)(1 + X_2)) = \det(1 + X_1) \det(1 + X_2)$$

for $X_1, X_2 \in \mathcal{I}_1$.

Item (v) is proven in Proposition A.3. Item (vi) follows from the uniqueness in Proposition A.3. ☐

Theorem A.4 is essentially the same as Theorem 3 in [38] (see also [40] and [16]).

Corollary A.5. *If A, A' are both finite rank perturbations of B with all three operators self-adjoint, we have*

$$\int_{\mathbb{R}} |\xi(\lambda; A, B) - \xi(\lambda; A', B)| \, d\lambda \le \|A - A'\|_{\mathcal{I}_1}. \tag{A.36}$$

Proof. By (A.33),

$$\xi(\,\cdot\,; A, B) - \xi(\,\cdot\,; A', B) = \xi(\,\cdot\,; A, A').$$

Thus, (A.36) follows from (A.31). ☐

This yields the principal result on existence and properties of the KrSSF (see [38] or [40]).

Theorem A.6. *Let A, B be bounded self-adjoint operators with $(A - B) \in \mathcal{I}_1$. Then,*

(i) *There exists a unique function $\xi(\,\cdot\,; A, B) \in L^1(\mathbb{R}; d\lambda)$ such that* (A.17) *holds for any f satisfying the hypotheses of Theorem A.1.*

(ii)
$$\int_{\mathbb{R}} |\xi(\lambda; A, B)| \, d\lambda \le \|A - B\|_{\mathcal{I}_1}. \tag{A.37}$$

(iii) (A.32) *holds.*

(iv) *$\xi(\lambda) = 0$ if $\lambda \le \min(\sigma(A) \cup \sigma(B))$ or $\lambda \ge \max(\sigma(A) \cup \sigma(B))$.*

(v) *If $(A - B)$ and $(B - C)$ are both trace class,* (A.33) *holds.*

(vi) *If $(A - B)$ and $(A' - B)$ are trace class,* (A.36) *holds.*

Proof. Find A_n so $(A_n - B) \underset{n\to\infty}{\longrightarrow} (A - B)$ in $\mathcal{I}_1$ and $(A_n - B)$ is finite rank. By (A.36), $\xi(\,\cdot\,; A_n, B)$ is Cauchy in $L^1(\mathbb{R})$ so converges to an $L^1(\mathbb{R})$ function by (A.36). Thus, items (i), (ii), (iii), (v), and (vi) hold by taking limits (using $\|\cdot\|_{\mathcal{I}_1}$-continuity of the mapping $C \to \det(I+C)$. Uniqueness and (iv) follow from Proposition A.3. □

We refer to [50] (see also [51]) for a description of a class of functions f for which this theorem holds.

We note that there are interesting extensions of the trace formula (A.17) to classes of operators A, B different from self-adjoint or unitary operators. While we cannot possibly list all such extensions here, we refer, for instance, to Adamjan and Neidhardt [1], Adamjan and Pavlov [2], Jonas [29], [30], Krein [41], Langer [44], Neidhardt [47], [48], Rybkin [56], Sakhnovich [57], and the literature cited therein.

Theorem A.7. *Let B_n, B, $n \in \mathbb{N}$, be uniformly bounded self-adjoint operators such that $B_n \underset{n\to\infty}{\longrightarrow} B$ strongly. Let X_n, X, $n \in \mathbb{N}$, be a sequence of self-adjoint trace class operators such that $\|X - X_n\|_{\mathcal{I}_1} \underset{n\to\infty}{\longrightarrow} 0$. Then for any continuous function, g,*

$$\int_{\mathbb{R}} g(\lambda)\xi(\lambda; B_n + X_n, B_n)\, d\lambda \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} g(\lambda)\xi(\lambda; B + X, B)\, d\lambda. \quad \text{(A.38)}$$

Proof. By Theorem A.1, we have (A.38) for $g \in C_0^\infty(\mathbb{R})$. Note that the $\|\cdot\|_\infty$-norm closure of C_0^∞ includes the continuous functions of compact support. Thus, by an approximation argument using uniform $L^1(\mathbb{R}; d\lambda)$-bounds on ξ, we get (A.38) for continuous functions of compact support. Since $\xi(\lambda; A, B) = 0$ for $\lambda \in [-\max(\|A\|, \|B\|), \max(\|A\|, \|B\|)]$, the result for continuous functions of compact support extends to any continuous function. □

We want to note the following. Define

$$\xi(\mathcal{I}_1) = \{\xi(\,\cdot\,; A, B) \mid A, B \text{ bounded and self-adjoint}, (A - B) \in \mathcal{I}_1\}.$$

Proposition A.8. *$\xi(\mathcal{I}_1)$ is the set of $L^1(\mathbb{R}; d\lambda)$-elements of compact support.*

Proof. Since A, B are bounded and self-adjoint, any $\xi(\,\cdot\,; A, B) \in \xi(\mathcal{I}_1)$ necessarily lies in $L^1(\mathbb{R}; d\lambda)$ and has compact support (cf. Theorem A.6 (i) and (iv)).

Next, let $g \in L^1(\mathbb{R}; d\lambda)$ satisfy $0 \le g(\lambda) \le 1$ and $\operatorname{supp}(g) \subset (a, b)$ for some $-\infty < a < b < \infty$. Define

$$G(z) = \exp\biggl(\frac{1}{\pi}\int_a^b \frac{g(\lambda)\, d\lambda}{\lambda - z}\biggr), \quad \operatorname{Im}(z) > 0. \quad \text{(A.39)}$$

Then G satisfies the following items (i)–(iii):

(i) $\mathrm{Im}(G(z)) > 0$ (for $\mathrm{Im}(z) > 0$) since
$$0 \leq \mathrm{Im}\biggl(\int_a^b \frac{g(\lambda)\,d\lambda}{\lambda - z}\biggr) \leq \mathrm{Im}\biggl(\int_a^b \frac{d\lambda}{\lambda - z}\biggr) \leq \pi$$
on account of $0 \leq g \leq 1$.

(ii) $\mathrm{Im}(G(\lambda + i0)) = 0$ if $\lambda < a$ or $\lambda > b$.

(iii) $G(z) \to 1$ as $\mathrm{Im}(z) \to \infty$ since $g \in L^1(\mathbb{R}; d\lambda)$. It follows that there is $\alpha > 0$ and a probability measure $d\mu$ on $[a, b]$ with
$$G(z) = 1 + \alpha \int_a^b \frac{d\mu(\lambda)}{\lambda - z}. \tag{A.40}$$

Let B be multiplication by λ on $L^2((a,b); d\mu)$, φ is the function 1 in $L^2((a,b); d\mu)$ and $A = B + \alpha(\varphi, \cdot\,)\varphi$. Then, by (A.5), (A.6), (A.7), and (A.14), $\xi(\lambda; A, B) = \pi^{-1} g(\lambda)$ for a.e. $\lambda \in (a, b)$, and $\alpha = \pi^{-1} \int_a^b g(\lambda)\,d\lambda$. Thus, we have the theorem if $0 \leq g \leq 1$ or (by interchanging A and B) if $0 \geq g \geq -1$. Since any $L^1(\mathbb{R}; d\lambda)$-function is a sum of such g's converging in $L^1(\mathbb{R}; d\lambda)$ (simple functions are dense in $L^1(\mathbb{R}; d\lambda)$), we obtain the general result. $\square$

We note that a similar result for the finite rank case can be found in [42].

Finally, we prove invariance of the absolutely continuous spectrum under trace class perturbations using the KrSSF and perturbation determinants, that is, without directly relying on elements from scattering theory.

We start with the following observations:

Lemma A.9. *Let A, B be bounded self-adjoint operators with $X = (A - B)$ of rank one. Then,*
$$\sigma_{\mathrm{ac}}(A) = \sigma_{\mathrm{ac}}(B)$$
and
$$\xi(\lambda; A, B) \in \{-1, 0, 1\}$$
for a.e. $\lambda \in \mathbb{R}\backslash\sigma_{\mathrm{ac}}(B)$.

Proof. $\xi(\lambda; A, B) \in \{-1, 0, 1\}$ follows from (A.5)–(A.7), (A.11), and (A.12). $\sigma_{\mathrm{ac}}(A) = \sigma_{\mathrm{ac}}(B)$ follows in the usual manner by computing the normal boundary values to the real axis of the imaginary part of F_α in terms of that of F_0 using (A.8). $\square$

Lemma A.10. *Let A, B be bounded self-adjoint operators with $X = (A - B) \in \mathcal{I}_1$. Then for a.e. $\lambda \in \mathbb{R}\backslash\sigma_{\mathrm{ac}}(B)$ one has*
$$\lim_{\varepsilon\downarrow 0} \det(I + X(B - \lambda - i\varepsilon)^{-1}) \in \mathbb{R}.$$

Proof. By (A.32), it suffices to prove that

$$\xi(\,\cdot\,; A, B) \in \mathbb{Z} \text{ a.e. on } \mathbb{R}\backslash\sigma_{\rm ac}(B). \tag{A.41}$$

Introducing

$$X = \sum_{n=1}^{\infty} x_n(\phi_n, \,\cdot\,)\phi_n, \quad X_0 = 0,\ X_N = \sum_{n=1}^{N} x_n(\phi_n, \,\cdot\,)\phi_n,\ N \in \mathbb{N},$$

the rank-by-rank construction of $\xi(\,\cdot\,; A, B)$ alluded to in the proof of Theorem A.6 yields the $L^1(\mathbb{R}; d\lambda)$-convergent series

$$\xi(\,\cdot\,; A, B) = \sum_{n=1}^{\infty} \xi(\,\cdot\,; B + X_n, B + X_{n-1}). \tag{A.42}$$

By Lemma A.9, each term in the above series is integer-valued a.e. on $\mathbb{R}\backslash\sigma_{\rm ac}(B)$ and hence so is the left-hand side of (A.42), which yields (A.41). □

Lemma A.11. *Let A, B be bounded self-adjoint operators in the Hilbert space $\mathcal{H}$ with $X = (A - B) \in \mathcal{I}_1$ and $\varphi \in \mathcal{H}$, $\|\varphi\| = 1$. Denote $P_\varphi = (\varphi, \,\cdot\,)\varphi$. Then,*

$$1-(\varphi, (B-z)^{-1}\varphi) = \frac{\det(I - (X + P_\varphi)(A - z)^{-1})}{\det(I - X(A - z)^{-1})}, \quad z \in \mathbb{C}_+. \tag{A.43}$$

Proof. One computes

$$\begin{aligned} I - P_\varphi(B - z)^{-1} &= (B - P_\varphi - z)(A - z)^{-1}(A - z)(B - z)^{-1} \\ &= (B - P_\varphi - z)(A - z)^{-1}[(B - z)(A - z)^{-1}]^{-1} \\ &= [I - (X + P_\varphi)(A - z)^{-1}][I - X(A - z)^{-1}]^{-1}. \end{aligned} \tag{A.44}$$

Taking determinants in (A.44) then yields

$$\begin{aligned} \frac{\det(I - (X + P_\varphi)(A - z)^{-1})}{\det(I - X(A - z)^{-1})} &= \det(I - P_\varphi(B - z)^{-1}) \\ &= 1 - (\varphi, (B - z)^{-1}\varphi). \end{aligned}$$ □

Theorem A.12. *Let A, B be bounded self-adjoint operators in the Hilbert space $\mathcal{H}$ with $(A - B) \in \mathcal{I}_1$. Then,*

$$\sigma_{\rm ac}(A) = \sigma_{\rm ac}(B).$$

Proof. By symmetry between A and B, it suffices to prove $\sigma_{\rm ac}(B) \subseteq \sigma_{\rm ac}(A)$. Suppose to the contrary that there exists a set $\mathcal{E} \subseteq \sigma_{\rm ac}(B)$ such that $|\mathcal{E}| > 0$ and $\mathcal{E} \cap \sigma_{\rm ac}(A) = \emptyset$. Choose an element $\varphi \in \mathcal{H}$ such that $\lim_{\varepsilon\downarrow 0} \operatorname{Im}((\varphi, (B - \lambda - i\varepsilon)^{-1})\varphi) > 0$ for a.e. $\lambda \in \mathcal{E}$. Thus, for a.e. $\lambda \in \mathcal{E}$, the imaginary part of the limit $z \to \lambda + i0$ of the left-hand side of (A.43) is nonzero. On the other hand, by Lemma A.10, the right-hand side of (A.43) is real for a.e. $\lambda \in \mathcal{E}$, a contradiction. $\square$

Remark. Employing $\det(I - A) = \det_2(I - A)e^{\mathrm{Tr}(A)}$ for $A \in \mathcal{I}_1$, and using an approximation of Hilbert–Schmidt operators by trace class operators in the norm $\|\cdot\|_{\mathcal{I}_2}$, one rewrites (A.43) in the case where $X = (A - B) \in \mathcal{I}_2$ as

$$1 - (\varphi, (B - z)^{-1}\varphi) = \frac{\det_2(I - (X + P_\varphi)(A - z)^{-1})}{\det_2(I - X(A - z)^{-1})} e^{(\varphi,(A-z)^{-1}\varphi)}, \quad z \in \mathbb{C}_+. \tag{A.45}$$

Since in the proof of Theorem A.12 one assumes $\mathcal{E} \subseteq \sigma_{\mathrm{ac}}(B)$, $|\mathcal{E}| > 0$, and $\mathcal{E} \cap \sigma_{\mathrm{ac}}(A) = \emptyset$, one concludes that

$$(\varphi, (A - \lambda - i0)^{-1}\varphi) \text{ is real-valued for a.e. } \lambda \in \mathcal{E}.$$

Moreover, if the boundary values of $\det_2(I - X(A - \lambda - i0)^{-1})$ exist for a.e. $\lambda \in \sigma_{\mathrm{ac}}(B)$, by (A.45), so do those of $\det_2(I - (X + P_\varphi)(A - \lambda - i0)^{-1})$. Hence, if one can assert real-valuedness of

$$\frac{\det_2(I - (X + P_\varphi)(A - \lambda - i0)^{-1})}{\det_2(I - X(A - \lambda - i0)^{-1})} \text{ for a.e. } \lambda \in \sigma_{\mathrm{ac}}(B), \tag{A.46}$$

using input from some other sources, one can follow the proof of Theorem A.12 step by step to obtain invariance of the a.c. spectrum.

In the special case of Schrödinger (and similarly for Jacobi) operators with real-valued potentials $V \in L^p([0, \infty))$, $p \in [1, 2]$, the existence of the boundary values of $\det_2(I - X(A - \lambda - i0)^{-1})$ is indeed known due to Christ–Kiselev [17] (for $p \in [1, 2)$ using some heavy machinery) and Killip–Simon [34] (for $p = 2$). We will return to this circle of ideas in [25].

References

[1] V. M. Adamjan and H. Neidhardt, *On the summability of the spectral shift function for pair of contractions and dissipative operators*, J. Operator Theory **24** (1990), 187–206.

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

[3] R. Alicki and M. Fannes, *Quantum Dynamical Systems*, Oxford University Press, Oxford, 2001.

[4] N. Aronszajn, *On a problem of Weyl in the theory of singular Sturm–Liouville equations*, Amer. J. Math. **79** (1957), 597–610.

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

[6] J. Avron, R. Seiler, and B. Simon, *The index of a pair of projections*, J. Funct. Anal. **120** (1994), 220–237.
[7] H. Baumgärtel and M. Wollenberg, *Mathematical Scattering Theory*, Operator Theory: Advances and Applications, Vol. 9, Birkhäuser, Boston, 1983.
[8] J. Bendat and S. Sherman, *Monotone and convex operator functions*, Trans. Amer. Math. Soc. **79** (1955), 58–71.
[9] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.
[10] M. Sh. Birman, *On a test for the existence for wave operators*, Soviet Math. Dokl. **3** (1962), 1747–1748; Russian original: Dokl. Akad. Nauk SSSR **147**, no. 5 (1962), 1008–1009.
[11] M. Sh. Birman, *Existence conditions for wave operators*, (in Russian) Izv. Akad. Nauk SSSR, Ser. Mat. **27** (1963), 883–906.
[12] M. Sh. Birman and M. Solomyak, *On double Stieltjes operator integrals*, Soviet Math. Dokl. **6** (1965), 1567–1571; Russian original: Dokl. Akad. Nauk SSSR **165** (1965), 1223–1226.
[13] M. Sh. Birman and M. Solomyak, *Double Stieltjes operator integrals*, Topics in Mathematical Physics, Vol. 1, pp. 25–54, Consultants Bureau, New York, 1967; Russian original: 1966 Problems of Mathematical Physics, No. 1, Spectral Theory and Wave Processes, pp. 33–67, Izdat. Leningrad. Univ., Leningrad.
[14] M. Sh. Birman and M. Solomyak, *Remarks on the spectral shift function*, J. Soviet Math. **3** (1975), 408–419; Russian original: Boundary value problems of mathematical physics and related questions in the theory of functions, 6, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **27** (1972), 33–46.
[15] M. Sh. Birman and M. Solomyak, *Double operator integrals in a Hilbert space*, Integral Equations Operator Theory **47** (2003), 131–168.
[16] 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** (1993), 833–870.
[17] M. Christ and A. Kiselev, *Absolutely continuous spectrum for one-dimensional Schrödinger operators with slowly decaying potentials: Some optimal results*, J. Amer. Math. Soc. **11** (1998), 771–797.
[18] P. A. Deift and R. Killip, *On the absolutely continuous spectrum of one-dimensional Schrödinger operators with square summable potentials*, Comm. Math. Phys. **203** (1999), 341–347.
[19] M. Dostanić, *Trace formula for nonnuclear perturbations of selfadjoint operators*, Publ. Inst. Mathématique **54** (68) (1993), 71–79.
[20] N. Dunford and J. T. Schwartz, *Linear Operators. Part II. Spectral Theory*, Interscience, New York, 1988.
[21] P. L. Duren, *On the Bloch–Nevanlinna conjecture*, Colloq. Math. **20** (1969), 295–297.
[22] P. L. Duren, *Theory of H^p Spaces*, Academic Press, Boston, 1970.
[23] E. G. Effros, *Why the circle is connected: An introduction to quantized topology*, Math. Intelligencer **11** (1989), 27–34.
[24] F. Gesztesy, K. A. Makarov, and A. K. Motovilov, *Monotonicity and concavity properties of the spectral shift function*, in *Stochastic Processes, Physics*

and Geometry: New Interplays. II. A Volume in Honor of Sergio Albeverio, F. Gesztesy, H. Holden, J. Jost, S. Paycha, M. Röckner, and S. Scarlatti (eds.), Canadian Mathematical Society Conference Proceedings, Vol. 29, pp. 207–222, American Mathematical Society, Providence, R.I., 2000.

[25] F. Gesztesy, A. Pushnitski, and B. Simon, in preparation.

[26] F. Gesztesy and B. Simon, *Rank one perturbations at infinite coupling*, J. Funct. Anal. **128** (1995), 245–252.

[27] I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Transl. Math. Monographs, Vol. 18, American Mathematical Society, Providence, R.I., 1969.

[28] R. B. Israel, *Convexity in the Theory of Lattice Gases*, Princeton University Press, Princeton, 1979.

[29] P. Jonas, *On the trace formula for perturbation theory. I*, Math. Nachr. **137** (1988), 257–281.

[30] P. Jonas, *On the trace formula for perturbation theory. II*, Math. Nachr. **197** (1999), 29–49.

[31] T. Kato, *On finite-dimensional perturbations of self-adjoint operators*, J. Math. Soc. Japan **9** (1957), 239–249.

[32] T. Kato, *Perturbation of continuous spectra by trace class operators*, Proc. Japan Acad. **33** (1957), 260–264.

[33] T. Kato, *Perturbation Theory for Linear Operators*, corr. printing of the 2nd edition, Grundlehren der Mathematischen Wissenschaften, Vol. 132, Springer, Berlin, 1980.

[34] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. (2) **158** (2003), 253–321.

[35] O. Klein, *Zur quantenmechanischen Begründung des zweiten Haupsatzes der Wärmelehre*, Z. Phys. **72** (1931), 767–775.

[36] L. S. Koplienko, *Trace formula for nontrace-class perturbations*, Siberian Math. J. **25** (1984), 735–743; Russian original: Sibirsk. Mat. Zh. **25** (1984), 62–71.

[37] F. Kraus, *Über konvexe Matrixfunktionen*, Math. Z. **41** (1936), 18–42.

[38] M. G. Krein, *On the trace formula in perturbation theory*, (in Russian) Mat. Sbornik N. S. **33(75)** (1953), 597–626.

[39] M. G. Krein, *Perturbation determinants and a formula for the traces of unitary and self-adjoint operators*, Soviet Math. Dokl. **3** (1962), 707–710; Russian original: Dokl. Akad. Nauk SSSR **144** (1962), 268–271.

[40] 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.), Operator Theory, Advances and Applications, Vol. 7, pp. 107–172, Birkhäuser, Basel, 1983.

[41] M. G. Krein, *On perturbation determinants and a trace formula for certain classes of pairs of operators*, Amer. Math. Soc. Transl. (2) **145** (1989), 39–84.

[42] M. G. Krein and V. A. Yavryan, *Spectral shift functions that arise in perturbations of a positive operator*, (in Russian) J. Operator Theory **6** (1981), 155–191.

[43] S. T. Kuroda, *On a theorem of Weyl–von Neumann*, Proc. Japan Acad. **34** (1958), 11–15.

[44] H. Langer, *Eine Erweiterung der Spurformel der Störungstheorie*, Math. Nachr. **30** (1965), 123–135.
[45] E. Lieb and G. K. Pedersen, *Convex multivariable trace functions*, Rev. Math. Phys. **14** (2002), 631–648.
[46] K. Löwner, *Über monotone Matrixfunktionen*, Math. Z. **38** (1934), 177–216.
[47] H. Neidhardt, *Scattering matrix and spectral shift of the nuclear dissipative scattering theory*, in *Operators in Indefinite Metric Spaces, Scattering Theory and Other Topics*, H. Helson and G. Arsene (eds.), Operator Theory, Advances and Applications, Vol. 24, pp. 237–250, Birkhäuser, Basel, 1987.
[48] H. Neidhardt, *Scattering matrix and spectral shift of the nuclear dissipative scattering theory. II*, J. Operator Theory **19** (1988), 43–62.
[49] H. Neidhardt, *Spectral shift function and Hilbert–Schmidt perturbation: Extensions of some work of Koplienko*, Math. Nachr. **138** (1988), 7–25.
[50] V. V. Peller, *Hankel operators in the perturbation theory of unitary and self-adjoint operators*, Funct. Anal. Appl. **19** (1985), 111–123; Russian original: Funkts. Analiz Prilozhen. **19** (1985), 37–51.
[51] V. V. Peller, *Hankel operators in the perturbation theory of unbounded self-adjoint operators*, in *Analysis and Partial Differential Equations*, C. Sadosky (ed.), Lecture Notes in Pure and Appl. Math., Vol. 122, pp. 529–544, Dekker, New York, 1990.
[52] V. V. Peller, *An extension of the Koplienko–Neidhardt trace formulae*, J. Funct. Anal. **221** (2005), 456–481.
[53] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, IV: Analysis of Operators*, Academic Press, New York, 1978.
[54] M. Rosenblum, *Perturbation of the continuous spectrum and unitary equivalence*, Pacific J. Math. **7** (1957), 997–1010.
[55] D. Ruelle, *Statistical Mechanics. Rigorous Results*, reprint of the 1989 edition, World Scientific Publishing, River Edge, N.J.; Imperial College Press, London, 1999.
[56] 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** (1996), 683–699.
[57] L. A. Sahnovič, *Dissipative operators with absolutely continuous spectrum*, Trans. Moscow Math. Soc. **19** (1968), 233–297; Russian original: Trudy Moskov. Mat. Obšč. **19** (1968), 211–270.
[58] B. Simon, *The Statistical Mechanics of Lattice Gases, Vol. 1*, Princeton University Press, Princeton, 1993.
[59] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory; Part 2: Spectral Theory*, AMS Colloquium Series, American Mathematical Society, Providence, R.I., 2005.
[60] B. Simon, *Trace Ideals and Their Applications*, 2nd edition, Mathematical Surveys and Monographs, Vol. 120, American Mathematical Society, Providence, R.I., 2005.
[61] K. B. Sinha and A. N. Mohapatra, *Spectral shift function and trace formula*, Proc. Indian Acad. Sci. (Math. Sci.) **104** (1994), 819–853.
[62] E. M. Stein and R. Shakarchi, *Fourier Analysis. An Introduction*, Princeton Lectures in Analysis, 1, Princeton University Press, Princeton, N.J., 2003.

[63] E. C. Titchmarsh, *On expansions in eigenfunctions (VI)*, Quart. J. Math., Oxford **12** (1941), 154–166.

[64] D. Voiculescu, *On a trace formula of M. G. Kreĭn*, in *Operators in Indefinite Metric Spaces, Scattering Theory and Other Topics (Bucharest, 1985)*, pp. 329–332, Oper. Theory Adv. Appl., 24, Birkhäuser, Basel, 1987.

[65] J. von Neumann, *Charakterisierung des Spektrums eines Integraloperators*, Actualités Sci. Industr. **229** (1935), 3–20.

[66] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, 12th printing, Princeton Landmarks in Mathematics, Princeton Paperbacks, Princeton University Press, Princeton, N.J., 1996.

[67] H. Weyl, *Über beschränkte quadratische Formen, deren Differenz vollstetig ist*, Rend. Circ. Mat. Palermo **27** (1909), 373–392.

[68] D. R. Yafaev, *Mathematical Scattering Theory*, Transl. Math. Monographs, Vol. 105, American Mathematical Society, Providence, R.I., 1992.

[69] D. R. Yafaev, *Perturbation determinants, the spectral shift function, trace identities, and all that*, preprint, 2007.

[70] A. Zygmund, *Trigonometric Series, Vols. I & II*, 2nd edition, Cambridge University Press, Cambridge, 1990.