Trace Class Operators and Lidskii's Theorem

Tom Phelan

Semester 2 2009

1 Introduction

The purpose of this paper is to provide the reader with a self-contained derivation of the celebrated Lidskii Trace Theorem for operators on a Hilbert space. Our exposition will necessarily involve digressions into a selection of techniques of modern mathematical physics, including antisymmetric tensor products and construction of the determinant function. This approach is not standard throughout the literature on this subject and this report owes much to the work of Barry Simon and Michael Reed, who themselves departed somewhat from Lidskii's original paper in favour of a more modern expository style.

Lidskii's Theorem asserts the following: If $A \in \mathcal{T}_1$ and $\{\lambda_i\}$ are the eigenvalues of A then for any orthornormal basis $\{\eta_j\}$ we have

$$\sum_n \lambda_n(A) = Tr(A) := \sum_m (\eta_m, A\eta_m) \tag{1}$$

Of course to make sense of this expression one must first digress to explain the meaning of the terms and concepts involved. This is the subject of our first section.

2 Definitions and Preliminary Material

2.1 The Spectrum

Definition 2.1. Let $T \in \mathcal{L}(X)$ and $\lambda \in \mathbb{C}$. Then λ is said to be in the *resolvent set* $\rho(T)$ of A if the linear transformation $\lambda I - T$ is a bijection with a bounded inverse. In this case, the inverse $R_\lambda(T) = (\lambda I - T)^{-1}$ is called the *resolvent* of T at λ. If $\lambda \notin \rho(T)$, then λ is in the spectrum $\sigma(T)$ of T.

By the bounded inverse theorem demonstrated in our Banach space theory lectures, this last requirement that $\lambda I - T$ have a bounded inverse is actually redundant whenever $\lambda I - T$ is a bijection. Elements of the spectrum $\sigma(T)$ fall under two headings: eigenvalues and elements of the residual spectrum.

Definition 2.2. Let $T \in \mathcal{L}(X)$.

1. If there exists a non-zero $x \in X$ such that $Tx = \lambda x$, then λ is termed an *eigenvalue* of T. If λ is an eigenvalue, then $\lambda I - T$ is not injective.

2. If λ is not an eigenvalue, then $Ran(\lambda I - T)$ is not dense and λ is said to be in the *residual spectrum* of T.

2.2 Positive Operators on a Hilbert Space

Here we make mention of a special decomposition of operators on a Hilbert space that serves an analogous purpose to the polar decomposition of complex numbers, $z = |z|e^{i\ arg\ z}$. The complete construction of this decomposition is not crucial to our analysis and so we omit the details, which may be found in [1]. We are primarily interested in the notion of the square root of an operator A, for which the following definition is required.

Definition 2.3. Let $\mathcal{H}$ be a Hilbert space. A linear operator A is called *positive* if for all $x \in \mathcal{H}$, $(Ax, x) \geq 0$. One writes $B \geq 0$ if B is positive and $A \geq B$ if $A - B \geq 0$.

Positive operators play a role similar to that of the positive real numbers in polar decomposition, but this fact is not a feature of our exposition. Note that for any A, the operator A^*A is positive since for all $x \in \mathcal{H}$,

$$(A^*Ax, x) = (Ax, Ax) = ||Ax||^2 \geq 0$$

This indicates that we may want to define $|A| = \sqrt{A^*A}$ just as we defined $|z| = \sqrt{\bar{z}z}$ for the complex numbers. To do this one must first show the notion of square root is well-defined for positive operators, for which we need the following elementary lemma, stated without proof:

Lemma 2.1. *The power series $\sum_{n=1}^{\infty} c_n z^n$ for $\sqrt{1-z}$ converges absolutely for all complex numbers z satisfying $|z| \leq 1$.*

Theorem 2.2. *Let $A \in \mathcal{L}(\mathcal{H})$ be a positive operator on a Hilbert space $\mathcal{H}$. Then there exists a unique $B \in \mathcal{L}(\mathcal{H})$ such that $B \geq 0$ and $B^2 = A$.*

Proof. Due to the linearity of the $||.||$ function, it is sufficient to consider the case where $||A|| \leq 1$. Then $I - A$ is positive and since

$$||I - A|| = \sup_{||\phi||=1} |((I - A)\phi, \phi)| \leq 1$$

one may apply the above lemma to see that the series

$$I + c_1(I - A) + c_2(I - A)^2 + \dots \tag{2}$$

converges in norm by the triangle inequality,

$$||I + c_1(I - A) + c_2(I - A) + \dots || \ \leq \ 1 + \sum_i^{\infty} c_i ||I - A||^i$$

which ensures that (2) converges in norm to some operator B. Now, since the convergence of (2) is absolute one may square and rearrange terms without affecting the sum. Computation shows $B^2 = A$, and further that since the series defining B converges absolutely we know B commutes with any operator that commutes with A. Now suppose there is another operator $B' \in \mathcal{L}(\mathcal{H})$ such that $B' \geq 0$ and $B'^2 = A$. Then

$$B'A = (B')^3 = AB'$$

shows B' commutes with A and hence with B, using the above. Thus we may write

$$(B - B')B(B - B') + (B - B')B'(B - B') = (B^2 - B'^2)(B - B') = 0$$

Since the terms on the left-hand side above sum to zero and are both positive, they must be equal to zero. Their difference $(B - B')^3$ must then also be equal to zero. Since $B - B'$ is self-adjoint, from the first assignment of MATH3325 we know that $||B - B'||^4 = ||(B - B')^4|| = 0$, which shows $B = B'$. □

This derivation justifies the following definition:

Definition 2.4. If $A \in \mathcal{L}(\mathcal{H})$, then $|A| = \sqrt{A^*A}$

Some remarks are in order about the use of the symbol $|.|$ in this context,

1. For any $\lambda \in \mathbb{C}$, $|\lambda A| = |\lambda||A|$
2. It is not true in general that $|AB| = |A||B|$ or that $|A| = |A^*|$.
3. It is not true in general that $|A + B| \leq |A| + |B|$.
4. It is true that $|.|$ is norm continuous, but is not known whether it is Lipschitz.

It may be shown (see for instance, [1]) that given the above notion of the square root of an operator A, the following holds:

Theorem 2.3 (Polar Decomposition)**.** *Let A be a bounded linear operator on a Hilbert space $\mathcal{H}$. Then there exists a partial isometry U (that is, an isometry when restricted to $Ker(U)^{\perp}$) such that $A = |A|U$.*

In light of these results we make the following definition:

Definition 2.5. If A is a compact operator then the set $\{\mu_i(A)\}$ of eigenvalues of $|A|$ are called the *singular values* of A.

We are now in a position to characterise those operators that are the subject of Lidskii's Theorem, the elements of the trace class $\mathcal{T}_1$.

3 Trace Class Operators

Definition 3.1. Let $\mathcal{H}$ be a Hilbert space and $\{\phi_i\}$ an orthonormal basis for $\mathcal{H}$. Then for any positive operator $A \in \mathcal{L}(\mathcal{H})$ we define the *trace* of A to be

$$tr\ A = \sum_{i=1}^{\infty}(\phi_i, A\phi_i) \tag{3}$$

Some elementary properties of the *trace* are listed below:

Proposition 3.1. *1. The quantity $tr\ A$ is independent of the choice $\{\phi_i\}$ of orthonormal basis.*

2. $tr(A + B) = tr\ A + tr\ B$.

3. $tr(\lambda A) = \lambda tr\ A$ for all $\lambda \geq 0$.

Proof. The second and third properties ought to be obvious to the reader and so we shall only prove the first. For a given orthonormal basis $\{\phi_i\}$, define $tr_\phi\ A = \sum_{i=1}^{\infty}(\phi_i, A\phi_i)$. If $\{\psi_i\}$ is another orthonormal basis, we have

$$tr_\phi A = \sum_{i=1}^{\infty}(\phi_i, A\phi_i) = \sum_{i=1}^{\infty} ||\sqrt{A}\phi_i||^2 \tag{4}$$

where we have used the result above that every positive operator has a unique square root that is necessarily self-adjoint. Since the $\{\psi_i\}$ form a basis, we may proceed

$$tr_\phi A = \sum_{i=1}^{\infty}\big(\sum_{j=1}^{\infty}|(\psi_j, \sqrt{A}\phi_i)|^2\big) = \sum_{i=1}^{\infty}\big(\sum_{j=1}^{\infty}|(\sqrt{A}\psi_j, \phi_i)|^2\big) \tag{5}$$

Simplifying, (5) becomes

$$tr_\phi A = \sum_{j=1}^{\infty}||\sqrt{A}\psi_j||^2 = \sum_{j=1}^{\infty}(\psi_j, A\psi_j) = tr_\psi A$$

$\square$

Definition 3.2. An operator $A \in \mathcal{L}(\mathcal{H})$ is called *trace class* if and only if $tr\ |A| < \infty$. The family of all trace class operators is denoted by $\mathcal{T}_1$ and the map $tr : \mathcal{T}_1 \to \mathbb{C}$ defined by $A \mapsto \sum_{i=1}^{\infty}(\phi_i, A\phi_i)$ is called the *trace* of the operator A. By the above, this is well-defined indpendent of the choice of orthogonal basis $\{\phi_i\}$.

The reader will note that it is precisely those operators A in the trace class $\mathcal{T}_1$ to which the Lidskii result refers. We list now without proof some elementary properties of the trace function and trace class operators.

Theorem 3.2. *Let $||.||_1$ be defined on $\mathcal{T}_1$ by $||A||_1 = tr||A||$. Then $\mathcal{T}_1$ forms a Banach space under the norm $||.||_1$.*

It turns out there is a simple connection between operators in the trace class $\mathcal{T}_1$ and compact operators. Since it employs no new techniques relevant to our discussion, we omit the proof of the following theorem. For details see [1].

Theorem 3.3. *Every $A \in \mathcal{T}_1$ is compact. Conversely, a compact operator A is in $\mathcal{T}_1$ if and only if $\sum \lambda_i < \infty$, where $\{\lambda_i\}$ are the singular values of A.*

The following corollary serves to provide one with an intuitive understanding of the size of $\mathcal{T}_1$, and thus allow for the application of approximation arguments in what follows. For details, the reader is again referred to [1].

Corollary 3.4. *The finite rank operators are desne in $\mathcal{T}_1$ under the $||.||_1$ topology.*

4 Antisymmetric Tensor Products and the Determinant

There are various ways in which one may construct appropriate tools for the derivation of the Lidskii result. We have chosen to present here a brief development of what is often termed alternating algebra, or the theory of anti-symmetric tensor products. This will allow for an elegant extension of the familiar determinant function on a finite dimensional space to the generality of the infinite dimensional case. Several links between the trace and determinant of an operator $A \in \mathcal{T}_1$ then arise in a natural way. In the following sections we shall simply state those properties of this analytical framework crucial for our exposition. For further details the reader is referred to [3].

Definition 4.1. Let A be a compact operator on a Hilbert space $\mathcal{H}$. A *Schur basis* for the operator A is an orthonormal set $\{e_i\}_{i=1}^{N(A)}$ such that for each $0 \leq n \leq N(A)$

$$Ae_n = \lambda_n(A)e_n + \sum_{i=1}^{n-1} v_{nm}e_i$$

for some constants v_{nm}. In particular,

$$(e_n, Ae_n) = \lambda_n(A) \tag{6}$$

Note that a Schur basis is defined for a given operator A and need not actually form a basis for the Hilbert space $\mathcal{H}$. Our interest in Schur bases arises from the following result, found in [2]:

Proposition 4.1. *Every compact operator A has a Schur basis $\{e_i\}_{i=1}^{N(A)}$.*

The following definitions introduce the tools we need to study the propertires of the determinant function:

Definition 4.2. Given a Hilbert space $\mathcal{H}$, $\otimes^n\mathcal{H}$ is defined as the vector space of multilinear functionals on $\mathcal{H}$. Given $f_1, \ldots, f_n \in \mathcal{H}$ we define $f_1 \otimes \cdots \otimes f_n \in \otimes^n\mathcal{H}$

$$(f_1 \otimes \cdots \otimes f_n)(g_1, \ldots, g_n) = (f_1, g_1) \ldots (f_n, g_n)$$

for an arbitrary vector $(g_1, \ldots, g_n) \in \mathcal{H}^n$. Further, for any $A \in \mathcal{L}(\mathcal{H})$ there is a natural operator $\Gamma_n(A) \in \mathcal{L}(\otimes^n\mathcal{H})$ defined by

$$\Gamma_n(A)((f_1 \otimes \cdots \otimes f_n)) = Af_1 \otimes \cdots \otimes Af_n$$

Definition 4.3. Let $\mathcal{P}_n$ denote the set of all permutations of n letters, and let $\epsilon : \mathcal{P}_n \to \{-1, 1\}$ be the function equal to $+1$ on even permutations and -1 on odd permutations. Then we may define $\phi_1 \wedge \cdots \wedge \phi_n \in \otimes^n\mathcal{H}$ to be

$$\phi_1 \wedge \cdots \wedge \phi_n = \sum_{\pi \in \mathcal{P}_n} \epsilon(\pi)[\phi_{\pi(1)} \otimes \cdots \otimes \phi_{\pi(n)}]$$

Now we may define $\Lambda^n(\mathcal{H})$ to be the subspace of $\otimes^n\mathcal{H}$ spanned by all vectors of the form $\{\phi_1 \wedge \cdots \wedge \phi_n\}$. For any $A \in \mathcal{L}(\mathcal{H})$, we define the operator $\Lambda^n(A)$ to be the restriction $\Gamma_n(A)$ to $\Lambda^n(\mathcal{H})$. We now turn to the construction fo the determinant function.

There are various definition of the determinant function in the standard literature, and their equivalence is the content of Lidskii's Theorem. For further details of alternative expositions, we refer to reader to [2]. To motivate our discussion, we shall first suppose that A is an operator on a finite dimensional space $\mathcal{H}$ where $dim\ \mathcal{H} = n$. Then we may take as our definition of the determinant the function

$$\det(a_{ij}) = \sum_{\pi in \mathcal{P}_n} \epsilon(\pi) a_1\pi_1 \ldots a_n\pi_n$$

Then if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A and $e_1, \ldots, e_n$ form a Schur basis for A, it is easy to observe (see [3] for details) that

$$\det(1 + A) = (e_1 \wedge \cdots \wedge e_n, (1 + A)e_1 \wedge \cdots \wedge (1 + A)e_n) = \prod_{j=1}^{n}(1 + \lambda_j)$$

and

$$tr(\Lambda^n(A)) = \sum_{1 \leq i_1 \leq n} ((e_{i_1} \wedge \ldots e_{i_k}), (Ae_{i_1} \wedge \ldots Ae_{i_k})) = \sum_{1 \leq i_1 \leq n} \lambda_{i_1} \ldots \lambda_{i_n}$$

yielding

$$\det(1 + A) = \sum_{k=0}^{\infty}(tr\Lambda^k(A))$$

for the finite dimensional case. For the general case, we simply define the determinant in a manner consistent with the above identity:

Definition 4.4. If A is a trace class operator on a Hilbert Space $\mathcal{H}$, then we define

$$\det(1 + zA) = \sum_{k=0}^{\infty} z^k (tr\Lambda^k(A)) \tag{7}$$

Of critical importance to our exposition is the assertion that $\det(1 + .)$ is continuous. For this we use the following result, a proof of which may be found in [3].

Lemma 4.2. *For any $A, B \in \mathcal{T}_1$,*

$$|\det(1 + A) - \det(1 + B)| \leq ||A - B||_1 exp(||A||_1 + ||B||_1 + 1) \tag{8}$$

where $exp(.)$ is the familiar real-valued exponential function.

Proposition 4.3. *The function $\det(1 + .)$ is continuous.*

Proof. First note that by the continuity of the exponential function, for any fixed operator A and all $\epsilon_1 > 0$ there exists δ_1 such that

$$||A - B||_1 < \delta_1 \implies exp(||A||_1 + ||B||_1 + 1) < exp(2||A||_1 + 1) + \epsilon_1 \tag{9}$$

We wish to show $\det(1 + .)$ is continuous at A. Fixing $1 > \epsilon_2 > 0$ we note that by (8) it will suffice to show there exists δ_2 such that

$$||A - B||_1 < \delta_2 \implies ||A - B||_1 exp(||A||_1 + ||B||_1 + 1) < \epsilon_2 \tag{10}$$

For this one need only select ϵ_1 and δ_1 satisfying (9) such that

$$\delta_1(B^{-1} + \epsilon_1) < \epsilon_2$$

Where $B = exp(2||A||_1 + 1)$. Then choosing $\delta_2 < \delta_1$ will suffice to demonstrate (9), in light of (10). Since ϵ_2 was arbitrarily small, this completes the proof. □

Together with our comments regarding anti-symmetric tensor products above, this allows us to demonstrate some useful properties of $\det(1 + A)$, summarised in the following proposition.

Proposition 4.4. *Let A and B be trace class operators. Then,*

1. $\det(1 + A)\det(1 + B) = \det(1 + A + B + AB)$
2. $(1 + A)$ *is invertible if and only if* $\det(1 + A) \neq 0$
3. *If $-\mu^{-1}$ is an eigenvalue of of A, then $\det(1 + zA)$ has a zero of order n at $z = \mu$, where n is the algebraic multiplicity of $-\mu^{-1}$*

4. *For any $\epsilon > 0$ there exists a constant C_ϵ such that*

$$|\det(1+zA)| \leq C_\epsilon exp(\epsilon|z|) \tag{11}$$

Sketch. 1. Since we know that the finite rank operators are dense in $\mathcal{T}_1$ and $\det(1+.)$ is continuous by the above, we need only verify this first identity for finite rank operators and apply a limiting argument. This effectively amounts to showing that for finite-dimensional operators A and B,

$$\det(AB) = \det(A)\det(B)$$

2. If $1+A$ is not invertible, then $\det(1+A) = 0$ by the third identity. If $1+A$ is invertible, then define $B = -A(1+A)^{-1}$. Then $\det(1+A+B+AB) = 1$, and so $\det(1+A) \neq 0$ by the first identity.

3. and 4. We omit these proofs, see [3] for details.

5 Lidskii's Theorem

Given the above constructions, our derivation of the Lidskii result rests largely upon the following:

Theorem 5.1. *For any $A \in \mathcal{T}_1$,*

$$det(1+zA) = \prod_{j=1}^{N(A)} (1+z\lambda_j(A)) \tag{12}$$

where the $\{\lambda_j\}_{j=1}^{N(A)}\}$ $N(A)$ are the eigenvalues of A.

For the proof, we shall require the following lemma from complex analysis,

Lemma 5.2 (Borel-Caratheodory)**.** *Let f be analytic in a neighbourhood of $|z| \leq R$. Then for all $r \leq R$,*

$$\max_{|z|=r} |f(z)| = \frac{2r}{R-r} \max_{|z|=R} [Re(f(z))] + \frac{R+r}{R-r} |f(0)| \tag{13}$$

Proof. For suppose that $|f(0)| = 0$. Without loss, we may also suppose that f is not identically zero for otherwise the identity (13) is trivial. If we define $A = \max_{|z|=R}[Re(f(z))]$ the maximum modulus principle applied to the function e^f then implies that $Re(f(z)) \leq A$ for $|z| \leq R$. Thus the function

$$g(z) \equiv \frac{f(z)}{z(2A - f(z))} \tag{14}$$

is analytic on the domain $|z| < R$. Writing f as $f = u + iv$ for some real functions u, v we find that for $|z| = R$,

$$|g(z)|^2 = \frac{1}{R^2}\frac{u^2+v^2}{(2A-u)^2+v^2} \leq \frac{1}{R^2}$$

Invoking the maximum modulus principle, it follows that $|g(z)| \leq R^{-1}$ for all $|z| \leq R$. Re-arranging (14) for $f(z)$,

$$f(z) = \frac{2Azg(z)}{1+zg(z)}$$

we see that for $|z| = r$,

$$|f(z)| \leq \frac{2Ar}{R-r} \tag{15}$$

Invoking the maximum modulus principle for the thrid time, (15) implies (14). Now suppose that $f(0) \neq 0$. Then since we know (14) holds for $h(z) = f(z) - f(0)$, we may write

$$\max_{|z|=r}|h(z)| \leq \frac{2r}{R-r}\max_{|z|=R}[Re(f(z)-f(0))] \leq \frac{2r}{R-r}\max_{|z|=R}[Re(f(z))] + \frac{2r}{R-r}|f(0)| \tag{16}$$

Adding $|f(0)|$ to both sides of (16) and applying the triangle inequality, it follows that

$$\begin{aligned}\max_{|z|=r}|f(z)| &\leq \frac{2r}{R-r}\max_{|z|=R}[Re(f(z))] + \frac{2r}{R-r}|f(0)| + |f(0)| \\ &= \frac{2r}{R-r}\max_{|z|=R}[Re(f(z))] + \frac{R+r}{R-r}|f(0)|\end{aligned}$$

proving (13) for a general analytic function $f(z)$. □

We are now in a position to prove (12).

Proof. let us temporarily denote the left-hand side by $f(z)$ and the right-hand side by $g(z)$. Our aim is then to show that $f(z) = g(z)$ for all $z \in \mathbb{C}$. The purpose of our earlier digression into the determinant function now becomes apparent. Recalling our earlier result we see that f and g have the same zeroes including multiplicity, and so the quotient $\frac{f}{g}$ is a nowhere vanishing entire analytic function. By elementary complex analysis, we may write

$$f(z) = g(z)e^{h(z)}$$

for some entire function h with $h(0) = 0$. For $|z| < R$ we define three families of functions

1. $$h_R(z) = ln[f_R(z)]$$

2.

$$k_R(z) = -\sum_{\{j||\lambda_j|^{-1}>R\}} ln(1 + z\lambda_j(A))$$

3.

$$f_R(z) = f(z)[\prod_{\{j||\lambda_j|^{-1}\leq R\}} (1 + z\lambda_j)]^{-1}$$

Observe that for any R, f_R is an entire function and that since $e^h(z) = \frac{f}{g}$ we find

$$\begin{aligned} h_R + k_R &= lnf(z) - \sum_{\{j||\lambda_j|^{-1}\leq R\}} ln(1 + z\lambda_j(A)) - \sum_{\{j||\lambda_j|^{-1}>R\}} ln(1 + z\lambda_j(A)) \\ &= lnf(z) - lng(z) \end{aligned}$$

Thus it will suffice to show that for $|z| < 1$ we have $|h_R| \to 0$ and $|k_R| \to 0$ as $R \to \infty$.

Now observe that since $ln(1 + x)$ vanishes at $x = 0$ and is analytic in a neighbourhood of the set $\{x \ : \ |x| \leq \frac{1}{2}\}$, there exists a constant C such that for $|x| \leq \frac{1}{2}$

$$|ln(1 + x)| \leq C|x|$$

Applying this observation to our present situation, we find that for $R \geq 2$ and $|z| \leq 1$,

$$|k_R(z)| \leq |z| \sum_{\{j||\lambda_j|^{-1}>R\}} |\lambda_j(A)| \tag{17}$$

Since the sum on the right-hand side of (17) is convergent, this shows $|k_R(z)| \to 0$ as $R \to \infty$. We now turn to the task of showing $|h_R| \to 0$ as $R \to \infty$.

Consider the entire function $f_R(z)$. If $|\lambda_i|^{-1} \leq R$, then $|1 + z\lambda_i| \leq 1$ and thus $|f_R(z)| \leq |f(z)|$, so applying our earlier result (11) we see

$$|f_R(z)| \leq C_\epsilon exp(2\epsilon R)$$

Applying the maximum moduls principle again, it follows that

$$Re(h_R(z)) \leq ln\ C_\epsilon + 2\epsilon R$$

Applying the Borel-Caratheodory result proved above to the entire function h_R, and noting that $h_R(0) = 0$, one obtains the estimate

$$max_{|z|\leq 1}|h_R(z)| \leq 2(R - 1)^{-1}[ln\ C_\epsilon + 2\epsilon R]$$

Thus for any $|z| \leq 1$, it follows that

$$\limsup_{R\to\infty} |h_R(z)| \leq 4\epsilon$$

Since ϵ was arbitrary, this shows $|h_R(z)| \to 0$ as $R \to \infty$. $\square$

Our interest in this theorem lies in the observation that if two complex-valued functions f and g agree for every $z \in \mathbb{C}$, then their Taylor series expansions about $z = 0$ are term-by-term identical. Together with the above, this observation allows us to show

Theorem 5.3 (Lidskii's Theorem)**.** *If $A \in \mathcal{T}_1$ and $\{\lambda_i\}$ are the eigenvalues of A, then*

$$\sum_n \lambda_n(A) = Tr(A)$$

Proof. Differentiating the left-hand side of (12) and evaluating at $z = 0$, we find we find that the constant term in the Taylor expansion of $\det(1 + zA)$ about $z = 0$ is

$$\frac{d}{dz} \prod_{j=1}^{N(A)} (1 + z\lambda_j(A))|_{z=0} = \sum_{j=1}^{N(A)} \lambda_j \tag{18}$$

whilst the constant term in the Taylor expansion for the right-hand side of (12) is

$$\frac{d}{dz} \sum_{k=0}^{\infty} z^k (tr\Lambda^k(A))|_{z=0} = tr\ A \tag{19}$$

Since $\Lambda^1(A) = A$. Combining (18) and (19) immediately yields our desired result. $\square$

References

[1] Michael Reed and Barry Simon, Methods of Mathematical Physics Volume I, Academic Press, 1978.

[2] Barry Simon, Trace Ideals and Their Applications, Mathematical Surveys and Monographs Volume 120, 1979

[3] Michael Reed and Barry Simon, Methods of Mathematical Physics Volume IV, Academic Press, 1978.