dimension of the generalized eigenspace.

The trace formula - Lidskii's Theorem (1959)

The trace of a trace class operator on a Hilbert space is the sum of its eigenvalues (including multiplicity), i.e.,

$$\operatorname{tr} T = \sum_{j=1}^{\infty} \lambda_j(T).$$

(the trace is exactly the sum of the eigenvalues and generalized eigenvalues. Nothing else).

Proof:

Let T be a trace class operator. Then, by definition,
$\operatorname{tr} T = \sum_{n=1}^{\infty} \langle T f_n, f_n \rangle$, $\{f_n\}$ ONB.

If T is normal then we can choose an ONB consisting of eigenvectors of T, and thus we get

$$\operatorname{tr} T = \sum_{n=1}^{\infty} \langle T f_n, f_n \rangle = \sum_{n=1}^{\infty} \lambda_n(T),$$

as wanted.

If T is not normal the eigenvectors are in general not orthogonal, and we may have generalized eigenvectors, so

$$T w_n = \lambda_n w_n, \quad \text{or} \quad T w_n = \lambda_n w_n + w_{n-1}$$

We use Gram-Schmidt to orthonormalize them and get an orthonormal set $\{f_n\}$ so that

$$T f_n = \lambda_n f_n + \text{lin. comb. of } f_1, \dots, f_{n-1}.$$

Since the f_n's are orthonormal, $(Tf_n, f_n) = \lambda_n$.
Summing all n would yield the trace formula if $\{f_n\}$ form a basis of the whole Hilbert space.
They do if the eigenvectors and generalized eigenvectors span the whole space, but this is not always true.
Therefore $\{f_n\}$ might have to be supplemented by an orthonormal basis $\{h_m\}$ for the orthogonal complement of the span of f_n's, $(\text{span } f_n)^{\perp}$.
The expression of the trace of T now reads

$$\operatorname{tr} T = \sum_{n=1}^{\infty} (Tf_n, f_n) + \sum_{n=1}^{\infty} (Th_m, h_m) = \sum_{n=1}^{\infty} \lambda_n(T) + \sum_{n=1}^{\infty} (Th_m, h_m),$$

and the goal is to show that the last sum is zero.
For this we need some lemmas.

Lemma 6 Let T be a compact operator on a Hilbert space H, and let K be the orthogonal complement of its eigenvectors and generalized eigenvectors. Then
(i) K is an invariant subspace of T^*.
(ii) The spectrum of T^* over K consists of the single point $\lambda = 0$, $\sigma(T^*) = \{0\}$ on K.

Proof: Let e be an eigenvector, possibly generalized, of T,

$$Te = \lambda e + f,$$

where f is another generalized eigenvector.
Suppose that u is orthogonal to e and f.

Then we find

$$\langle e, T^* u\rangle = \langle Te, u\rangle = \langle \lambda e + f, u\rangle = \lambda\langle e, u\rangle + \langle f, u\rangle = 0,$$

and thus $T^* : K \to K$ as wanted.

(ii) Since T is compact then so is T^*.
If $\lambda \neq 0$ is an eigenvalue of T^* on K then $\bar{\lambda}$ will be an eigenvalue of T on H (of finite multiplicity), i.e., $\dim(N(T-\bar{\lambda})) < \infty$.
(According to a theorem by Riesz) there exists $i \in \mathbb{Z}$ such that

$$N((T^*-\lambda)^i) = N((T^*-\lambda)^{i+1}) \supseteq N((T^*-\lambda)^{i-1}).$$

Let $u \in K$ be in $N((T^*-\lambda)^i) \setminus N((T^*-\lambda)^{i-1})$.
Then the equation

$$(T^*-\lambda)v = u$$

has no solution.
For a solution v would belong to $N((T^*-\lambda)^{i+1})$ but not to $N((T^*-\lambda)^i)$ which is a contradiction.
Then according to the Fredholm alternative there must be an eigenvector w of T, $(T-\bar{\lambda})w = 0$, that is not orthogonal to u.

$u \in \operatorname{Ran} T \iff \langle u, h\rangle = 0 \ \forall h \in \ker T^*$

But this is a contradiction, because $u \in K$. $\square$

$$(T^*-\lambda)^i v = \overbrace{(T^*-\lambda)^{i-1} u}^{\neq 0 \text{ since } u \notin N((T^*-\lambda)^{i-1})}$$

$$(T^*-\lambda)^{i+1} v = (T^*-\lambda)^i u = 0$$

If T is trace class over H then so is T^*, and $T^*|_K$ is then of trace class as well.

Now we look again at the sum we wanted to show was zero. We can rewrite it as follows

$$\sum_{n=1}^{\infty} \langle Th_m, h_m \rangle = \sum_{n=1}^{\infty} \langle h_m, T^* h_m \rangle = \sum_{n=1}^{\infty} \overline{\langle T^* h_m, h_m \rangle}$$

Since h_m is an ONB for K this is the complex conjugate of the trace of T^* over K,

$$\overline{\operatorname{tr} T^*|_K} = \sum_{n=1}^{\infty} \overline{\langle T^* h_m, h_m \rangle}.$$

Because of Lemma 6 which we just proved the vanishing of this sum can be described as this:

Lidskii's Lemma. Let T be a trace class operator that has no eigenvalues except zero. Then $\operatorname{tr} T = 0$.

The proof of this lemma goes through several intermediate results, the fiost one iis an estimate of the eigenvalues of a compact operator in terms of its singular values.

Recall:
When T is compact then so is the absolute value $A := |T|$. The nonzero eigenvalues of A we denote $\{s_j\}$. They are positive numbers that tend to zero, and we index them in decreasing order.
The numbers s_j are then called the singular values of the operator T, and we write $s_j(T)$.

It can be shown that for each j, $s_j(T)$ is a continuous function of T (in norm topology).

Lemma 7. Let T be a compact operator with nonzero eigenvalues $\lambda_1, \lambda_2, \dots$, arranged in decreasing order of their absolute value (including multiplicity).
Then for any $N \in \mathbb{N}$,

$$\prod_{j=1}^{N} |\lambda_j| \leq \prod_{j=1}^{N} s_j(T),$$

where $s_j(T)$ are the singular values of T.

Proof: Let E_N be the space spanned by the first N eigenvectors of T and let P_N be an orthogonal projection onto E_N.
Denote by T_N the restriction of T to E_N and by A_N the absolute value of T_N, that is,

$$T_N = U_N A_N \qquad \text{(polar decomposition)}$$

Since the eigenvalues λ_j are nonzero T_N is invertible.
Thus U_N is as well and therefore it is unitary
Taking determinants we get

$$|\det(T_N)| = |\det(U_N A_N)| = |\det(U_N)||\det(A_N)| = \det(A_N),$$

since $\det(A_N) \geq 0$.
Since the determinant of a matrix is the product of its eigenvalues we can write this as

$$\prod_{j=1}^{N} |\lambda_j| = \prod_{j=1}^{N} \lambda_j(A_N). \qquad (*)$$

The operator TP_N acts on E_N as the matrix T_N. On $E_N^{\perp}$ $TP_N = 0$. Therefore we must have that the absolute value of TP_N is A_n on E_N, i.e.,

$$TP_N = T_N = U_N A_N \text{ on } E_N$$

and $\qquad TP_N = 0$ on $E_N^{\perp}$.
From this we see that $\lambda_j(A_N) = s_j(TP_N)$, $j = 1, \dots, N$, and since the singular values of adjoint operators are the same we find

$$s_j(TP_N) = s_j(P_N^* T^*) \leq \|P_N^*\| \, s_j(T^*) = \|P_N\| \, s_j(T) \leq s_j(T).$$

inequality proven earlier in book.

Now, inserting into (*) we obtain

$$\prod_{j=1}^{N} |\lambda_j| = \prod_{j=1}^{N} \lambda_j(A_N) = \prod_{j=1}^{N} s_j(TP_N) \leq \prod_{j=1}^{N} s_j(T),$$

proving the assertion. $\square$

We can use the following principle to deduce further inequalities between the absolute value of the eigenvalues $|\lambda_j|$ and the singular values s_j:

<u>Lemma 8</u> Let $a_1 \geq a_2 \geq \dots$ and $b_1 \geq b_2 \geq \dots$ be two decreasing sequences of real numbers, satisfying for each $N \in \mathbb{N}$

$$\sum_{j=1}^{N} a_j \leq \sum_{j=1}^{N} b_j .$$

Let F be a convex function defined on $\mathbb{R}$ that tends to zero as its argument tends to $-\infty$, that is, $F(x) \to 0$ as $x \to -\infty$. Then

$$\sum_{j=1}^{N} F(a_j) \leq \sum_{j=1}^{N} F(b_j) \qquad \text{for every } N \in \mathbb{N}.$$

<u>Proof:</u> Omitted (for now). $\square$

We proved earlier that $\prod_{j=1}^{N} |\lambda_j| \leq \prod_{j=1}^{N} s_j$, so taking logarithms we have

$$\log\left(\prod_{j=1}^{N} |\lambda_j|\right) \leq \log\left(\prod_{j=1}^{N} s_j\right)$$

$$\Leftrightarrow \quad \sum_{j=1}^{N} \log|\lambda_j| \leq \sum_{j=1}^{N} \log s_j .$$

Therefore we can use Lemma 8 with $a_j = \log|\lambda_j|$ and $b_j = \log s_j$. Choose $F(x) = e^x$. This is clearly convex, and $e^x \to 0$ as $x \to -\infty$.

Thus Lemma 8 yields

$$(**)\qquad \sum_{j=1}^{N} |\lambda_j| \leq \sum_{j=1}^{N} s_j .$$

If we instead choose $F(x) = \log(1+re^x)$, $r>0$, we obtain (from Lemma 8)

$$\sum_{j=1}^{N} \log(1+r|\lambda_j|) \leq \sum_{j=1}^{N} \log(1+rs_j),$$ which implies

$$\log\Big(\prod_{j=1}^{N}(1+r|\lambda_j|)\Big) \leq \log\Big(\prod_{j=1}^{N}(1+rs_j)\Big),$$ yielding

$$(+)\qquad \prod_{j=1}^{N}(1+r|\lambda_j|) \leq \prod_{j=1}^{N}(1+rs_j) .$$

To estimate the trace of T we approximate T by finite dimensional operators, that is, cut-downs of T by projections.
Let $\{h_n\}$ be an arbitrary orthonormal basis of the Hilbert space, and let P_N be the orthogonal projection onto $\operatorname{span}\{h_1, \dots, h_N\}$. Let T_N be the cut-down of T by P_N,

$$T_N := P_N T P_N .$$

Then the following holds:

Lemma 9 Suppose that T is a trace class operator with no nonzero eigenvalues. Let $T_N = P_N T P_N$ as before. Then

(i) T_N approaches T uniformly, that is

$$\lim_{N\to\infty} \|T_N - T\| = 0.$$

(ii) $\lim_{N\to\infty} \operatorname{tr} T_N = \operatorname{tr} T$

(iii) Denote the spectral radius of T_N by σ_N.

(Recall that $\sigma_N = \overset{\max}{\sup}\{|\lambda| ; \lambda \in \sigma(T_N)\}$).

Then $\sigma_N \longrightarrow 0$ as $N \to \infty$.

Proof:

(i) This is true for any compact operator T.

Indeed, let's prove it for rank one-operators, because we know that finite rank operators are dense in the compact operators.

So let $T_{e,f}\, x = \langle x, e\rangle f$, $\|e\|=\|f\|=1$. We then wish to show that $\|P_N T_{e,f} P_N - T_{e,f}\| \to 0$, $N \to \infty$.

We find that

$$P_N T_{e,f}(P_N x) = P_N \langle P_N x, e\rangle f = \langle x, P_N e\rangle P_N f = T_{P_N e, P_N f}\, x$$

and

$$\|T_{e,f} - T_{P_N e, P_N f}\| = \sup_{\|x\|=1} \|\langle x,e\rangle f - \langle x, P_N e\rangle P_N f\|$$

$$= \sup_{\|x\|=1} \|\langle x,e\rangle f - \langle x, P_N e\rangle f + \langle x, P_N e\rangle f - \langle x, P_N e\rangle P_N f\|$$

$$\overset{\boxed{\Delta\text{-ineq}}}{\leq} \sup_{\|x\|=1} \left(\|\langle x, e - P_N e\rangle f\| + \|\langle x, P_N e\rangle (f - P_N f)\|\right)$$

$$\overset{\boxed{C\text{-}S}}{\leq} \underset{1}{\underset{\parallel}{\|x\|}}\, \underset{0}{\underset{\downarrow}{\|e - P_N e\|}}\, \underset{1}{\underset{\parallel}{\|f\|}} + \underset{1}{\underset{\parallel}{\|x\|}}\, \underset{\leq 1}{\|P_N e\|}\, \underset{0}{\underset{\downarrow}{\|f - P_N f\|}} \longrightarrow 0,$$

as wanted.

T trace-class. Then $\operatorname{tr} T := \lim_{N\to\infty} \sum_{n=1}^{N} \langle Te_n, e_n \rangle$ $\{e_n\}$ ONB

(ii) This is the definition of trace,

(iii) Since T has no non-zero eigenvalues $T-\lambda$ is invertible for all $\lambda \neq 0$.
Given any $\delta > 0$ we define

$$m(\delta) = \max_{|\lambda| \geq \delta} \|(T-\lambda)^{-1}\|.$$

By (i) we can choose $M(\delta)$ so large that for $N > M(\delta)$

$$\|T_N - T\| < \frac{1}{m(\delta)}.$$

For such N and $|\lambda| \geq \delta$ we have

$$\|(T_N - T)(T-\lambda)^{-1}\| \leq \|T_N - T\| \, \|(T-\lambda)^{-1}\|$$
$$< \frac{1}{m(\delta)} m(\delta) = 1,$$

so if we write

$$T_N - \lambda = T_N - T + T - \lambda = \left((T_N - T)(T-\lambda)^{-1} + I \right)(T-\lambda),$$

we see that this is invertible when $|\lambda| \geq \delta$.
Therefore $\sigma_N < \delta$, and since δ was arbitrary > 0 this proves the assertion.

$\square$

Denote the eigenvalues of T_N as $\lambda_j^{(N)}$, $j = 1, \ldots, N$, and denote by D_N the polynomial

$$D_N(\lambda) := \prod_{j=1}^{N} (1-\lambda\lambda_j^{(N)})$$

Lemma 10 We have
$$\lim_{N\to\infty} D_N(\lambda) = e^{-\lambda\alpha}, \quad \alpha = \operatorname{tr} T$$
uniformly on every bounded set of complex numbers λ

Proof: Taking logarithm of $D_N(\lambda)$ gives
$$\log D_N(\lambda) = \sum_{j=1}^{N} \log(1-\lambda\lambda_j^{(N)}),$$
and then taking derivatives wrt. λ gives
$$\frac{D_N'}{D_N} = -\sum_{j=1}^{N} \frac{\lambda_j^{(N)}}{1-\lambda\lambda_j^{(N)}}.$$

We now (for $|\lambda| < 1/\sigma_N$) expand each term in the series on the right hand side as a geometric series; giving us
$$\frac{D_N'}{D_N} = -\sum_{j=1}^{N} \frac{\lambda_j^{(N)}}{1-\lambda\lambda_j^{(N)}} = -\frac{1}{\lambda}\sum_{j=1}^{N} \frac{\lambda\lambda_j^{(N)}}{1-\lambda\lambda_j^{(N)}} = -\frac{1}{\lambda}\sum_{j=1}^{N}\sum_{k=1}^{\infty} (\lambda\lambda_j^{(N)})^k$$
$$= -\sum_{j=1}^{N}\sum_{k=1}^{\infty} \lambda^{k-1}\lambda_j^{(N)k} \quad (\text{possible since } |\lambda_j^{(N)}| \le \sigma_N).$$

Set $S_k^{(N)} := \sum_{j=1}^{N} \lambda_j^{(N)k}$, then

(8) $$\frac{D_N'}{D_N} = -\sum_{k=1}^{\infty} S_k^{(N)} \lambda^{k-1}.$$

trace formula for matrices

For $k=1$ we notice that $S_1^{(N)} = \sum_{j=1}^{N} \lambda_j^{(N)} = \operatorname{tr} T_N$.
For $k>1$ we estimate $S_k^{(N)}$ as follows:

$$|S_k^{(N)}| = \left|\sum_{j=1}^{N} \lambda_j^{(N)k}\right| \le \sum_{j=1}^{N} |\lambda_j^{(N)k}| \le \sum_{j=1}^{N} |\lambda_j^{(N)}|\,\sigma_N^{k-1}$$
$$= \sigma_N^{k-1} \sum_{j=1}^{N} |\lambda_j^{(N)}|$$

Now we use inequality (**) to T_N, that is,

$$\sum_{j=1}^{M} |\lambda_j^{(N)}| \le \sum_{j=1}^{M} s_j(T_N) \quad \text{and get, still } k>1,$$

$$|S_k^{(N)}| \le \sigma_N^{k-1} \sum_{j=1}^{M} |\lambda_j^{(N)}| \le \sigma_N^{k-1} \sum_{j=1}^{M} s_j(T_N) \le \sigma_N^{k-1} \|T_N\|_{tr}$$

$$\le \sigma_N^{k-1} \|T\|_{tr}.$$

We rewrite (✿)

$$\frac{D_N'}{D_N} + trT = trT - S_1^{(N)} - \sum_{k=2}^{\infty} S_k^{(N)} \lambda^{k-1}$$

$$= trT - trT_N - \sum_{k=2}^{\infty} S_k^{(N)} \lambda^{k-1},$$

and taking absolute values and using the estimate we just made, we get, for $|\lambda| < 1/\sigma_N$,

$$\left|\frac{D_N'}{D_N} + trT\right| = \left| trT - trT_N - \sum_{k=2}^{\infty} S_k^{(N)} \lambda^{k-1} \right|$$

$$\le \left| trT - trT_N - \sum_{k=1}^{\infty} \sigma_N^{k} \|T\|_{tr} \lambda^{k} \right|$$

$$\le \left| trT - trT_N - \|T\|_{tr} \sum_{k=1}^{\infty} \sigma_N^{k} \lambda^{k} \right|$$

$$= |trT - trT_N| + \|T\|_{tr} \left| \frac{1}{1-\sigma_N \lambda} - 1 \right|$$

$$= |trT - trT_N| + \|T\|_{tr} \left| \frac{\sigma_N \lambda}{1-\sigma_N \lambda} \right|$$

$$\le |trT - trT_N| + \|T\|_{tr} \frac{\sigma_N |\lambda|}{1-\sigma_N |\lambda|}$$

Now let $N \to \infty$. From part (ii) and (iii) of Lemma 9 we have $\sigma_N \to 0$ and $\mathrm{tr}T_N \to \mathrm{tr}T$, therefore we get that,

$$\lim_{N\to\infty} \left| \frac{D_N'}{D_N} + \mathrm{tr}T \right| = 0$$

uniformly for all λ in a compact set.
Integrating this wrt. λ we get (since the convergence is uniform we can take the limit out of the integral)
$\lim_{N\to\infty} \int_0^\lambda \frac{D_N'}{D_N} + \mathrm{tr}T \, d\lambda' = 0$, and calculating the lefthandside yields

$$0 = \lim_{N\to\infty} \int_0^\lambda \frac{D_N'}{D_N} + \mathrm{tr}T \, d\lambda' = \lim_{N\to\infty} \left[\log D_N(\lambda') + \lambda' \mathrm{tr}T\right]_0^\lambda$$

$$\boxed{D_N(0)=1} = \lim_{N\to\infty} \left(\log D_N(\lambda) + \lambda \,\mathrm{tr}T \right)$$

Using continuity of $f(x) = e^x$ this implies

$$e^0 = e^{\lim_{N\to\infty} (\log D_N(\lambda) + \lambda \mathrm{tr}T)} = \lim_{N\to\infty} D_N(\lambda) e^{\lambda \mathrm{tr}T}$$

which means exactly that

$$\lim_{N\to\infty} D_N(\lambda) = e^{-\lambda \mathrm{tr}T},$$ proving the assertion.

(λ in bounded set $\subseteq \mathbb{C}$).

$\square$

We now use the definition of D_N to do the following estimate

$$|D_N(\lambda)| = \left|\prod_{j=1}^{N}(1-\lambda\lambda_j^{(N)})\right| \leq \prod_{j=1}^{N}(1+|\lambda||\lambda_j^{(N)}|)$$

Using the inequality (+) with $r=|\lambda|$ applied to T_N we get

$$\prod_{j=1}^{N}(1+|\lambda||\lambda_j^{(N)}|) \leq \prod_{j=1}^{N}(1+|\lambda|s_j(T_N)),$$

and since $s_j(T_N) \leq s_j(T)$ we find, all in all

$$|D_N(\lambda)| \leq \prod_{j=1}^{N}(1+|\lambda|s_j(T)).$$

If we then let $N \to \infty$ this yields

$$|e^{-\lambda \operatorname{tr} T}| \leq \prod_{j=1}^{\infty}(1+|\lambda|s_j(T)).$$

Here we use that $1+r \leq e^r$ on all but the first M factors to obtain

$$\begin{aligned} |e^{-\lambda \operatorname{tr} T}| &\leq \prod_{j=1}^{M}(1+|\lambda|s_j(T)) \prod_{j=M+1}^{\infty} e^{|\lambda|s_j(T)} \\ &= \prod_{j=1}^{M}(1+|\lambda|s_j(T)) \exp\left(\sum_{j=M+1}^{\infty}|\lambda|s_j(T)\right) \\ &= P_M(|\lambda|)\exp\left(|\lambda|\sum_{j=M+1}^{\infty}s_j(T)\right) \end{aligned}$$

where $P_M(|\lambda|)$ is a polynomial of degree M.

Choosing the argument of λ so that $-\lambda \operatorname{tr} T$ is positive and letting $|\lambda| \to \infty$ we deduce that

$$|\operatorname{tr} T| \le \sum_{M+1}^{\infty} s_j$$

because a polynomial grows more slowly than any exponential, so

$$\infty$$

$$1 \le P_M(|\lambda|) \left| \frac{e^{|\lambda|\varepsilon_M}}{e^{-\lambda\alpha}} \right| = P_M(|\lambda|)\, e^{|\lambda|\varepsilon_M + \lambda\alpha} \quad <0$$

so for this to hold we must have

$$|\lambda|\varepsilon_M \ge |\lambda\alpha| ,$$

implying $\varepsilon_M \ge |\alpha|$.

Since $\sum_{M+1}^{\infty} s_j \to 0$ as $M \to \infty$ it follows that

$\operatorname{tr} T = 0$, and this completes the proof of Lidskii's Lemma.

$\square$

Let K be a one-dimensional integral operator of the form

$$(*) \qquad (Ku)(s) = \int_0^1 K(s,t)\,u(t)\,dt$$

acting on the Hilbert space $L^2[0,1]$.

Then we have the following nice result:

Theorem 12. Let K be an integral operator of the form $(*)$ of trace class with a continuous kernel. Then the trace of K equals the integral of its kernel along the diagonal:

Proof: We first consider the case when the kernel is smooth.

Intermediate Theorem: An integral operator with smooth kernel is trace class.