SPECTRAL THEOREM FOR COMPACT, SELF-ADJOINT OPERATORS

ANDREAS STRÖMBERGSSON

The following theorem is known as the spectral theorem for compact, self-adjoint operators. We will give two proofs which connects as much as possible with Kreyszig's book.

Theorem 1. *Let $T : H \to H$ be a compact and self-adjoint operator on a Hilbert space H. Then there is a finite or infinite sequence $\{\lambda_n\}_{n=1}^N$ (i.e. $N \in \mathbb{Z}^+$ or $N = \infty$) of real eigenvalues $\lambda_n \neq 0$, and a corresponding orthonormal sequence $\{e_n\}_{n=1}^N$ in H such that:*

$$\begin{aligned} &\textit{(a)}\quad Te_n = \lambda_n e_n, \qquad \textit{for all } n \textit{ with } 1 \leq n \leq N;\\ &\textit{(b)}\quad \mathcal{N}(T) = \mathit{Span}(\{e_n\}_{n=1}^N)^\perp;\\ &\textit{(c)}\quad \textit{if } N = \infty, \textit{ then: } \lim_{n\to\infty} \lambda_n = 0. \end{aligned}$$

Note that this theorem is a *generalization of the spectral theorem for self-adjoint operators in finite dimensional spaces*, which you are hopefully already familiar with. [To be precise: If H is finite dimensional, say $\dim H = m$, then since $e_1, e_2, ...$ is an orthonormal sequence, and hence linearly dependent, we must have N *finite* in Theorem 1, and in fact $N \leq m$. Also $\dim \mathcal{N}(T) = m - N$ by (b), and hence (if $N > m$) we can choose an orthonormal basis $e_{N+1}, e_{N+2}, ..., e_m$ for $\mathcal{N}(T)$; since these vectors lie in $\mathcal{N}(T)$ they are *also* eigenvectors of T, with eigenvalues 0. Now $\{e_n\}_{n=1}^m$ is an ON basis for H. Hence: If $\dim H < \infty$ and T is self-adjoint, then Theorem 1 says that *H has an ON basis where each basis vector is an eigenvector of T*. This is the usual spectral theorem in finite dimension.]

Let us note in the general case that Theorem 1 gives us a *very explicit formula for T!* Namely, let $Y = \overline{\text{Span}(\{e_n\}_{n=1}^N)}$; then $Y^\perp = \text{Span}(\{e_n\}_{n=1}^N)^\perp$, so that by Theorem 1 (b) and [Kreyszig, Thm. 3.3-4], we have

$$H = Y \oplus \mathcal{N}(T).$$

Hence every vector $v \in H$ can be uniquely written as $v = y + z$ with $y \in Y$ and $z \in \mathcal{N}(T)$. Furthermore $y \in Y = \overline{\text{Span}(\{e_n\}_{n=1}^N)}$ can be uniquely expressed as

$$y = \sum_{n=1}^{N} \alpha_n e_n \qquad \text{for some } \alpha_n \in K \text{ with } \sum_{n=1}^{N} |\alpha_n|^2 < \infty.$$

This follows from [Kreyszig, Theorem 3.5-2] *in the form which I stated in class* (one uses in particular my "(d)" of that theorem, together with the fact that $\{e_n\}_{n=1}^N$ is a *total* orthonormal sequence in Y, by definition). Now $T(z) = 0$ since $z \in \mathcal{N}(T)$, and $T(y) = T(\sum_{n=1}^N \alpha_n e_n) = \sum_{n=1}^N \alpha_n T(e_n) = \sum_{n=1}^N \lambda_n \alpha_n e_n$ since T is continuous. To summarize:

Consequence from Theorem 1. *Every vector $v \in H$ can be uniquely written (recall $N \in \mathbb{Z}^+$ or $N = \infty$ as in Theorem 1)*

$$v = \Big(\sum_{n=1}^{N} \alpha_n e_n\Big) + z, \quad \textit{where } \alpha_n \in K, \ \sum_{n=1}^{N} |\alpha_n|^2 < \infty, \ z \in \mathcal{N}(T),$$

and for each such vector v we have

$$T(v) = \sum_{n=1}^{N} \lambda_n \alpha_n e_n.$$

We will give two proofs of Theorem 1. The first proof is modelled on one of the standard methods of proofs *in the finite dimensional case,* although the details are of course more difficult in infinite dimension: Namely, one starts by finding a unit vector $x_0 \in H$ with $||Tx_0|| = ||T||$ (see Lemma 2 below); very often x_0 itself is an eigenvector of T, and if it is not then one can still easily construct an eigenvector from x_0 (see Lemma 3 below). Once *one* eigenvector has been found, one can restrict attention to the *orthogonal complement* of this eigenvector, and repeat the whole procedure. In the end one has found all eigenvectors.

The second proof uses the much more advanced spectral representation theorem of self-adjoint operators proved in Chapter 9, and I give it since it is satisfactory to see that Theorem 1 is indeed a special case of this, and also since it gives and opportunity to become more familiar with the language and tools used in Chapter 9. If you take the course Functional Analysis for 4p, then please ignore the second proof! (Note: the second proof is probably what the author had in mind in [Kreyszig, §9.9: Problem 6].

First proof of Theorem 1.

Lemma 1. *Let X be a* reflexive *normed space and assume that X'* is separable. *Then every bounded sequence (x_n) in X has a subsequence which is weakly convergent.*

Proof. Let (x_n) be a bounded sequence in X. Recall that "bounded" means that there is a number $B > 0$ such that $||x_n|| \leq B$ for all n. Let $\{f_1, f_2, ...\}$ be a countable dense subset of X' (this exists because X' is separable). Now $|f_1(x_j)| \leq ||f_1|| \cdot ||x_j|| \leq ||f_1|| \cdot B$ for all j, i.e. $(f_1(x_j))$ is a bounded sequence of numbers. Hence by Bolzano-Weierstrass theorem there is a subsequence which converges, i.e. there is a subsequence $(x_{1,j})$ of (x_j) such that $\lim_{j\to\infty} f_1(x_{1,j})$ exists. By the same argument we see that $(x_{1,j})$ has a subsequence $(x_{2,j})$ such that $\lim_{j\to\infty} f_2(x_{2,j})$ exists. This argument is repeated to form sequences $(x_{3,j})$, $(x_{4,j})$, ..., which are successive subsequences of each other. Finally, we let (y_m) be the "diagonal sequence", i.e. $y_m = x_{m,m}$ for all $m \geq 1$.

Now (y_m) is a subsequence of (x_j), and for each $n \geq 1$ the limit $\lim_{m\to\infty} f_n(y_m)$ exists. (Proof: Fix any $n \geq 1$. Then the sequence $(y_n, y_{n+1}, y_{n+2}, ...)$ is by construction a subsequence of $(x_{n,1}, x_{n,2}, x_{n,3}, ...)$, and we know that $\lim_{j\to\infty} f_n(x_{n,j})$ exists. Hence $\lim_{m\to\infty} f_n(y_m)$ exists.)

Now let f be an *arbitrary* element in X'. Let $\varepsilon > 0$. Then since $\{f_1, f_2, ...\}$ is dense in X' there is some n such that $||f_n - f|| < \frac{\varepsilon}{10B}$. Furthermore, by what we have proved, the sequence $\{f_n(y_k)\}_{k=1}^{\infty}$ is Cauchy and hence there is some $K \geq 1$ such that $|f_n(y_k) - f_n(y_{k'})| < \frac{\varepsilon}{10}$ for all $k, k' \geq K$. Hence, for all $k, k' \geq K$,

$$
\begin{aligned}
&|f(y_k) - f(y_{k'})| = \Big|(f - f_n)(y_k) + (f_n(y_k) - f_n(y_{k'})) + (f_n - f)(y_{k'})\Big| \\
&\leq \Big|(f - f_n)(y_k)\Big| + \Big|f_n(y_k) - f_n(y_{k'})\Big| + \Big|(f_n - f)(y_{k'})\Big| \\
&< ||f - f_n|| \cdot ||y_k|| + \frac{\varepsilon}{10} + ||f - f_n|| \cdot ||y_{k'}|| \\
&\leq \frac{\varepsilon}{10B} \cdot B + \frac{\varepsilon}{10} + \frac{\varepsilon}{10B} \cdot B < \varepsilon.
\end{aligned}
$$

But here $\varepsilon > 0$ was arbitrary; hence the sequence $\{f(y_k)\}_{k=1}^{\infty}$ is a Cauchy sequence, and thus $\lim_{k\to\infty} f(y_k)$ exists. We may now define:

$$g(f) := \lim_{k\to\infty} f(y_k), \qquad \text{for each } \ f \in X'.$$

Then g is a map $g : X' \to K$, and it is easily checked that g is linear. Furthermore, for each $f \in X'$ we have

$$|g(f)| = \lim_{k\to\infty} |f(y_k)| \leq \limsup_{k\to\infty} ||f|| \cdot ||y_k|| \leq B \cdot ||f||.$$

Hence g is a bounded linear functional, $g \in X''$, and $||g|| \leq B$. Now, *since X is reflexive*, there exists some $x \in X$ such that $g = g_x$, where (as in [Kreyszig, p. 239]) $g_x(f) := f(x), \forall f \in X'$. Hence, for all $f \in X'$,

$$f(x) = g_x(f) = g(f) = \lim_{k\to\infty} f(y_k).$$

Hence the sequence y_k converges weakly to $x \in X$. $\square$

Lemma 2. *Let $T : H \to H$ be a compact and self-adjoint operator on a Hilbert space H. Let $S_1 = S(0;1)$ be the unit sphere in H. Then there is a vector $x_0 \in S_1$ such that $||Tx_0|| = ||T||$.*

Proof. By the definition of $||T||$ as $||T|| = \sup_{x \in S_1} ||Tx||$, there exists a sequence $x_1, x_2, ...$ in S_1 such that $\lim_{n\to\infty} ||Tx_n|| = ||T||$. Since T is compact we may assume, after having replaced $\{x_n\}$ by a subsequence (and changed notation so that $\{x_n\}$ denotes this subsequence), that

$$y_0 = \lim_{n\to\infty} Tx_n \quad \text{exists in } H. \tag{1}$$

Let $Y = \overline{\text{Span}\{x_1, x_2, ...\}}$. This is a closed subspace of H, hence a Hilbert space. Y is reflexive by [Kreyszig, Thm. 4.6-6], and Y is separable by construction (a dense subset in Y is, e.g., $\{q_1x_1 + q_2x_2 + ... + q_nx_n \mid 1 \leq n < \infty, q_j \in \mathbb{Q}\}$). Hence Lemma 1 applies, so that there is a subsequence of (x_n), say $(x_{j_1}, x_{j_2}, x_{j_3}, ...)$ with $1 \leq j_1 < j_2 < j_3 < ...$, such that x_{j_k} converge weakly *in* Y to some point $x_0 \in Y$. Now, for each $z \in Y$,

$$\langle Tx_0, z\rangle = \langle x_0, Tz\rangle = \lim_{k\to\infty} \langle x_{j_k}, Tz\rangle = \lim_{k\to\infty} \langle Tx_{j_k}, z\rangle = \langle y_0, z\rangle. \tag{2}$$

Here the *second* equality holds since x_{j_k} converge weakly to x_0 (for we know that $x \mapsto \langle x, Tz\rangle$ is a bounded linear functional on Y), and the *last* equality holds because we know from (1) that $\lim_{k\to\infty} Tx_{j_k} = y_0$ exists, and $\langle \cdot, \cdot \rangle$ is continuous ([Kreyszig, Lemma 3.2-2]). Since (2) holds for all $z \in Y$, we conclude $Tx_0 = y_0$. Hence

$$||Tx_0|| = ||y_0|| = || \lim_{k\to\infty} T(x_n)|| = \lim_{k\to\infty} ||T(x_n)|| = ||T||.$$

Also note that since x_{j_k} converge weakly to x_0 we have $\langle x_0, x_0\rangle = \lim_{k\to\infty}\langle x_{j_k}, x_0\rangle$, and $|\langle x_{j_k}, x_0\rangle| \leq ||x_{j_k}|| \cdot ||x_0|| = 1$ for all k, hence $||x_0|| \leq 1$. We *cannot* have $||x_0|| < 1$, since then $||Tx_0|| \leq ||T|| \cdot ||x_0|| < ||T||$, contrary to what we have seen above. Hence $||x_0|| = 1$, i.e. $x_0 \in S_1$. This shows that x_0 has all the required properties. $\square$

Lemma 3. *Let notation be as in Lemma 2. Then T has an eigenvector with eigenvalue $||T||$ or $-||T||$.*

Proof. We write $\lambda = ||T||$. We may assume $\lambda > 0$, for if $||T|| = \lambda = 0$ then $T = 0$ and we can choose *any* vector $y \in H - \{0\}$ as eigenvector.

By Lemma 2 there is a point $x_0 \in S_1$ such that $||Tx_0|| = \lambda$. Since T is self-adjoint we have

$$\langle T^2 x_0, x_0\rangle = \langle Tx_0, Tx_0\rangle = ||Tx_0||^2 = \lambda^2 \tag{3}$$

On the other hand, by Schwarz inequality we have

$$\big|\langle T^2 x_0, x_0\rangle\big| \le ||T^2 x_0|| \cdot ||x_0|| \le ||T^2|| \cdot 1 \le ||T||^2 = \lambda^2. \tag{4}$$

Because of (3), we must have *equalities* throughout in (4). In particular we have equality in Schwarz inequality $|\langle T^2 x_0, x_0\rangle| \le ||T^2 x_0|| \cdot ||x_0||$, and hence by [Kreyszig, Lemma 3.2-1(a)], x_0 and $T^2 x_0$ are linearly independent; since $x_0 \neq 0$ this implies $T^2 x_0 = \alpha x_0$ for some $\alpha \in K$. Plugging this into (3) we get $\alpha = \lambda^2$, i.e. $T^2 x_0 = \lambda^2 x_0$. Using this we compute:

$$T(\lambda x_0 + Tx_0) = \lambda(\lambda x_0 + Tx_0); \qquad T(\lambda x_0 - Tx_0) = -\lambda(\lambda x_0 - Tx_0).$$

This means that *if* $\lambda x_0 + Tx_0 \neq 0$ then $\lambda x_0 + Tx_0$ is an eigenvector of T with eigenvalue λ, and *if* $\lambda x_0 - Tx_0 \neq 0$ then $\lambda x_0 - Tx_0$ is an eigenvector of T with eigenvalue $-\lambda$. But the sum of these two vectors is $2\lambda x_0 \neq 0$, hence at least one of the vectors must be non-zero. Hence there is an eigenvector with eigenvalue λ or an eigenvector with eigenvalue $-\lambda$. $\square$

Lemma 4. *Let $T : H \to H$ be a bounded self-adjoint operator on a Hilbert space H, and let $Y \subset H$ be a subspace such that $T(Y) \subset Y$. Then $T(Y^\perp) \subset Y^\perp$, and $T_{|Y^\perp} : Y^\perp \to Y^\perp$ is a bounded self-adjoint linear operator on the Hilbert space $Y^\perp$, with norm $||T_{|Y^\perp}|| \le ||T||$.*

Proof. Take $z \in Y^\perp$. For each $y \in Y$:

$$\langle Tz, y\rangle = \langle z, Ty\rangle = 0,$$

since $Ty \in T(Y) \subset Y$ and $z \in Y^\perp$. It follows that $Tz \in Y^\perp$. Hence we have proved $T(Y^\perp) \subset Y^\perp$. The remaining statements are obvious. $\square$

Proof of Theorem 1. If $T = 0$ then the theorem is trivial; we simply choose $N = 0$ so that $\{\lambda_n\}_{n=1}^N$ and $\{e_n\}_{n=1}^N$ are empty sequences.

Now assume $T \neq 0$. Then by Lemma 3 there is an eigenvector e_1 with eigenvalue $\lambda_1 = ||T||$ or $\lambda_1 = -||T||$. Since $T \neq 0$ we have $\lambda_1 \neq 0$. After scaling we may assume $||e_1|| = 1$. Let $H_1 = \text{Span}\{e_1\}^\perp$. Then by Lemma 4 (for $Y = \text{Span}\{e_1\}$), the restriction $T_{|H_1}$ is a self-adjoint operator $T_{|H_1} : H_1 \to H_1$, and its norm is $||T_{|H_1}|| \le ||T|| = \lambda_1$.

If $T_{|H_1} = 0$ then we stop here. If not, then we repeat the above application of Lemma 3 over and over: The above was **step** 1. In general, after **step** $n-1$ we have found an orthonormal sequence of eigenvectors $e_1, e_2, ..., e_{n-1}$ in H with corresponding real eigenvalues $\lambda_1, \lambda_2, ..., \lambda_{n-1}$, such that $|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_{n-1}| > 0$ and such that if $H_{n-1} = \text{Span}\{e_1, ..., e_{n-1}\}^\perp$ then $T_{|H_{n-1}}$ is a self-adjoint operator $T_{|H_{n-1}} : H_{n-1} \to H_{n-1}$ of norm $||T_{|H_{n-1}}|| \leq |\lambda_{n-1}|$. If $T_{|H_{n-1}} = 0$ then we stop after this step. Otherwise, if $T_{|H_{n-1}} \neq 0$, then we continue to **step** n, which consists in applying Lemma 3 to the operator $T_{|H_{n-1}} : H_{n-1} \to H_{n-1}$; we thus find an eigenvector $e_n \in H_{n-1}$ with eigenvalue $\lambda_n = ||T_{|H_{n-1}}||$ or $\lambda_n = -||T_{|H_{n-1}}||$. By our assumptions we then have $0 < |\lambda_n| \leq |\lambda_{n-1}|$. After scaling we may also assume $||e_n|| = 1$. Since $e_n \in H_{n-1} = \text{Span}\{e_1, ..., e_{n-1}\}^\perp$ we also have $\langle e_n, e_k \rangle = 0$ for $k = 1, 2, ..., n-1$. Hence $e_1, e_2, ..., e_n$ is an orthonormal sequence. Let $H_n = \text{Span}\{e_1, ..., e_n\}^\perp$. Then by Lemma 4 (for $Y = \text{Span}\{e_1, ..., e_n\}$), the restriction $T_{|H_n}$ is a self-adjoint operator $T_{|H_n} : H_n \to H_n$. Since $H_n = \text{Span}\{e_1, ..., e_n\}^\perp \subset \text{Span}\{e_1, ..., e_{n-1}\}^\perp = H_{n-1}$ (see [Kreyszig, p. 150 exercise 7 (b)]), we have $||T_{|H_n}|| \leq ||T_{|H_{n-1}}|| = |\lambda_n|$. Hence we are in the same situation as after **step** $n-1$, and the process can be repeated.

If the above process stops after **step** N we have $T_{|H_N} = 0$; hence $\text{Span}(\{e_n\}_{n=1}^N)^\perp = H_N \subset \mathcal{N}(T)$. On the other hand, if $x \in \mathcal{N}(T)$ then for each n,

$$\lambda_n \langle e_n, x \rangle = \langle Te_n, x \rangle = \langle e_n, Tx \rangle = \langle e_n, 0 \rangle = 0, \tag{5}$$

and hence $x \perp e_n$, since $\lambda_n \neq 0$. Hence $x \perp \text{Span}(\{e_n\}_{n=1}^N)^\perp$. This proves that $\mathcal{N}(T) = \text{Span}(\{e_n\}_{n=1}^N)^\perp$, and hence Theorem 1 holds in this case.

Now assume that the process never stops. We then obtain an infinite sequence $\{\lambda_n\}_{n=1}^\infty$ of real non-zero eigenvalues with $|\lambda_1| \geq |\lambda_2| \geq ...$, and a corresponding orthonormal sequence $\{e_n\}_{n=1}^\infty$ of eigenvectors (i.e. $Te_n = \lambda_n e_n$ for all $n \geq 1$). It now remains to prove that (b) and (c) hold in Theorem 1. Since T is compact and $\{e_n\}_{n=1}^\infty$ is a bounded sequence in H, [Kreyszig, Thm. 8.1-3] says that there is a subsequence $1 \leq n_1 < n_2 < ...$ such that (Te_{n_j}) converges in H as $j \to \infty$. Since every convergent sequence is Cauchy we have

$$||Te_{n_j} - Te_{n'_j}|| \to 0 \qquad \text{as } j, j' \to \infty. \tag{6}$$

But $Te_{n_j} = \lambda_{n_j} e_{n_j}$, and these vectors are mutually orthogonal for distinct j's. Hence by Pythagoras' formula we have for all $j < j'$,

$$||Te_{n_j} - Te_{n'_j}||^2 = ||\lambda_{n_j} e_{n_j}||^2 + ||\lambda_{n_{j'}} e_{n_{j'}}||^2 = |\lambda_{n_j}|^2 + |\lambda_{n_{j'}}|^2. \tag{7}$$

Together, (6) and (7) imply that $\lim_{j\to\infty} |\lambda_{n_j}|^2 = 0$. Hence, since $|\lambda_1| \geq |\lambda_2| \geq ...$, the *full* sequence (λ_n) converges to 0, i.e. (c) holds.

To prove (b), first note that if $x \in \mathcal{N}(T)$ then $x \perp e_n$ for each n, using (5) and $\lambda_n \neq 0$. Hence $\mathcal{N}(T) \subset \operatorname{Span}(\{e_n\}_{n=1}^{\infty})^{\perp}$. Conversely, assume $x \in \operatorname{Span}(\{e_n\}_{n=1}^{\infty})^{\perp}$. Then $x \in H_n$ for all n, and since $||T_{|H_n}|| \leq |\lambda_n|$ we have $||Tx|| \leq |\lambda_n| \cdot ||x||$. This is true for all n, and $\lim_{n\to\infty} \lambda_n = 0$, hence $||Tx|| = 0$, i.e. $x \in \mathcal{N}(T)$. Hence $\mathcal{N}(T) \supset \operatorname{Span}(\{e_n\}_{n=1}^{\infty})^{\perp}$, and thus $\mathcal{N}(T) = \operatorname{Span}(\{e_n\}_{n=1}^{\infty})^{\perp}$. Now (b) is completely proved.
$\square$

Second proof of Theorem 1

Let us restrict to the case $K = \mathbb{C}$, since Kreyszig only works with this case in his Chapter 9.

Let $T : H \to H$ be as in Theorem 1, i.e. T is a compact and self-adjoint operator. Since T is compact it is bounded, and hence the Spectral Theorem [Kreyszig, Thm. 9.9-1] applies to T, i.e. if (E_λ) is the spectral family associated with T then

$$T = \int_{m-0}^{M} \lambda \, dE_\lambda$$

Here

$$m = \inf_{||x||=1} \langle Tx, x \rangle, \qquad M = \inf_{||x||=1} \langle Tx, x \rangle,$$

and we know that (E_λ) is a spectral family *on the interval* $[m, M] \subset \mathbb{R}$.

For convenience we fix some arbitrary real numbers $A < 0 < B$ such that $A < m$ and $M < B$; then

$$T = \int_{A}^{B} \lambda \, dE_\lambda$$

Let us write $Y_\lambda = E_\lambda(H)$; this is a closed subspace of H, and E_λ is the projection of H onto Y_λ

Let us assume that there is some $\lambda_0 > 0$ *such that* $\dim Y_{\lambda_0}^{\perp} = \infty$. (We will show that this leads to a contradiction.) We have $E_\lambda E_{\lambda_0} = E_\lambda$ for all $\lambda \leq \lambda_0$ and $E_\lambda E_{\lambda_0} = E_{\lambda_0}$ for all $\lambda \geq \lambda_0$ (see [Kreyszig, Thm. 9.6-1]); hence $TE_{\lambda_0} = \left(\int_A^B \lambda \, dE_\lambda\right) E_{\lambda_0} = \int_A^{\lambda_0} \lambda \, dE_\lambda$ (since the corresponding identity holds for each approximating Riemann sum) and thus

$$T(I - E_{\lambda_0}) = \int_{\lambda_0}^{B} \lambda \, dE_\lambda \geqq \int_{\lambda_0}^{B} \lambda_0 \, dE_\lambda = \lambda_0(E_B - E_{\lambda_0}) = \lambda_0(I - E_{\lambda_0}).$$

(The $\geqq$-relation holds because it is clearly true for each approximating Riemann sum.) But note that $(I - E_{\lambda_0})$ is projection of H onto $Y_{\lambda_0}^{\perp}$, hence for each vector $y \in Y_{\lambda_0}^{\perp}$ we have $(I - E_{\lambda_0})(y) = y$ and thus, by the definition of "$\geqq$" ([Kreyszig, p. 470(1)]), $\langle Ty, y \rangle \geq \langle \lambda_0 y, y \rangle = \lambda_0 ||y||^2$.

Combining this with Schwarz' inequality, $||Ty|| \cdot ||y|| \geq \langle Ty, y\rangle$, we have proved

$$||Ty|| \geq \lambda_0 ||y||, \qquad \forall y \in Y_{\lambda_0}^{\perp}. \tag{8}$$

But $\dim Y_{\lambda_0}^{\perp} = \infty$; hence there is an infinite orthonormal sequence $e_1, e_2, ...$ in $Y_{\lambda_0}^{\perp}$, and from (8) we get, for all $j \neq i$,

$$||Te_j - Te_i|| \geq \lambda_0 ||e_j - e_i|| = \sqrt{2} \cdot \lambda_0.$$

This shows that there can *not* exist any subsequence $j_1 < j_2 < ...$ such that (Te_{j_ℓ}) is Cauchy. But (e_j) is a bounded sequence in H, hence this *contradicts* the fact that T is compact (see [Kreyszig, Thm. 8.1-3]).

This contradiction shows that we must discard our assumption that $\dim Y_{\lambda_0}^{\perp} = \infty$. Hence we have proved that $\dim Y_{\lambda_0}^{\perp}$ is *finite* for each $\lambda_0 > 0$.

Hence $d(\lambda) = \dim Y_{\lambda}^{\perp}$ is a function from $\mathbb{R}^+$ to $\mathbb{Z}_{\geq 0}$, the set of non-negative integers. Since $d(\lambda)$ is decreasing and right continuous (by the definition of spectral family), we see that for each $n \in \mathbb{Z}_{\geq 0}$ the set $\{\lambda > 0 \mid d(\lambda) = n\}$ is either empty or an interval of the form $\mathbb{R}^+ \cap [\mu, \mu')$. It follows that there is an infinite sequence $B = \mu_1 > \mu_2 > ...$ of positive numbers such that $d(\lambda)$ *is constant on each interval* $[\mu_j, \mu_{j-1})$, and $\lim_{j \to \infty} \mu_j = 0$. (Note that we do *not* require $d(\lambda)$ to have a jump discontinuity at *each* point μ_j; the reason for this is that we want to have a notation which works also if $\lim_{\lambda \to 0+} d(\lambda)$ is finite: In this case we may choose $B = \mu_1 > \mu_2 > ... > \mu_M > 0$ so that $d(\lambda)$ is constant on $(0, \mu_M)$ and on each $[\mu_j, \mu_{j-1})$; then add extra "dummy" points tending to 0, say $\mu_{M+j} = 2^{-j}\mu_M$ for $j = 1, 2, 3, ...$)

It now follows easily from the definition of the "spectral Riemann-Stieltjes integral" that:

$$\int_0^B \lambda \, dE_\lambda = \sum_{j=1}^{\infty} \mu_j (E_{\mu_j} - E_{\mu_{j+1}}) \qquad \text{(convergence in } B(H, H)\text{)}. \tag{9}$$

[Detailed proof: Let P_n be the partition of $[0, B]$ which is obtained by splitting at the points $0 < \mu_n < \mu_{n-1} < ... < \mu_1 = B$ and then, for each $2 \leq j \leq n$, *subdividing* the interval (μ_j, μ_{j-1}) into n equal parts. Then by construction $\eta(P_n) \to 0$ as $n \to \infty$, and hence $\lim_{n \to \infty} s(P_n) = \int_0^B \lambda \, dE_\lambda$. But $d(\lambda)$ is constant on each interval $[\mu_j, \mu_{j-1})$ and hence so is E_λ. Hence the contribution from the interval $[\mu_j, \mu_{j-1})$ to $s(P_n)$ equals $\mu_j(E_{\mu_j} - E_{\mu_{j+1}})$, regardless of the subdivision of this interval. Thus $s(P_n) = \mu_n(E_{\mu_n} - E_0) + \sum_{j=1}^{n-1} \mu_j (E_{\mu_j} - E_{\mu_{j+1}})$. But $E_{\mu_n} - E_0$ is a projection and hence of norm ≤ 1, so that

$$||s(P_n) - \sum_{j=1}^{n-1} \mu_j(E_{\mu_j} - E_{\mu_{j+1}})|| = ||\mu_n(E_{\mu_n} - E_0)|| \leq \mu_n \to 0 \quad \text{as } n \to \infty.$$

From this fact together with $\lim_{n\to\infty} s(P_n) = \int_0^B \lambda\, dE_\lambda$ it follows that $\lim_{n\to\infty} \sum_{j=1}^{n-1} \mu_j(E_{\mu_j} - E_{\mu_{j+1}}) = \int_0^B \lambda\, dE_\lambda$, i.e. (9) holds.]

Now for each $j \geq 1$, $E_{\mu_j} - E_{\mu_{j+1}}$ is the projection of H onto $Z_j := Y_{\mu_j} \ominus Y_{\mu_{j+1}}$, the orthogonal complement of $Y_{\mu_{j+1}}$ in Y_{μ_j} (see [Kreyszig, Thm. 9.6-2(b)]). Note that $Z_j = Y_{\mu_{j+1}}^{\perp} \ominus Y_{\mu_j}^{\perp}$ and hence Z_j has finite dimension, $\dim Z_j = d(\mu_{j+1}) - d(\mu_j)$ (this may be 0). The spaces $Z_1, Z_2, ...$ are mutually orthogonal, for if $1 \leq j < i$ then $\mu_{j+1} \geq \mu_i$ and $Z_i \subset Y_{\mu_i} \subset Y_{\mu_{j+1}}$, and $Y_{\mu_{j+1}}$ is orthogonal to Z_j by definition. Hence if we choose an ON basis $e_1, ..., e_{d_1}$ in Z_1, and then choose an ON basis $e_{d_1+1}, ..., e_{d_1+d_2}$ in Z_2, etc., we obtain an orthonormal sequence $e_1, e_2, ...$ in H. This sequence may be finite or infinite. Since $E_{\mu_j} - E_{\mu_{j+1}}$ is projection onto Z_j we have $(E_{\mu_j} - E_{\mu_{j+1}})(x) = \sum_{e_k \in Z_j} \langle x, e_k \rangle e_k$, for each j. Hence, if we define $\lambda_k := \mu_j$ for each $e_k \in Z_j$, it follows from (9) that

$$\Big(\int_0^B \lambda\, dE_\lambda\Big)(x) = \sum_k \lambda_k \langle x, e_k \rangle e_k, \qquad \forall x \in H. \tag{10}$$

Note that by our construction we have $\lambda_1 \geq \lambda_2 \geq ...$, all these numbers are positive, and if this sequence is infinite then $\lim_{k\to\infty} \lambda_k = 0$. Note also $Z_j \subset Y_0^{\perp}$ for each j, since $Z_j \perp Y_{\mu_{j+1}}$ and $Y_0 \subset Y_{\mu_{j+1}}$.

Negative λ can be treated in a very similar way, and we will just give an outline: To start, let $\lambda_0 < 0$ be arbitrary. One notes $TE_{\lambda_0} \leqq \int_A^{\lambda_0} \lambda_0\, dE_\lambda = \lambda_0 E_{\lambda_0}$, and hence $\langle Ty, y \rangle \leq \lambda_0 ||y||^2 < 0$ for all $y \in Y_{\lambda_0}$; by Schwarz' inequality this leads to

$$||Ty|| \geq |\lambda_0| \cdot ||y||, \forall y \in Y_{y_0}.$$

Using this one proves by a similar argument as before that $\dim Y_{\lambda_0} < \infty$. (Note: now this is for Y_{λ_0}, *not* $Y_{\lambda_0}^{\perp}$!) Using this one can see that there is an infinite sequence $A = \nu_1 < \nu_2 < ...$ of negative numbers such that $\dim Y_\lambda$ *is constant on each interval* $[\nu_j, \nu_{j+1})$, and $\lim_{j\to\infty} \nu_j = 0$. It follows that

$$\int_A^0 \lambda\, dE_\lambda = \sum_{j=2}^{\infty} \nu_j(E_{\nu_j} - E_{\nu_{j-1}}) \qquad \text{(convergence in } B(H,H)). \tag{11}$$

This formula implies that there is an orthonormal (finite or infinite) sequence $f_1, f_2, ...$ in Y_0 and corresponding negative numbers $\lambda_1' \leq \lambda_2' \leq$

..., with $\lim_{k\to\infty}\lambda'_k = 0$ if the sequence is infinite, such that

$$\Big(\int_A^0 \lambda \, dE_\lambda\Big)(x) = \sum_k \lambda'_k \langle x, f_k\rangle f_k, \qquad \forall x \in H. \tag{12}$$

Together (10) and (12) give, for all $x \in H$,

$$Tx = \Big(\int_A^B \lambda \, dE_\lambda\Big)(x) = \sum_k \lambda_k \langle x, e_k\rangle e_k + \sum_k \lambda'_k \langle x, f_k\rangle f_k. \tag{13}$$

Note here that $e_k \perp f_{k'}$ for all k, k', since $e_k \in Y_0^\perp$ and $f_{k'} \in Y_0$. Hence any merging of the sequences (e_k) and (f_k) is a new orthonormal sequence. It now follows immediately from (13) that $Te_k = \lambda_k e_k$ for all e_k, and $Te_k = \lambda'_k f_k$ for all f_k. We also see from (13) that $Tx = 0$ if and only if $\langle x, e_k\rangle = 0$ and $\langle x, f_k\rangle = 0$ for all e_k, f_k, that is,

$$\mathcal{N}(t) = \operatorname{Span}(\{e_k\} \cup \{f_k\})^\perp.$$

We now see that all claims in Theorem 1 are fulfilled, except that we have slightly different notation: To obtain the notation in Theorem 1 we may simply form the "merged" sequence $e_1, f_1, e_2, f_2, \ldots$ (appropriately modified if one or both of $\{e_k\}$ and $\{f_k\}$ are finite), and the corresponding sequence of eigenvalues.

$\square$

Problems

1. Let $A : H \to H$ be a compact, self-adjoint operator. Assume that A is *positive*, i.e. $\langle Ax, x\rangle \geq 0$ for all $x \in H$. Let $n \geq 2$. Prove that there exists a bounded operator $B : H \to H$ such that $B^n = A$.

2. Prove that the compactness assumption is necessary in Lemma 2: In precise terms, construct an example of a bounded self-adjoint operator $T : H \to H$ such that $||Tx_0|| < ||T||$ for all unit vectors x_0.

References

[Kreyszig] E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, 1989.