12.1. Spectra of shifts

Let $S\colon \ell^2(\mathbb{N},\mathbb{C}) \to \ell^2(\mathbb{N},\mathbb{C})$ be the right shift on $\ell^2(\mathbb{N},\mathbb{C})$, i.e.,

$$S((x_1,x_2,x_3,\ldots)) = (0,x_1,x_2,\ldots) \quad \text{for all } (x_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N},\mathbb{C}).$$

(a) Calculate the operator norm $\|S\|_{L(\ell^2(\mathbb{N},\mathbb{C}),\ell^2(\mathbb{N},\mathbb{C}))}$ and the spectral radius r_S of S.

Solution: It holds for all $x \in \ell^2(\mathbb{N},\mathbb{C})$ that $\|Sx\|_{\ell^2(\mathbb{N},\mathbb{C})} = \|x\|_{\ell^2(\mathbb{N},\mathbb{C})}$. It follows for all $n \in \mathbb{N}$, $x \in \ell^2(\mathbb{N},\mathbb{C})$ that $\|S^n x\|_{\ell^2(\mathbb{N},\mathbb{C})} = \|x\|_{\ell^2(\mathbb{N},\mathbb{C})}$. Thus, we obtain $\|S^n\|_{L(\ell^2(\mathbb{N},\mathbb{C}),\ell^2(\mathbb{N},\mathbb{C}))} = 1$ for all $n \in \mathbb{N}$. This implies that $\|S\|_{L(\ell^2(\mathbb{N},\mathbb{C}),\ell^2(\mathbb{N},\mathbb{C}))} = 1$ and $r_S = 1$.

(b) Determine the point spectrum $\sigma_p(S)$, the continuous spectrum $\sigma_c(S)$ and the residual spectrum $\sigma_r(S)$ of S.

Solution: For $x = (x_n)_{n\in\mathbb{N}} \in \ell^2$, $\lambda \in \mathbb{C}$, the relation $\lambda x = Sx$ implies that $\lambda x_1 = 0$ and $\lambda x_{n+1} = x_n$ for every $n \in \mathbb{N}$. For $\lambda \neq 0$, this leads to $x = 0$. That is, $\sigma_p(S) \subseteq \{0\}$. Since S is an isometry, S is injective and therefore, $0 \notin \sigma_p(S)$. Hence, $\sigma_p(S) = \emptyset$.

Note that, for every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$ it holds that $x^{(\lambda)} := (\lambda^{n-1})_{n\in\mathbb{N}} \in \ell^2(\mathbb{N},\mathbb{C})$ (for $\lambda = 0$, $x^{(\lambda)} = e_1 = (1,0,0,\ldots)$) and $\lambda x_\lambda = S^* x_\lambda$. In particular, for every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$, the range of $\lambda - S$ cannot be dense as $\ker(\overline{\lambda} - S^*) \neq \{0\}$. Thus, $\{\lambda \in \mathbb{C} \mid |\lambda| < 1\} \subseteq \sigma_r(S) \cup \sigma_p(S) = \sigma_r(S)$ (and we saw during the proof that $\{\lambda \in \mathbb{C} \mid |\lambda| < 1\} \subseteq \sigma_p(S^*)$). Moreover, since $\{\lambda \in \mathbb{C} \mid |\lambda| < 1\} \subseteq \sigma(S) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$ and $\sigma(S)$ is closed (as the resolvent set is open), we know at this stage that $\sigma(S) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$.

For $\lambda \in \mathbb{C}$ with $|\lambda| = 1$, $x \in \ker(\lambda - S^*)$ implies that $\|S^* x\|_{\ell^2(\mathbb{N},\mathbb{C})} = \|\lambda x\|_{\ell^2(\mathbb{N},\mathbb{C})} = \|x\|_{\ell^2(\mathbb{N},\mathbb{C})}$, i.e., $x_1 = 0$. But this implies $x_2 = 0$, $x_3 = 0$ ... and inductively $x_n = 0$ for all $n \in \mathbb{N}$. Hence, for every $\lambda \in \mathbb{C}$ with $|\lambda| = 1$, we have that $\ker(\overline{\lambda} - S^*) = \{0\}$ (in other words, $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\} \cap \sigma_p(S^*) = \emptyset$) and, therefore, $\operatorname{im}(\lambda - S)$ is dense. Thus, $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\} \cap \sigma_r(S) = \emptyset$. Since we know already that $\sigma(S) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$ and $\sigma_p(S) = \emptyset$, it follows that $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\} \subseteq \sigma_c(S)$.

To sum up, we found that

$$\sigma_p(S) = \emptyset, \quad \sigma_c(S) = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}, \quad \text{and} \quad \sigma_r(S) = \{\lambda \in \mathbb{C} \mid |\lambda| < 1\}.$$

(c) Do the same for S^*, the left shift.

Solution: First, note that $\lambda \in \sigma(S^*)$ if and only if $\overline{\lambda} \in \sigma(S)$. Hence, we obtain from $\sigma(S) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$ that $\sigma(S^*) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$. Moreover, having already seen in the part (b) that $\sigma_p(S^*) \supseteq \{\lambda \in \mathbb{C} \mid |\lambda| < 1\}$ and that $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\} \cap \sigma_p(S^*) = \emptyset$, we obtain that $\sigma_p(S^*) = \{\lambda \in \mathbb{C} \mid |\lambda| < 1\}$. In addition, since for every

$\lambda \in \sigma_r(S^*)$ we would need to have $\overline{\lambda} \in \sigma_p(S)$, we see that $\sigma_r(S^*) = \emptyset$. Consequentially, $\sigma_c(S^*) = \sigma(S^*) \setminus (\sigma_p(S^*) \cup \sigma_r(S^*)) = \sigma(S^*) \setminus \sigma_p(S^*) = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$.

To sum up:

$$\sigma_p(S^*) = \{\lambda \in \mathbb{C} \mid |\lambda| < 1\}, \quad \sigma_c(S^*) = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}, \quad \text{and} \quad \sigma_r(S^*) = \emptyset.$$

12.2. Fredholm's alternative (on Hilbert spaces)

Let H be a Hilbert space and let $K \in L(H)$ be a compact operator. Prove the following statements. (The goal of this exercise lies in (d) and (e) below.)

(a) $\dim(\ker(I-K)) < \infty$.

Solution: Assume that $\dim(\ker(I-K)) = \infty$. Then there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \ker(I-K)$ with $\langle x_n, x_m\rangle = \delta_{nm}$ for all $n, m \in \mathbb{N}$. In particular, $(x_n)_{n\in\mathbb{N}}$ does not have a converging subsequence. By compactness of K and by $x_n = Kx_n$ for every $n \in \mathbb{N}$, the sequence $(x_n)_{n\in\mathbb{N}}$ should have a converging subsequence, though.

Alternatively, restricting K to the closed (and therefore complete) subspace $\ker(I-K)$, we are in the situation of a Hilbert/Banach space on which the identity operator is a compact operator or, put differently, in which the closed unit ball is compact. This only ever happens in finite dimensions.

(b) $\operatorname{im}(I-K)$ is closed.

Solution: We claim that there exists $\gamma \in (0,\infty)$ so that $\|x\| \le \gamma\|x - Kx\|$ for all $x \in (\ker(I-K))^\perp$. Indeed, if this was not the case, then there would exist a sequence $(x_n)_{n\in\mathbb{N}} \subseteq (\ker(I-K))^\perp$ satisfying $1 = \|x_n\| > n\|x_n - Kx_n\|$ for all $n \in \mathbb{N}$. This would imply that $x_n - Kx_n \to 0$ as $n \to \infty$. On the other hand, by compactness of K, we may assume (by passing to a subsequence, if necessary) that $Kx_n \to y$ as $n \to \infty$ for some $y \in H$. Consequentially, we would have that $x_n = (x_n - Kx_n) + Kx_n \to 0 + y = y$ as $n \to \infty$. Hence, we would obtain $y \in (\ker(I-K))^\perp$, $\|y\| = \lim_{n\to\infty}\|x_n\| = 1$, and $Ky = \lim_{n\to\infty} Kx_n = y$. But this is not possible as $y \in (\ker(I-K))^\perp$ and $Ky = y$ (i.e., $y \in \ker(I-K)$) would imply that $y = 0$, contradicting $\|y\| = 1$.

With $\gamma \in (0,\infty)$ so that $\|x\| \le \gamma\|x - Kx\|$ for all $x \in (\ker(I-K))^\perp$, we can now conclude that $\operatorname{im}(I-K)$ is closed: Let $(y_n)_{n\in\mathbb{N}} \subseteq \operatorname{im}(I-K)$ be an arbitrary sequence converging to y_∞ in H. Let $(x_n)_{n\in\mathbb{N}} \subseteq H$ satisfy for all $n \in \mathbb{N}$ that $y_n = x_n - Kx_n$. Denoting by $P \in L(H)$ the orthogonal projection onto the closed subspace $(\ker(I-K))^\perp$, we obtain that $(Px_n)_{n\in\mathbb{N}} \subseteq (\ker(I-K))^\perp$ (and therefore $x_n - Px_n \in \ker(I-K)$) so that $Px_n - KPx_n = x_n - Kx_n = y_n$ for every $n \in \mathbb{N}$.

Now, we can use the previously obtained inequality to verify that $(Px_n)_{n\in\mathbb{N}} \subseteq H$ is a Cauchy sequence:

$$\limsup_{N\to\infty} \sup_{m,n\geq N} \|Px_n - Px_m\| \leq \limsup_{N\to\infty} \sup_{m,n\geq N} \gamma\|y_n - y_m\| = 0.$$

Thus, there exists a limit $x_\infty \in H$ of $(Px_n)_{n\in\mathbb{N}}$ and $x_\infty - Kx_\infty = \lim_{n\to\infty}(I-K)Px_n = \lim_{n\to\infty} y_n = y_\infty$, i.e., $y_\infty \in \operatorname{im}(I-K)$.

(c) $\operatorname{im}(I-K) = (\ker(I-K^*))^\perp$.

Solution: This follows immediately from the fact that $\overline{\operatorname{im}(I-K)} = (\ker(I-K^*))^\perp$ and the fact that $\operatorname{im}(I-K)$ is closed (cp. part (b)).

(d) $\ker(I-K) = \{0\}$ if and only if $\operatorname{im}(I-K) = H$.

Solution: "($\Rightarrow$)": Assume for a contradiction that $\ker(I-K) = \{0\}$ and $\operatorname{im}(I-K) \neq H$. We first show by induction that $(I-K)^{k+1}(H) \subsetneq (I-K)^k(H)$ for every $k \in \mathbb{N}_0$. Indeed, for $k = 0$, this is just the previous assumption. And if $k \in \mathbb{N}$ is such that $(I-K)^k(H) \subsetneq (I-K)^{k-1}(H)$ but $(I-K)^{k+1}(H) = (I-K)^k(H)$, then we obtain that $x_0 \in (I-K)^{k-1}(H) \setminus (I-K)^k(H)$ gets mapped by $I-K$ to $(I-K)x_0 \in (I-K)^k(H) = (I-K)^{k+1}(H) = (I-K)((I-K)^k(H))$ so that there has to exist $x_1 \in (I-K)^k(H)$ satisfying $(I-K)x_0 = (I-K)x_1$. Hence, $0 \neq x_0 - x_1 \in \ker(I-K)$ (since $x_0 \neq x_1$ as $x_0 \notin (I-K)^k(H)$ while $x_1 \in (I-K)^k(H)$), which contradicts that $I-K$ is injective.

Knowing that – under the assumption that $\ker(I-K) = \{0\}$ and $\operatorname{im}(I-K) \neq H$ – it has to hold for every $k \in \mathbb{N}_0$ that $(I-K)^{k+1}(H) \subsetneq (I-K)^k(H)$ and since $(I-K)^k(H)$ is closed for every $k \in \mathbb{N}$ by part (b), we can now choose a sequence $(x_k)_{k\in\mathbb{N}} \subseteq H$ such that $\|x_k\| = 1$ and $x_k \in (I-K)^k(H) \cap ((I-K)^{k+1}(H))^\perp$ for every $k \in \mathbb{N}$. Moreover, note that for all $k, l \in \mathbb{N}$ with $k < l$ it holds that

$$x_k - (Kx_k - Kx_l) = \underbrace{(x_k - Kx_k)}_{\in (I-K)^{k+1}(H)} - \underbrace{(x_l - Kx_l)}_{\in (I-K)^{l+1}(H)} + \underbrace{x_l}_{\in (I-K)^l(H)} \in (I-K)^{k+1}(H),$$

i.e., $\|Kx_k - Kx_l\| \geq \operatorname{dist}(x_k, (I-K)^{k+1}(H)) = \|x_k\| = 1$ (since, sloppily speaking, $Kx_k - Kx_l$ has to cover at least the part of x_k perpendicular to $(I-K)^{k+1}(H)$). In particular, $(Kx_k)_{k\in\mathbb{N}}$ does not have a converging subsequence, although $(x_k)_{k\in\mathbb{N}} \subseteq H$ is a bounded sequence and K is compact.

"($\Leftarrow$)": $\operatorname{im}(I-K) = H$ implies that $\ker(I-K^*) = \{0\}$. By Schauder's theorem (cp. also Problem 11.2 (*Schauder's theorem*)) K^* is compact. The previous part of the proof hence implies that $\operatorname{im}(I-K^*) = H$. Hence, $\ker(I-K) = \{0\}$.

(e) $\dim(\ker(I-K)) = \dim(\ker(I-K^*))$.

Solution: Assume for a contradiction that $\dim(\ker(I-K)) < \dim(\ker(I-K^*))$. Since $\ker(I-K^*) = \operatorname{im}(I-K)^\perp$, we are assuming that $\dim(\ker(I-K)) < \dim(\operatorname{im}(I-K)^\perp)$. Since $\ker(I-K)$ is finite-dimensional by part (a) and $\dim(\ker(I-K)) < \dim(\operatorname{im}(I-K)^\perp)$, there exists an injective, but not surjective map $A_0\colon \ker(I-K) \to \operatorname{im}(I-K)^\perp$. Moreover, since $\ker(I-K)$ is finite-dimensional, A_0 has finite rank and is therefore compact. Define $A\colon H \to \operatorname{im}(I-K)^\perp$ via $A(x+y) = A_0x$ for $x \in \ker(I-K)$, $y \in (\ker(I-K))^\perp$. Since A is a compact linear map, $K+A$ is also a linear map (from H to H). Note that $(I-K-A)x = 0$ implies that $Ax = (I-K)x \in \operatorname{im}(I-K) \cap (\operatorname{im}(I-K))^\perp = \{0\}$, hence $x \in \ker(I-K) \cap \ker(A) = \ker(A_0) = \{0\}$. On the other hand, for every $x \in H$ it holds that $(I-K-A)x = (I-K)x - Ax \in \operatorname{im}(I-K) \oplus \operatorname{im}(A) \subsetneq \operatorname{im}(I-K) \oplus (\operatorname{im}(I-K))^\perp = H$ since $\operatorname{im}(A) \subsetneq (\operatorname{im}(I-K))^\perp$. Hence, we have $\ker(I-K-A) = \{0\}$ and $\operatorname{im}(I-K-A) \neq H$, contradicting part (d). This contradiction now shows $\dim(\ker(I-K)) \geq \dim(\ker(I-K^*))$. Since K^* is, by Schauder's theorem, compact as well, we obtain by the above argument that $\dim(\ker(I-K^*)) \geq \dim(\ker(I-K))$.

Remark. The statement remains true in the Banach space setting. (The proof gets slightly more technical.) In particular, we just saw – as mentioned earlier – that the extra symmetry assumption on the kernel k in Problem 11.5 (*Integral operators*) was not really necessary.

12.3. Symmetry vs. self-adjointness

Let H be a $\mathbb{C}$-Hilbert space and let $A\colon D_A \subseteq H \to H$ be a densely defined symmetric linear operator. Prove that the following statements are equivalent:

(i) A is self-adjoint.

(ii) A is closed and $\ker(A^*+i) = \{0\} = \ker(A^*-i)$.

(iii) $\operatorname{im}(A+i) = H = \operatorname{im}(A-i)$.

Solution: "$(i) \Rightarrow (ii)$": Since A^* is closed and $A = A^*$ by assumption, A is closed. Moreover, for every $x \in D_A$ it holds that (since $A = A^*$)

$$\langle Ax, x\rangle = \langle x, A^*x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle},$$

i.e., $\langle Ax, x\rangle \in \mathbb{R}$. On the other hand, it holds for every $x \in \ker(A^*+i)$ that

$$i\|x\|^2 = \langle x, -ix\rangle = \langle x, A^*x\rangle = \langle Ax, x\rangle \in \mathbb{R},$$

which results in $x = 0$. Similarly do we obtain for every $x \in \ker(A^* - i)$ that $-i\|x\|^2 = \langle Ax, x\rangle \in \mathbb{R}$, which again implies that $x = 0$.

"$(ii) \Rightarrow (iii)$": We show that $\operatorname{im}(A+i)$ is closed and dense. Let $(y_n)_{n\in\mathbb{N}} \subseteq \operatorname{im}(A+i)$ be a sequence converging to $y_\infty \in H$ as $n \to \infty$ and let $(x_n)_{n\in\mathbb{N}} \subseteq D_A$ satisfy $y_n = (A+i)x_n$ for every $n \in \mathbb{N}$. Then it holds for every $n \in \mathbb{N}$ that

$$\begin{aligned}\|y_n\|\|x_n\| &\geq |\langle y_n, x_n\rangle| = |\langle Ax_n + ix_n, x_n\rangle| = |\langle Ax_n, x_n\rangle + i\|x_n\|^2| \\ &= \sqrt{\langle Ax_n, x_n\rangle^2 + \|x_n\|^4} \geq \|x_n\|^2.\end{aligned}$$

It follows that $(x_n)_{n\in\mathbb{N}} \subseteq H$ is a Cauchy sequence and therefore converges to some limit $x_\infty \in H$. Since $y_n = Ax_n + ix_n \to y_\infty$ as $n \to \infty$, it follows that $Ax_n = y_n - ix_n \to y_\infty - ix_\infty$ as $n \to \infty$. The assumption that A is closed now implies that $x_\infty \in D_A$ and $y_\infty - ix_\infty = Ax_\infty$. Hence, $y_\infty = (A+i)x_\infty \in \operatorname{im}(A+i)$ and $\operatorname{im}(A+i)$ is closed. In an analogous way it can be shown that $\operatorname{im}(A-i)$ is closed and dense.

"$(iii) \Rightarrow (i)$": Since A is symmetric, we know that $A \subseteq A^*$. It thus remains to show that $A^* \subseteq A$, i.e., that $D_{A^*} \subseteq D_A$ (and, of course, $A^*x = Ax$ for every $x \in D_{A^*}$, but this is then clear). For this, let $x \in D_{A^*}$. Since $A + i$ is assumed to be surjective, there exists $z \in D_A$ such that $A^*x + ix = Az + iz$. Then it holds for all $y \in D_A$ that

$$\langle x, Ay - iy\rangle = \langle A^*x + ix, y\rangle = \langle Az + iz, y\rangle = \langle z, Ay - iy\rangle.$$

Moreover, since $\operatorname{im}(A - i) = H$, this implies that $x = z \in D_A$.

12.4. Special construction of self-adjoint operators

Let H and K be $\mathbb{K}$-Hilbert spaces (with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) and let $J \in L(K, H)$ be an injective operator with dense range.

(a) Prove that $JJ^* \in L(H)$ is an injective operator with dense range.

Solution: For all $x \in \ker(JJ^*)$ it holds that $\|J^*x\|_K^2 = \langle JJ^*x, x\rangle_H = 0$. Hence, $\ker(JJ^*) \subseteq \ker(J^*) = (\operatorname{im}(J))^\perp = \{0\}$ since J is assumed to have dense range. Moreover, for every $x \in \operatorname{im}(JJ^*)^\perp$, it holds that $0 = \langle JJ^*x, x\rangle_H = \|J^*x\|_K^2$. That is, $(\operatorname{im}(JJ^*))^\perp \subseteq \ker(J^*) = \{0\}$. Thus, $\operatorname{im}(JJ^*)$ lies dense in H.

(b) Prove that $S := (JJ^*)^{-1}$ (i.e., the operator $S\colon D_S \subseteq H \to H$, defined by $D_S = \operatorname{im}(JJ^*)$ and $S(JJ^*x) = x$ for all $x \in H$) is self-adjoint.

Solution: First, we consider the case of $\mathbb{K} = \mathbb{C}$. We show that S is a densely defined symmetric operator satisfying $\operatorname{im}(S+i) = H = \operatorname{im}(S-i)$ (where we show density and closedness of $\operatorname{im}(S+i)$ and $\operatorname{im}(S-i)$ for the latter) and invoke Problem 12.3 (*Symmetry vs. self-adjointness*). For the symmetry of S, let $x_1, x_2 \in D_S$ be arbitrary. Necessarily, there exist $w_1, w_2 \in H$ such that $x_1 = JJ^*w_1$ and $x_2 = JJ^*w_2$. Self-adjointness of $JJ^* \in L(H)$ ensures that

$$\langle Sx_1, x_2\rangle_H = \langle w_1, JJ^*w_2\rangle_H = \langle JJ^*w_1, w_2\rangle_H = \langle x_1, Sx_2\rangle_H.$$

For the density of $\operatorname{im}(S+i)$, consider $x \in \operatorname{im}(S+i)^\perp$. Since $JJ^*x \in D_S$, it thus holds that

$$0 = \langle (S+i)JJ^*x, x\rangle_H = \|x\|_H^2 + i\|J^*x\|_K^2,$$

showing that $x = 0$. Hence, $\overline{\operatorname{im}(S+i)} = H$. Analogously, one can show that $\overline{\operatorname{im}(S-i)} = H$.

For the closedness of $\operatorname{im}(S+i)$, consider a sequence $(y_n)_{n\in\mathbb{N}} \subseteq \operatorname{im}(S+i)$ with limit y_∞ and let $(x_n)_{n\in\mathbb{N}} \subseteq D_S$ be given by $y_n = Sx_n + ix_n$ for every $n \in \mathbb{N}$. Since it holds for all $u \in D_S$ that

$$\|(S+i)u\|_H\|u\|_H \geq |\langle Su + iu, u\rangle_H| = \Big|i\|J^*Su\|_K^2 + \|u\|_H^2\Big| \geq \|u\|_H^2,$$

we infer that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in H. Therefore, $(x_n)_{n\in\mathbb{N}}$ converges to some $x_\infty \in H$ and $Sx_n = y_n - ix_n \to y_\infty - ix_\infty$ as $n \to \infty$. Since JJ^* is continuous (and therefore closed), S is closed, and since S is closed, we conclude that $x_\infty \in D_S$ with $Sx_\infty + ix_\infty = y_\infty$. Thus, $\operatorname{im}(S+i)$ is closed. It can be proved analogously that $\operatorname{im}(S-i)$ is closed.

Problem 12.3 (*Symmetry vs. self-adjointness*) now ensures that S – as a symmetric, densely defined (by part (a), we know that $D_S = \operatorname{im}(JJ^*)$ is dense in H) operator with $\operatorname{im}(S+i) = H = \operatorname{im}(S-i)$ – is self-adjoint. Thus, the claim is proved in the case that $\mathbb{K} = \mathbb{C}$.

Next we consider the case that $\mathbb{K} = \mathbb{R}$. Let $\mathcal{H} := H^2$ and $\mathcal{K} := K^2$. By defining the vector operations as well as the scalar product on $\mathcal{H}$ (and analogously on $\mathcal{K}$) via

$$\begin{aligned}
(g_1, g_2) +_{\mathcal{H}} (h_1, h_2) &= (g_1 + h_1, g_2 + h_2),\\
(a_1 + ia_2) \cdot_{\mathcal{H}} (g_1, g_2) &= (a_1g_1 - a_2g_2, a_1g_2 + a_2g_1),\\
\langle (g_1, g_2), (h_1, h_2)\rangle_{\mathcal{H}} &= \langle g_1, h_1\rangle_H + \langle g_2, h_2\rangle_H + i(\langle g_2, h_1\rangle_H - \langle g_1, h_2\rangle_H),
\end{aligned}$$

for all $a_1, a_2 \in \mathbb{R}$, $g_1, g_2, h_1, h_2 \in H$, we equip $\mathcal{H}$ (and, analogously, $\mathcal{K}$) with a $\mathbb{C}$-Hilbert space structure (why?). Moreover, H is isometrically embedded in $\mathcal{H}$ via $(H \ni h \mapsto (h, 0) \in \mathcal{H})$ (and analogously K is isometrically embedded in $\mathcal{K}$).

Note that $\mathcal{J} \colon \mathcal{K} \to \mathcal{H}$, given by $\mathcal{J}(x, y) = (Jx, Jy)$ for all $x, y \in K$, is a bounded ($\mathbb{C}$-)linear map from $\mathcal{K}$ to $\mathcal{H}$. Moreover, $\operatorname{im}(\mathcal{J}) = \operatorname{im}(J) \times \operatorname{im}(J)$ is dense in $\mathcal{H}$ and

$\ker(\mathcal{J}) = \ker(J) \times \ker(J) = \{(0_K, 0_K)\} = \{0_{\mathcal{K}}\}$, i.e. $\mathcal{J}$ is injective and has dense range. In addition, it holds for all $x_1, x_2 \in K$, $y_1, y_2 \in H$ that

$$\begin{aligned}\langle \mathcal{J}(x_1,x_2),(y_1,y_2)\rangle_{\mathcal{H}} &= \langle (Jx_1,Jx_2),(y_1,y_2)\rangle_{\mathcal{H}}\\ &= \langle Jx_1,y_1\rangle_H + \langle Jx_2,y_2\rangle_H + i(\langle Jx_2,y_1\rangle_H - \langle Jx_1,y_2\rangle_H)\\ &= \langle x_1,J^*y_1\rangle_K + \langle x_2,J^*y_2\rangle_K + i(\langle x_2,J^*y_1\rangle_K - \langle x_1,J^*y_2\rangle_K)\\ &= \langle (x_1,x_2),(J^*y_1,J^*y_2)\rangle_{\mathcal{K}},\end{aligned}$$

i.e., $J^*(y_1,y_2) = (J^*y_1, J^*y_2)$ for all $y_1, y_2 \in H$. We know from our considerations of the case of $\mathbb{C}$-Hilbert spaces that $\mathcal{S} := (\mathcal{J}\mathcal{J}^*)^{-1}$ is self-adjoint. Since it holds for all $x, y \in H$ that $\mathcal{J}\mathcal{J}^*(x,y) = \mathcal{J}(J^*x, J^*y) = (JJ^*x, JJ^*y)$, we obtain that $D_{\mathcal{S}} = \operatorname{im}(\mathcal{J}\mathcal{J}^*) = \operatorname{im}(JJ^*) \times \operatorname{im}(JJ^*) = D_S \times D_S$ and that $\mathcal{S}(x,y) = (Sx, Sy)$ for all $x, y \in D_S$. This implies in particular that $D_{S^*} \times D_{S^*} \subseteq D_{\mathcal{S}^*}$. On the other hand, for every $(x_1, x_2) \in D_{\mathcal{S}^*}$, there exists $C \in [0,\infty)$ satisfying

$$|\langle \mathcal{S}(y_1,y_2),(x_1,x_2)\rangle_{\mathcal{H}}| \le C\|(y_1,y_2)\|_{\mathcal{H}} \quad \text{for all } y_1, y_2 \in D_S,$$

which implies for all $z \in D_S$ that

$$\begin{aligned}|\langle Sz, x_1\rangle_H| &= |\langle S(z,0),(x_1,x_2)\rangle_{\mathcal{H}}| \le C\|(z,0)\|_{\mathcal{H}} = C\|z\|_H,\\ |\langle Sz, x_2\rangle_H| &= |\langle S(0,z),(x_1,x_2)\rangle_{\mathcal{H}}| \le C\|(0,z)\|_{\mathcal{H}} = C\|z\|_H,\end{aligned}$$

which, in turn, results in $x_1, x_2 \in D_{S^*}$. Thus, we have $D_{S^*} \times D_{S^*} = D_{\mathcal{S}^*} = D_{\mathcal{S}} = D_S \times D_S$, i.e., $D_S = D_{S^*}$. Since it holds for all $x, y \in D_S$ that $\langle Sx, y\rangle_H = \langle Sx, JJ^*Sy\rangle_H = \langle JJ^*Sx, Sy\rangle_H = \langle x, Sy\rangle_H$, we have finally arrived at $S^* = S$.

12.5. Heisenberg's Uncertainty Principle

Let $(H, \langle\cdot,\cdot\rangle_H)$ be a Hilbert space over $\mathbb{C}$. Let $D_A, D_B \subseteq H$ be dense subspaces and let $A\colon D_A \subseteq H \to H$ and $B\colon D_B \subseteq H \to H$ be symmetric linear operators. Assume that

$$A(D_A \cap D_B) \subseteq D_B \quad \text{and} \quad B(D_A \cap D_B) \subseteq D_A,$$

and define the *commutator* of A and B as

$$[A,B]\colon D_{[A,B]} \subseteq H \to H, \qquad [A,B](x) := A(Bx) - B(Ax),$$

where $D_{[A,B]} := D_A \cap D_B$.

(a) Prove that

$$\Big|\langle x, [A,B]x\rangle_H\Big| \le 2\|Ax\|_H\|Bx\|_H \quad \text{for every } x \in D_{[A,B]}.$$

Solution: Let $x \in D_{[A,B]} := D_A \cap D_B$. Then, applying the Cauchy–Schwarz inequality,

$$\begin{aligned}\Big|\langle x, [A,B]x\rangle_H\Big| &\le \Big|\langle x, A(Bx)\rangle_H\Big| + \Big|\langle x, B(Ax)\rangle_H\Big| \\ &= \Big|\langle Ax, Bx\rangle_H\Big| + \Big|\langle Bx, Ax\rangle_H\Big| \\ &\le \|Ax\|_H\|Bx\|_H + \|Bx\|_H\|Ax\|_H \\ &= 2\|Ax\|_H\|Bx\|_H.\end{aligned}$$

(b) Define now the *standard deviation* of A

$$\varsigma(A,x) := \sqrt{\langle Ax, Ax\rangle_H - \langle x, Ax\rangle_H^2}$$

at each $x \in D_A$ with $\|x\|_H = 1$. Verify that $\varsigma(A,x)$ is well-defined for every x (i.e. that the radicand is real and non-negative) and prove that for every $x \in D_{[A,B]}$ with $\|x\|_H = 1$ there holds

$$\Big|\langle x, [A,B]x\rangle_H\Big| \le 2\varsigma(A,x)\,\varsigma(B,x).$$

Solution: Since A is a symmetric operator, $\langle x, Ax\rangle_H$ is real for every $x \in D_A \subseteq D_{A^*}$. Indeed,

$$\langle x, Ax\rangle_H = \langle A^*x, x\rangle_H = \langle Ax, x\rangle_H = \overline{\langle x, Ax\rangle}_H.$$

Moreover, for $x \in D_A$ with $\|x\|_H = 1$, we have

$$\langle x, Ax\rangle_H^2 \le \|x\|_H^2\|Ax\|_H^2 = \langle Ax, Ax\rangle_H.$$

Therefore, the radicand in the definition of the standard deviation is a non-negative real number and $\varsigma(A,x)$ is well-defined. For any $\lambda, \mu \in \mathbb{R}$, the commutators $[A,B]$ and $[A-\lambda, B-\mu]$ agree:

$$\begin{aligned}[A-\lambda, B-\mu] &= (A-\lambda)(B-\mu) - (B-\mu)(A-\lambda) \\ &= AB - \mu A - \lambda B + \lambda\mu - BA + \lambda B + \mu A - \lambda\mu = [A,B]\end{aligned}$$

on $D_{[A-\lambda,B-\mu]} = D_{A-\lambda} \cap D_{B-\mu} = D_A \cap D_B = D_{[A,B]}$. Since A is symmetric and $\lambda \in \mathbb{R}$, the operator $\tilde{A} = A - \lambda$ is also symmetric on $D_{\tilde{A}} = D_A$. Moreover, for any $x \in D_A$,

$$\begin{aligned}\|\tilde{A}x\|_H^2 = \langle \tilde{A}x, \tilde{A}x\rangle_H &= \langle Ax - \lambda x, Ax - \lambda x\rangle_H \\ &= \langle Ax, Ax\rangle_H - \lambda\langle x, Ax\rangle_H - \lambda\langle Ax, x\rangle_H + \lambda^2\langle x, x\rangle_H \\ &= \langle Ax, Ax\rangle_H - 2\lambda\langle x, Ax\rangle_H + \lambda^2\langle x, x\rangle_H.\end{aligned}$$

We observe that if we choose $\lambda = \langle x, Ax\rangle_H \in \mathbb{R}$ and if $\|x\|_H = 1$, then

$$\|\tilde{A}x\|_H^2 = \langle Ax, Ax\rangle_H - \langle x, Ax\rangle_H^2 = \varsigma(A,x)^2.$$

Now, let $x \in D_{[A,B]} = D_A \cap D_B$ with $\|x\|_H = 1$ be arbitrary. Since the operators $\tilde{A} := A - \langle x, Ax\rangle_H$ and $\tilde{B} := B - \langle x, Bx\rangle_H$ are symmetric, part (a) applies and yields

$$\Big|\langle x, [A,B]x\rangle_H\Big| = \Big|\langle x, [\tilde{A},\tilde{B}]x\rangle_H\Big| \leq 2\|\tilde{A}x\|_H\|\tilde{B}x\|_H = 2\varsigma(A,x)\,\varsigma(B,x).$$

Remark. The possible *states* of a quantum mechanical system are given by elements $x \in H$ with $\|x\|_H = 1$. Each *observable* is given by a symmetric linear operator $A\colon D_A \subseteq H \to H$. If the system is in state $x \in D_A$, we measure the observable A with uncertainty $\varsigma(A,x)$.

(c) Let $A\colon D_A \subseteq H \to H$ and $B\colon D_B \subseteq H \to H$ be as above. A, B is called *Heisenberg pair* if

$$[A,B] = i\,\mathrm{Id}\,|_{D_{[A,B]}}.$$

Show that, if A, B is a Heisenberg pair with B continuous (and $D_B = H$), then A cannot be continuous.

Solution: Suppose, $B \in L(H)$ and $A\colon D_A \subseteq H \to H$ satisfy

$$[A,B] = i\,\mathrm{Id}\,|_{D_{[A,B]}}.$$

By assumption, $D_{[A,B]} = D_A \cap H = D_A$ and $B(D_A) \subseteq D_A$. In particular, for any $n \in \mathbb{N}$ the inclusion $B^n(D_A) \subseteq D_A$ is satisfied, which is necessary to define $[A, B^n]$. We prove $[A, B^n]x = niB^{n-1}x$ for every $n \in \mathbb{N}$, $x \in D_A$ by induction. For $n = 1$, the claim holds by assumption. Suppose, it is true for some $n \in \mathbb{N}$. Then it holds for every $x \in D_A$ that

$$\begin{aligned}
[A, B^{n+1}]x &= AB^{n+1}x - B^{n+1}Ax \\
&= \Big(AB^n - B^nA + B^nA\Big)Bx - B^{n+1}Ax \\
&= \Big([A,B^n] + B^nA\Big)Bx - B^{n+1}Ax \\
&= niB^{n-1}Bx + B^nABx - B^{n+1}Ax \\
&= niB^nx + B^n[A,B]x = niB^nx + iB^nx = (n+1)iB^n.
\end{aligned}$$

A consequence is that B cannot be nilpotent: If $B^n = 0$ for some $n \in \mathbb{N}$, then $B^{n-1}x = \frac{1}{ni}[A, B^n]x = 0$ for all $x \in D_A$, i.e., $B^{n-1} = 0$, which iterates to $B = 0$ in contradiction to $[A,B] \neq 0|_{D_A}$. Suppose by contradiction that A has finite operator norm $\|A\|$. Then, we can assume w.l.o.g. that $D_A = H$ and

$$n\|B^{n-1}\| = \|[A,B^n]\| \leq \|AB^n\| + \|B^nA\| \leq 2\|A\|\|B^{n-1}\|\|B\|.$$

Since $\|B^{n-1}\| \neq 0$, we obtain $2\|A\| \geq \frac{n}{\|B\|} > 0$ for every $n \in \mathbb{N}$, thus $\|A\|$ cannot be finite and the contradiction is reached.

(d) Consider the Hilbert space $(H, \langle\cdot,\cdot\rangle_H) = \big(L^2([0,1],\mathbb{C}), \langle\cdot,\cdot\rangle_{L^2}\big)$ and the subspace

$$C_0^1([0,1],\mathbb{C}) := \{f \in C^1([0,1],\mathbb{C}) \mid f(0) = 0 = f(1)\}.$$

Recall that $C_0^1([0,1],\mathbb{C}) \subseteq L^2([0,1],\mathbb{C})$ is a dense subspace. The operators

$$\begin{aligned} P\colon C_0^1([0,1],\mathbb{C}) &\to L^2([0,1],\mathbb{C}), & Q\colon L^2([0,1],\mathbb{C}) &\to L^2([0,1],\mathbb{C}) \\ f(s) &\mapsto if'(s) & f(s) &\mapsto sf(s) \end{aligned}$$

correspond to the observables *momentum* and *position*. Check that P and Q are well-defined, symmetric operators. Check that $[P,Q]\colon C_0^1([0,1],\mathbb{C}) \to L^2([0,1],\mathbb{C})$ is well-defined.

Show that P and Q form a Heisenberg pair and conclude that the *uncertainty principle* holds: for every $f \in C_0^1([0,1],\mathbb{C})$ with $\|f\|_{L^2([0,1],\mathbb{C})} = 1$ there holds

$$\varsigma(P,f)\,\varsigma(Q,f) \geq \tfrac{1}{2}.$$

Thus we conclude: *The more precisely the momentum of some particle is known, the less precisely its position can be known, and vice versa.*

Solution: If $f \in C^1([0,1],\mathbb{C})$, then f' is bounded and in particular $f' \in L^2([0,1],\mathbb{C})$. Therefore, the linear operators

$$\begin{aligned} P\colon C_0^1([0,1],\mathbb{C}) &\to L^2([0,1],\mathbb{C}), & Q\colon L^2([0,1],\mathbb{C}) &\to L^2([0,1],\mathbb{C}) \\ f(s) &\mapsto if'(s) & f(s) &\mapsto sf(s) \end{aligned}$$

are indeed well-defined. They are also symmetric. For Q this follows immediately from $[0,1] \subseteq \mathbb{R}$. Indeed, for all $f,g \in D_Q = L^2([0,1],\mathbb{C})$ it holds that

$$\langle Qf, g\rangle_{L^2} = \int_0^1 sf(s)\overline{g(s)}\,ds = \int_0^1 f(s)\overline{sg(s)}\,ds = \langle f, Qg\rangle_{L^2}.$$

For P, symmetry follows via integration by parts. Indeed, given any $f,g \in D_P = C_0^1([0,1],\mathbb{C})$, we have

$$\langle Pf, g\rangle_{L^2} = \int_0^1 if'(s)\overline{g}(s)\,ds = -\int_0^1 if(s)\overline{g}'(s)\,ds = \int_0^1 f(s)\overline{ig'(s)}\,ds = \langle f, Pg\rangle_{L^2}.$$

When integrating by parts, the boundary terms vanish due to $f(0) = 0 = f(1)$. Hence, $P\colon C^1_0([0,1];\mathbb{C}) \to L^2([0,1];\mathbb{C})$ is symmetric (but *not* self-adjoint! see Beispiel 6.6.1).

Next, we verify that the commutator $[P,Q]$ is well-defined. Since $D_Q = L^2([0,1],\mathbb{C})$ is the whole space, the only thing to check is that $Qf\colon s \mapsto sf(s)$ is in $D_P = C^1_0([0,1],\mathbb{C})$ whenever $f \in C^1_0([0,1],\mathbb{C})$. But this follows from the product rule. Moreover,

$$([P,Q]f)(s) = (P(Qf))(s) - (Q(Pf))(s) = if(s) + isf'(s) - sif'(s) = if(s)$$

for almost every $s \in [0,1]$ which proves that P, Q is a Heisenberg pair. By part (b),

$$\forall f \in C^1_0,\ \|f\|_{L^2} = 1: \quad \varsigma(P,f)\,\varsigma(Q,f) \geq \tfrac{1}{2}\Big|\langle f, [P,Q]f\rangle_{L^2}\Big| = \tfrac{1}{2}\Big|\langle f, if\rangle_{L^2}\Big| = \tfrac{1}{2}.$$