Appendix A

Algebraic and Functional Analytic Results

Algebraic symbols are used when you do not know what you are talking about.

— Philippe Schnoebelen

In this appendix, we present a number of algebraic and functional analytic results that are needed in the main parts of this monograph.

In Section A.1, we mainly present algebraic results, such as "$(I-AB)^{-1} = I + A(I-BA)^{-1}B$" or "$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$", that are valid for matrices and more general linear operators or elements of certain rings.

In Section A.2, we very briefly introduce metric spaces and other topological spaces. In Section A.3, we list standard and extended concepts and facts about Hilbert and Banach spaces and Banach algebras. In Section A.4, we present strongly continuous (C_0) semigroups.

In the main part of the monograph, all vector spaces (e.g., Banach spaces) are assumed to be complex. However, in the appendices we generally assume that the scalar field is **K**, which the reader may read as either **C** or **R**. In Appendix D and Sections A.4 and F.3, we assume that $\mathbf{K} = \mathbf{C}$, as explicitly stated there; in the other sections in appendices we always state explicitly any such exceptions.

Thus, the concepts "vector space", "Banach space" and "Hilbert space" mean spaces of the corresponding type over the scalar field **K** (in particular, if some spaces in a theorem are assumed to be Banach spaces, all of them must be Banach spaces over the same **K** ($=$ **R** or **C**)).

A.1 Algebraic auxiliary results ($\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$)

If $A \in \mathcal{B}(X_1 \times X_2, Y_1 \times Y_2)$, where X_1, X_2, Y_1, Y_2 are Banach spaces, then A can be written as $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, where $A_{ij} \in \mathcal{B}(X_j, Y_i)$, as shown in Lemma A.1.1(a4) below. There are many simple, well-known algebraic rules how to handle this kind of *operator matrices*; e.g., if $A_{12} = 0$ and the diagonal blocks (operators) A_{11} and A_{22} are invertible, then so is A (see (b1)).

In Lemma A.1.1 we present a number of such results in a more general setting (we only need the ring operations and units); see the remarks following the lemma for generalizations.

Recall that "&" means "and", and that Banach spaces are topological vector spaces. Moreover, $X \times Y := \{(x,y) \,|\, x \in X,\ y \in Y\}$. If X and Y are normed spaces, we use the norm $\|(x,y)\|_{X\times Y} := (\|x\|_X^2 + \|y\|_Y^2)^{1/2}$; if X and Y are inner product spaces, we use the inner product $\langle (x,y),(x',y')\rangle_{X\times Y} := \langle x,x'\rangle_X + \langle y,y'\rangle_Y$; use induction for $\prod_{k=1}^n X_k := X_1 \times X_2 \times \cdots \times X_n$. Finally, $X^n := \prod_{k=1}^n X$.

Lemma A.1.1 (Operator matrix lemma) *We assume that (1.), (2.), (3.), (4.) or (5.) holds, where*

(1.) $\mathcal{X}$ is the collection of all vector spaces, and $\mathcal{A}(X,Y) = \mathrm{Hom}(X,Y)$ is the set of vector homomorphisms (i.e., linear mappings) $X \to Y$.

(2.) $\mathcal{X}$ is the collection of all topological vector spaces, and $\mathcal{A}(X,Y) = \mathcal{B}(X,Y)$ is the set of continuous linear mappings $X \to Y$.

(3.) $\mathcal{X}$ is the collection of all Banach spaces, and $\mathcal{A}$ is any of the symbols Hom, $\mathcal{B}$, $\mathcal{C}(\Omega;\mathcal{B}(\cdot,\cdot))$, $\mathrm{H}(\Omega;\mathcal{B}(\cdot,\cdot))$, $\mathrm{H}^\infty(\Omega;\mathcal{B}(\cdot,\cdot))$, $[\mathcal{B}+]\mathrm{H}^p_\infty$, $[\mathcal{B}+]\mathrm{H}^p_{\mathrm{strong},\infty}$, TI_ω, TIC_ω, $\mathrm{L}^\infty(Q,\mathcal{B}(\cdot,\cdot))$, *and* $\mathrm{L}^\infty_{\mathrm{strong}}(Q,\mathcal{B}(\cdot,\cdot))$, *where $\Omega \subset \mathbf{C}$ is open, Q and μ are as on p. 907, and $\omega \in \mathbf{R} \cup \{\infty\}$.*

(For $*\mathrm{TI}*$ and $\mathrm{H}*$ we require the elements of $\mathcal{X}$ to be complex Banach spaces; this applies (4.) and (5.) too.)

(4.) $\mathcal{X}$ is the collection of all complex Banach spaces, and $\mathcal{A}$ is any of the symbols defined in Definitions 2.6.1 and 2.6.3, except that if $\mathcal{A}$ is a symbol with a specified atomgroup $\mathbf{S}$, *we require that* $\mathbf{S} = \mathbf{S} - \mathbf{S} \subset \mathbf{R}$*; cf. Definition 2.6.3 and Theorem 2.6.4.*

(5.) $\mathcal{X}$ is the collection of all Banach spaces, and $\mathcal{A}$ is any of the symbols SR, SLR, SHPR, SVR, UR, ULR, UHPR *and* UVR, *(where these symbols are as in Definition 6.2.3) or $\mathcal{A}$ is the intersection* TIC_ω *and any of the above symbols for some $\omega \in \mathbf{R} \cup \{\infty\}$.*

We use the following notation: By $\mathcal{A}$ we mean any of the sets $\mathcal{A}(X,Y)$ ($X,Y \in \mathcal{X}$); the group operation and identity operator in (any) $\mathcal{A}$ are denoted by $+$ and $I = I_{\mathcal{A}}$, respectively. If $A \in \mathcal{A}(X,Y)$ and $B \in \mathcal{A}(Y,X)$ are s.t. $AB = I_{\mathcal{A}(Y)}$, *then we write* $A^{-1}_{\text{right}} = B$ *and* $B^{-1}_{\text{left}} = A$*; if, in addition,* $BA = I_{\mathcal{A}(X)}$*, then we write* $A^{-1} = B$*. By "*$\exists A^{-1}$*" we mean that* A *has an inverse (as above). Similarly, "*$\exists A^{-1}_{\text{right}}$*" (resp. "*$\exists A^{-1}_{\text{left}}$*") means that* A *has a right (resp. left) inverse.*

We assume that $X_k, Y_k, Z_k \in \mathcal{X}$ *for all* $k \in \mathbf{N}$*.*

With the above assumptions and notation, the following holds:

(a1) If $A \in \mathcal{A}(X_1, Y_1)$ *has a left inverse* B *and a right inverse* C*, then* $B = BAC = C$ *is the unique inverse of* A*.*

(a2) If $\exists A^{-1}, B^{-1}$*, then* $\exists (AB)^{-1} = B^{-1}A^{-1}$*, when* $A \in \mathcal{A}(X_1, Y_1)$ *and* $B \in \mathcal{A}(Y_1, Z_1)$*.*

(a3) If $\dim X_1 = \dim Y_1 < \infty$*, then any* $A \in \mathcal{A}(X_1, Y_1)$ *is left (resp. right) invertible iff it is invertible.*

(a4) Let $n, m, N \in \{1, 2, 3, \ldots\}$*, and let* $X := X_1 \times \cdots \times X_n$*,* $Y := Y_1 \times \cdots \times Y_m$ $Z := Z_1 \times \cdots \times Z_N$*. Let* $P_i : Y \to Y_i$ *be the canonical projection and* $L_j : X_j \to X$ *the canonical imbedding. Then* $P_i \in \mathcal{A}(Y, Y_i)$ *and* $L_j \in \mathcal{A}(X_j, X)$*.*

Let, in addition, $A \in \mathcal{A}(X, Y)$*. Then* $A_{ij} := P_i A L_j \in \mathcal{A}(X_j, Y_i)$ *for all i, j, and the representation*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nm} \end{bmatrix}, \tag{A.1}$$

satisfies the standard matrix multiplication rules $(A(x_1, \ldots, x_n))_i = \sum_{j=1}^n A_{ij} x_j \in Y_i$ *for* $(x_1, \ldots, x_n) \in X_1 \times \ldots \times X_n$*,* $i = 1, \ldots, n$*, and* $(AB)_{ik} = \sum_k A_{ij} B_{jk}$ *for* $B \in \mathcal{A}(Y, Z)$*,* $i = 1, \ldots, n$*,* $j = 1, \ldots, m$*,* $k = 1, \ldots, N$*. Moreover,* $(A + A')_{ij} := A_{ij} + A'_{ij}$ *for* $A' \in \mathcal{A}(X, Y)$*, and* $(-A)_{ij} := -A_{ij}$*.*

Conversely, if(f) $A_{ij} \in \mathcal{A}(X_j, Y_i)$ *for all i, j, then we can define* $A \in \mathcal{A}(X, Y)$ *by setting* $A := \sum_{i,j} P_i^* A_{ij} L_j^*$*; equivalently,* $(A(x_1, \ldots, x_n))_i := \sum_{j=1}^n A_{ij} x_j \in Y_i$*. We denote this by (A.1).*

If X_j *and* Y_i *are Hilbert spaces for all* i, j *and (A.1) holds, then* $(A^*)_{ij} = A^*_{ji}$ *for all* i, j*.*

In parts (b1)–(h1) we assume that $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathcal{A}(X_1 \times X_2, Y_1 \times Y_2)$ *and* $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \in \mathcal{A}(Y_1 \times Y_2, X_1 \times X_2)$*, and that also the other terms (operators) belong to the* $\mathcal{A}$*'s that are compatible with the formulae.*

(b1) $\exists A_{11}^{-1}, A_{22}^{-1} \implies \exists \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$ & $\exists \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1} A_{21} A_{11}^{-1} & A_{22}^{-1} \end{bmatrix}$.

If $\exists (A_{11})^{-1}_{\text{right}}, (A_{22})^{-1}_{\text{right}}$ *(resp.* $\exists (A_{11})^{-1}_{\text{left}}, (A_{22})^{-1}_{\text{left}}$*), then the above inverse matrices exist as right (resp. left) inverses, and the above formulae hold (*mutatis mutandis*).*

(b2) Conversely, if $\exists \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} =: \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$*, then* $\exists A_{11}^{-1}$, A_{22}^{-1} *iff any of the conditions (1)–(7) holds, where:*

(1) $\exists (A_{11})^{-1}_{\text{right}}$*; (2)* $\exists (A_{22})^{-1}_{\text{left}}$*; (3)* $B_{21} = 0$*; (4)* $A_{12} = 0$*; (5)* $\dim X_1 = \dim Y_1 < \infty$*; (6)* $\dim X_2 = \dim Y_2 < \infty$*; (7)* X_1, X_2, Y_1, Y_2 *are Banach spaces,* $\mathcal{A} = \mathcal{B}$*, and* $A_{11} \in \mathcal{GB} + \mathcal{BC}$ *or* $A_{22} \in \mathcal{GB} + \mathcal{BC}$.

(c1) Let $\exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} =: \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$. *Then* $\exists B_{11}^{-1} \Leftrightarrow \exists A_{22}^{-1} \Leftrightarrow \exists A_{22}^{-1} = B_{22} - B_{21}B_{11}^{-1}B_{12} \Leftrightarrow \exists B_{11}^{-1} = A_{11} - A_{12}A_{22}^{-1}A_{21}$.

If $\dim X_1 < \infty$ *or* $\dim X_2 < \infty$*, then, in addition,* $\exists (A_{22})^{-1}_{\text{left}} \Leftrightarrow \exists (A_{22})^{-1}_{\text{right}} \Leftrightarrow \exists A_{22}^{-1} \Leftrightarrow \exists (B_{11})^{-1}_{\text{left}} \Leftrightarrow \exists (B_{11})^{-1}_{\text{right}} \Leftrightarrow \exists B_{11}^{-1}$.

If $\exists B_{11}^{-1}$ *and* $A^{-1} = B$*, then the formulae of (d1) hold,* $B_{21}B_{11}^{-1} = -A_{22}^{-1}A_{21}$*, and*

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} I & B_{12} \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ A_{21} & I \end{bmatrix}^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ B_{21} & B_{22} \end{bmatrix}. \tag{A.2}$$

(c2) If $\exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}_{\text{right}} =: \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$*, then* $\exists (B_{11})^{-1}_{\text{right}} \Longrightarrow \exists (A_{22})^{-1}_{\text{right}} = B_{22} - B_{21}(B_{11})^{-1}_{\text{right}}B_{12}$
and $\exists (A_{22})^{-1}_{\text{left}} \Longrightarrow \exists (B_{11})^{-1}_{\text{left}} = A_{11} - A_{12}(A_{22})^{-1}_{\text{left}}A_{21}$.
If $\exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}_{\text{left}} =: \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$*, then* $\exists (B_{11})^{-1}_{\text{left}} \Longrightarrow \exists (A_{22})^{-1}_{\text{left}} = B_{22} - B_{21}(B_{11})^{-1}_{\text{left}}B_{12}$
and $\exists (A_{22})^{-1}_{\text{right}} \Longrightarrow \exists (B_{11})^{-1}_{\text{right}} = A_{11} - A_{12}(A_{22})^{-1}_{\text{right}}A_{21}$.

(c3) $\exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} \Leftrightarrow \exists \begin{bmatrix} A_{11} & -A_{12} \\ -A_{21} & A_{22} \end{bmatrix}^{-1}$.

(c4) Let $\exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} =: \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ *and* $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \in \mathrm{TIC}$.
If $\exists A_{11}^{-1} \in \mathrm{TIC}_\infty$*, then* $\exists (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} = B_{22} \in \mathrm{TIC}$*; if* $\exists A_{22}^{-1} \in \mathrm{TIC}_\infty$*, then* $\exists (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} = B_{11} \in \mathrm{TIC}$.

(d1) **(Schur decomposition)** *Let* $\exists A_{11}^{-1}$*. Then*

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & I \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \tag{A.3}$$

$$= \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & I \end{bmatrix} \tag{A.4}$$

$$= \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \tag{A.5}$$

$$= \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix}. \tag{A.6}$$

Therefore, $\exists A^{-1} \Leftrightarrow \exists (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$ *(see also Lemma 11.3.13). If, in* *addition,* $\exists A^{-1}$*, then*

$$A^{-1} = \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix}^{-1} \begin{bmatrix} A_{11} & 0 \\ A_{21} & I \end{bmatrix}^{-1} \tag{A.7}$$

$$= \begin{bmatrix} A_{11} & A_{12} \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix}^{-1} \tag{A.8}$$

$$= \begin{bmatrix} A_{11} & A_{12} \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ (A^{-1})_{21} & (A^{-1})_{22} \end{bmatrix} \tag{A.9}$$

$$= \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \tag{A.10}$$

$$= \begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ -(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}, \tag{A.11}$$

In particular, $\exists \begin{bmatrix} I & A_{12} \\ A_{21} & I \end{bmatrix}^{-1} \Leftrightarrow \exists (I-A_{12}A_{21})^{-1} \Leftrightarrow \exists (I-A_{21}A_{12})^{-1}$*, and the possible inverse is necessarily*

$$\begin{bmatrix} I & A_{12} \\ A_{21} & I \end{bmatrix}^{-1} = \begin{bmatrix} (I-A_{12}A_{21})^{-1} & -A_{12}(I-A_{21}A_{12})^{-1} \\ -(I-A_{21}A_{12})^{-1}A_{21} & (I-A_{21}A_{12})^{-1} \end{bmatrix}. \tag{A.12}$$

(d2) $\exists (A_{11})^{-1}_{\text{left}}$ & $\exists (A_{22}-A_{21}(A_{11})^{-1}_{\text{left}}A_{12})^{-1}_{\text{left}} \implies \exists A^{-1}_{\text{left}}$*, (and (A.4), (A.8 and (A.9) hold for these left inverses);*

$\exists (A_{11})^{-1}_{\text{right}}$ & $\exists (A_{22}-A_{21}(A_{11})^{-1}_{\text{right}}A_{12})^{-1}_{\text{right}} \implies \exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}_{\text{right}}$ *(and (A.3) and (A.7) hold for these right inverses);*

$\exists (A_{11})^{-1}_{\text{left}}$ & $\exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}_{\text{right}} \implies \exists (A_{22}-A_{21}(A_{11})^{-1}_{\text{left}}A_{12})^{-1}_{\text{right}}$;

$\exists (A_{11})^{-1}_{\text{right}}$ & $\exists \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}_{\text{left}} \implies \exists (A_{22}-A_{21}(A_{11})^{-1}_{\text{right}}A_{12})^{-1}_{\text{left}}$.

(e1) **(Coprime)** *Given* A_1, A_2, B_1*, there is* B_2 *s.t.* $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$ *iff* $A_1B_1 = I$ & $A_2B_1 = 0$ & $A_2B_2' = I$ *for some* B_2'*. If the latter holds, then* $B_2 := (I-B_1A_1)B_2'$ *is as above (but not necessarily unique).*

(e2) Let $\begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = I$*. Then* $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \Leftrightarrow A_1B_1 = I$ & $A_2B_2 = I$*.*

(e3) Let $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}^{-1} = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$*. Then, for a given* A_1'*, there is* B_2' *s.t.* $\begin{bmatrix} A_1' \\ A_2 \end{bmatrix} \begin{bmatrix} B_1 & B_2' \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$ *iff* $A_1'B_1 = I$*. If* $A_1'B_1 = I$*, then* $\begin{bmatrix} A_1' \\ A_2 \end{bmatrix} \begin{bmatrix} B_1 & B_2' \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \Leftrightarrow B_2' = (I - B_1A_1')A_2$ & $\begin{bmatrix} B_1 & B_2' \end{bmatrix} \begin{bmatrix} A_1' \\ A_2 \end{bmatrix} = I$*.*

(e4) Let $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}^{-1} = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$*. Then, for a given* A_2'*, there are* A_1', B_2' *s.t.* $\begin{bmatrix} A_1' \\ A_2' \end{bmatrix}^{-1} = \begin{bmatrix} B_1 & B_2' \end{bmatrix}$ *iff* $A_2'B_1 = 0$ *and* $\exists (A_2'B_2)^{-1}$*. If such a solution* A_1', B_2' *exists, then all*

solutions are given by

$$\left(\begin{bmatrix} I & T \\ 0 & I \end{bmatrix}\begin{bmatrix} A_1 \\ A_2' \end{bmatrix}\right)^{-1} = \begin{bmatrix} B_1 & B_2(A_2'B_2)^{-1} \end{bmatrix}\begin{bmatrix} I & -T \\ 0 & I \end{bmatrix}, \quad T \in \mathcal{A}. \qquad \text{(A.13)}$$

(e5) If $\left[\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right]\left[\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right] = \left[\begin{smallmatrix} I & 0 \\ 0 & I \end{smallmatrix}\right]$ *and* $\exists (A_{22})^{-1}_{\text{left}}, (B_{11})^{-1}_{\text{right}}$ *(resp.* $\exists (A_{11})^{-1}_{\text{left}}, (B_{22})^{-1}_{\text{right}}$*), then* $\exists A_{22}^{-1}, B_{11}^{-1}$ *(resp.* $\exists A_{11}^{-1}, B_{22}^{-1}$*) and* $BA = \left[\begin{smallmatrix} I & 0 \\ 0 & I \end{smallmatrix}\right] = AB$*.*

As above, in parts (f1)–(h1) we require that the operators belong to the $\mathcal{A}$*'s that are compatible with the formulae; that is,* $x \in \mathcal{A}(X,Y)$*,* $y, q \in \mathcal{A}(Y,X)$*,* $z \in \mathcal{A}(X)$ *and* $w \in \mathcal{A}(Y)$*, where* $X, Y \in \mathcal{X}$*.*

(f1) $\exists y^{-1} = (xy)^{-1}x$ *if* $\exists x^{-1}, (xy)^{-1}$ *or if* $\exists (xy)^{-1}, (yx)^{-1}$*.*

(f2) $\exists x^{-1} \Leftrightarrow \exists (x^n)^{-1}$ *for* $n \in \{1,2,3,\ldots\}$*.*

(f3) $(x+y)^{-1} - x^{-1} = -(x+y)^{-1}yx^{-1}$ *if* $\exists (x+y)^{-1}, x^{-1}$*.*

(f4) $z(I-z)^{-1} = (I-z)^{-1}z$ & $I + z(I-z)^{-1} = (I-z)^{-1}$ *if* $\exists (I-z)^{-1}$*.*
$I + z(I-z)^{-1}_{\text{right}} = (I-z)^{-1}_{\text{right}}$ *if* $\exists (I-z)^{-1}_{\text{right}}$*.*

(f5) $z(I+z)^{-1} = (I+z)^{-1}z$ & $z(I+z)^{-1} = I - (I+z)^{-1}$ *if* $\exists (I+z)^{-1}$*.*

(f6) $\exists (I-xy)^{-1} \Leftrightarrow \exists (I-yx)^{-1} = I + y(I-xy)^{-1}x$ & $y(I-xy)^{-1} = (I-yx)^{-1}y$*.*
$\exists (I-xy)^{-1}_{\text{left}} \Leftrightarrow \exists (I-yx)^{-1}_{\text{left}} = I + y(I-xy)^{-1}_{\text{left}}x$ *(analogously for* $()^{-1}_{\text{right}}$*).*

(f7) $\{y \mid \exists (I-xy)^{-1}\} = \{(I+qx)^{-1}q \mid \exists (I+qx)^{-1}\}$ *and* $\{q \mid \exists (I+qx)^{-1}\} = \{y(I-xy)^{-1} \mid \exists (I-xy)^{-1}\}$*, for a fixed x.*

(g1) $\exists (I \pm yw^{-1}x)^{-1} = I \mp y(w \pm xy)^{-1}x$ *if* $\exists w^{-1}, (w \pm xy)^{-1}$*.*

(g2) $\exists (z + yw^{-1}x)^{-1} = z^{-1} - z^{-1}y(w + xz^{-1}y)^{-1}xz^{-1}$ *if* $\exists w^{-1}, z^{-1}, (w + xz^{-1}y)^{-1}$*.*

(h1) Let $x, y, z \in \mathcal{A}(X)$ *be invertible,* $X \in \mathcal{X}$*. Then* $xyz = y^{-1} \Leftrightarrow zyx = y^{-1}$*.*

The claims in the lemma are "the best possible ones", i.e., there is nothing superfluous in the conditions and nothing (that one would expect) missing in the conclusions. For any candidate "better" claims there are counter-examples, even in the case where both X and Y are the Hilbert space $\ell^2(\mathbf{N})$ (and when $\mathcal{A}$ is any the symbols listed in (1.)–(5.)).

Recall that if $A \in \mathcal{B}(X,Y)$, where X and Y are Hilbert spaces (and $\mathcal{A} = \mathcal{B}$), then $\exists A^{-1}_{\text{left}} \Leftrightarrow A^*A \gg 0$ (and $\exists A^{-1}_{\text{right}} \Leftrightarrow AA^* \gg 0$), by Lemma A.3.1(c1) (if $\mathcal{A} = \text{Hom}$, then $\exists A^{-1}_{\text{left}} \Leftrightarrow \text{Ker}(A) = 0$).

Part (c4) is an example about how to apply the claims of the lemma in two different $\mathcal{A}$'s.

One often needs to apply the lemma with induction, e.g., an upper triangular matrix is invertible if its diagonal blocks (operators) are, by (b1) (e.g., consider $\left[\begin{smallmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{smallmatrix}\right]$ first partitioned as $\left[\begin{array}{c|cc} A_{11} & A_{12} & A_{13} \\ \hline 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{array}\right]$).

Exchanging two rows (resp. columns) of an operator matrix corresponds to exchanging the corresponding columns (resp. rows) of its inverse matrix.

In case of linear operators between vector spaces, multiplication of the kth row of an operator matrix by a scalar α corresponds to the multiplication of the kth column of its (left, right) inverse by α^{-1}.

(To be completely rigorous, we mention that we have tacitly used the convention that the zero mapping is the only homomorphism between vector spaces with different scalar fields. However, we do not even think that anybody would like to use the lemma for such homomorphisms.)

Finally, we note that we often define operators by using the converse part of (a4).

Proof of Lemma A.1.1: It is obvious that the assumptions (algebraic laws) of Remark A.1.3 are satisfied in case (1.) as well as when $\mathcal{A}(X,Y)$ is the set of functions $S \to \mathcal{B}(X,Y)$, for a fixed set S. Thus, the same laws hold for cases (2.)–(5.) too, except that one has to verify that $\mathcal{A}$ is closed under addition and multiplication. For $\mathrm{L}^\infty_{\mathrm{strong}}$, this verification is given in Lemma F.1.3(b); for H^p_∞ and $\mathrm{H}^p_{\mathrm{strong},\infty}$ this follows from Lemma F.3.5. For SR_ω, UR_ω and ULR_ω this is contained in Lemma 6.2.5 for complex Hilbert spaces, and the general case is analogous. Classes of Definitions 2.6.1 and 2.6.3 will also do, by Lemma 2.6.2 and Theorem 2.6.4. For all other classes this is straightforward.

Although we have assumed U and Y to be Hilbert spaces in the definition of $\mathrm{TI}(U,Y)$ and its subspaces, there is no need for this in the definition itself. Thus we have been able to state (3.)–(5.) for arbitrary complex Banach spaces (assuming the definitions to be extended correspondingly).

(a1)&(a2) These are obvious.

(a3) Case "Hom" is obvious, case "$\mathcal{B}$" is [Rud73, Theorem 2.12(b)], case "TI" is Lemma 2.2.1(b) (because any finite-dimensional Banach spaces are Hilbert spaces) and implies the other cases (because in them $\mathcal{A}$ is a subclass of TI; see also Theorem 2.1.2 for H^∞).

(a4) In cases (1.)–(5.) the assumptions of Lemma A.1.1 are clearly satisfied. The claims "$P_i \in \mathcal{A}(Y,Y_i)$" and "$L_j \in \mathcal{A}(X_j,X)$" are clearly true as well.

The first matrix multiplication claim follows from $\sum_j A_{ij}x_j = \sum_j P_i A L_j x_j = P_i A \sum_j L_j x_j = P_i A x$, and the second from

$$P_i ABx = \sum_j A_{ij} \sum_k B_{jk} y_k = \sum_k (\sum_j A_{ij} B_{jk}) y_k. \tag{A.14}$$

The final claims are even easier.

(b1) This is obvious.

(b2) (3)&(4) $B_{11}A_{11} = I = A_{11}B_{11}$, $A_{22}B_{22} = I = B_{22}A_{22}$, hence $\exists A_{11}^{-1}, A_{22}^{-1}$. (1)&(2) $A_{22}B_{21} = 0 = B_{21}A_{11}$, so, if $\exists (A_{11})^{-1}_{\mathrm{right}}$ tai $\exists (A_{22})^{-1}_{\mathrm{left}}$, then $B_{21} = 0$, so the claim follows from (3). (7) $B_{11}A_{11} = I_{X_1}$ and $A_{22}B_{22} = I_{Y_2}$, so if A_{11} or A_{22} belongs to $\mathcal{GB} + \mathcal{BC}$ (and X_1, X_2, Y_1, Y_2 are Banach spaces), then it is invertible, by Lemma A.3.4(B3), hence then the claim follows from (1) or (2). (5)&(6) Work as in (7).

(Similarly, $\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{bmatrix}$.)

“A counter-example” ($B_{12}=0$ is not enough): $\exists \begin{bmatrix} R & P \\ 0 & L \end{bmatrix}^{-1} =: \begin{bmatrix} L & 0 \\ P & R \end{bmatrix}$ (one can interpret these as the left and right shift on $\ell^2(\mathbf{Z}) = \ell^2(\mathbf{Z}_-)\times\ell^2(\mathbf{N})$, respectively); here R [L] is the right [left] shift on $\ell^2(\mathbf{N})$ and $P := I-RL$ is the projection to the first element; note that $RL = I-P$, $LR = I = P+RL$ and $LP = 0 = PR$.

(c1) & (c2) Assume that $\exists B_{11}^{-1}$ (alternatively, as a right inverse only) $AB = I \implies 0 = A_{21}B_{11}+A_{22}B_{21} \implies A_{21} = -A_{22}B_{21}B_{11}^{-1}$, $I = A_{21}B_{12}+A_{22}B_{22} = -A_{22}B_{21}B_{11}^{-1}B_{12}+A_{22}B_{22} = A_{22}(B_{22}-B_{21}B_{11}^{-1}B_{12})$ from this we get $A_{22}^{-1}A_{21} = -B_{21}B_{11}^{-1}$ if $\exists A_{22}^{-1}, B_{11}^{-1}$), in particular, $\exists (A_{22})^{-1}_{\text{right}}$.

Similarly, $I = BA$ implies that $0 = B_{11}A_{12}+B_{12}A_{22}$, hence $A_{12} = -B_{11}^{-1}B_{12}A_{22}$, and $I = B_{21}A_{12}+B_{22}A_{22} = B_{22}A_{22}-B_{21}B_{11}^{-1}B_{12}A_{22} = (B_{22}-B_{21}B_{11}^{-1}B_{12})A_{22}$ (this holds for merely left-invertible B_{11} too), hence $\exists A_{22}^{-1} = B_{22}-B_{21}B_{11}^{-1}B_{12}$.

The converse claims (assuming the invertibility of A_{22} are analogous.

If $\dim X_2 < \infty$, then $\exists (A_{22})^{-1}_{\text{left}} \Leftrightarrow \exists A_{22}^{-1} \Leftrightarrow \exists (A_{22})^{-1}_{\text{right}}$; if $\dim X_1 < \infty$, then (c2) implies that $\exists (A_{22})^{-1}_{\text{left}} \Leftrightarrow \exists (B_{11})^{-1}_{\text{left}} \Leftrightarrow \exists B_{11}^{-1} \Leftrightarrow \exists A_{22}^{-1}$ etc.

(c3) This follows from equation $\begin{bmatrix} A_{11} & -A_{12} \\ -A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$.

(c4) This follows from (c1) by setting $\mathcal{A} = \text{TIC}_\infty$ (we present this application for an easy reference).

(d1) $\begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$, hence these three operator matrices must be invertible (because two of them are, by the assumptions), therefore so is $A_{22}-A_{21}A_{11}^{-1}A_{12}$, by (b2)(1). Thus,

$$\begin{aligned}
\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} &= \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix}^{-1}\begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \\
&= \begin{bmatrix} I & -A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ 0 & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}\begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \\
&= \begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ -(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}.
\end{aligned}$$

The case for $\begin{bmatrix} I & A_{12} \\ A_{21} & I \end{bmatrix}$ follows from this and (f6).

(We may have $\nexists A_{22}^{-1}$, e.g., $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$.)

(d2) If $\exists (A_{11})^{-1}_{\text{left}}$, then (A.4) applies; if $\exists (A_{11})^{-1}_{\text{right}}$, then (A.3) applies; the other claims follow from these (and (b1)).

(e1) It is obvious that the three conditions are necessary. Conversely, by taking $B_2' = (I-B_1A_1')A_2$ we get $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}\begin{bmatrix} B_1 & B_2 \end{bmatrix} = I$.

(e2) It is obvious that the three conditions are necessary. For the converse, assume that $B_1A_1+B_2A_2 = I$ and $A_1B_1 = I$. By (d1), the invertibility of $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}\begin{bmatrix} B_1 & B_2 \end{bmatrix} = \begin{bmatrix} I & A_1B_2 \\ A_2B_1 & I \end{bmatrix}$ follows from that of $I-A_2B_1I^{-1}A_1B_2 = I-A_2(I-B_2A_2)B_2 = I-A_2B_2+A_2B_2A_2B_2 = I$.

(e3) The condition $FB_1 = I$ is obviously necessary. Conversely, for $E = (I-B_1F)B_2$, one soon verifies that $\begin{bmatrix} F \\ A_2 \end{bmatrix}\begin{bmatrix} B_1 & E \end{bmatrix} = I$ and $\begin{bmatrix} B_1 & E \end{bmatrix}\begin{bmatrix} F \\ A_2 \end{bmatrix} = I$, hence

$\begin{bmatrix} B_1 & E \end{bmatrix} = \begin{bmatrix} F \\ A_2 \end{bmatrix}^{-1}$ is unique.

(e4) The condition $GB_1 = 0$ is obviously necessary. Conversely, assume $GB_1 = 0$. Then $\begin{bmatrix} Z \\ G \end{bmatrix}$ is invertible iff $\begin{bmatrix} Z \\ G \end{bmatrix} \begin{bmatrix} B_1 & B_2 \end{bmatrix} = \begin{bmatrix} ZB_1 & ZB_2 \\ GB_1 & GB_2 \end{bmatrix} = \begin{bmatrix} ZB_1 & ZB_2 \\ 0 & GB_2 \end{bmatrix}$ is, i.e., iff ZB_1 and GB_2 are; therefore the condition $\exists (GB_2)^{-1}$ is necessary. It is also sufficient, because we can choose $Z = A_1,\ W = B_2$.

Whenever $\begin{bmatrix} Z \\ G \end{bmatrix}^{-1} = \begin{bmatrix} B_1 & W \end{bmatrix}$, we have $ZB_1 = I$ & $GB_2 = I$, hence then $\begin{bmatrix} Z \\ G \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} I & ZB_2 \\ 0 & I \end{bmatrix}^{-1}$; therefore all solutions are of the form (A.13); conversely, it is obvious that each $T \in \mathcal{A}(Y_2, Y_1)$ determines a solution.

(e5) We have $\exists A_{22}^{-1}, B_{11}^{-1}$, by (c2). Now $A_{22}^{-1} A_{21} = B_{21} B_{11}^{-1}$, so we obtain the (right-)invertibility of B from (d2), because $B_{22} - B_{21} B_{11}^{-1} B_{12} = B_{22} - A_{22}^{-1} A_{21} B_{12} = A_{22}^{-1}$ is invertible. The other case is obtained from this by permuting the rows and columns of A and B.

(f1) Clearly $(xy)^{-1}x$ is a left inverse of y, so we only have to show that $y(xy)^{-1}x$ is invertible (hence equal to I).

1° $xy(xy)^{-1}xx^{-1} = I \implies y(xy)^{-1}x = I$. 2° If $(yx)y(xy)^{-1}x = yx$ is invertible, then so is $y(xy)^{-1}x$.

(f2) Set $y := x,\ x := x^{n-1}$ in (f1), and use induction.

(f3) $\Leftrightarrow I - (x+y)x^{-1} = -yx^{-1} \Leftrightarrow -yx^{-1} = -yx^{-1}$.

(f4) $I + z(I-z)^{-1} = (I - z + z)(I-z)^{-1} = (I-z)^{-1} = (I-z)^{-1}(I - z + z) = I + (I-z)^{-1}z$.

(f5) Work as in (f4).

(f6) $[I + y(I-xy)^{-1}x][I - yx] = I - yx + y(I-xy)^{-1}x - y(I-xy)^{-1}xyx = I + y[-x + (I-xy)^{-1}(I-xy)x] = I$ & $[I - yx][I + y(I-xy)^{-1}x] = I - yx + y(I-xy)^{-1}x - yxy(I-xy)^{-1} = I + y[-x + (I-xy)(I-xy)^{-1}x] = I$. $(I-yx)^{-1}y = [I + y(I-xy)^{-1}x]y \overset{(f4)}{=} y + yxy(I-xy)^{-1} = y[I + xy(I-xy)^{-1}] = y[(e - xy + xy)(I-xy)^{-1}] = y(I-xy)^{-1}$.

N.B: $y(I-y)_{\text{left}}^{-1} \neq (I-y)_{\text{left}}^{-1} y$ when $y = I - R$, and R and $L = R_{\text{left}}^{-1}$ are the right and left translation on $\ell^2(\mathbf{N})$, respectively (because $(I-R)L \neq L(I-R)$).

(f7) 1° $y := (I+qx)^{-1}q \implies I - xy = I - xq(I+xq)^{-1}\ [(f6)]\ = (I + xq - xq)(I+xq)^{-1} = (I+xq)^{-1}$, hence $\exists (I-xy)^{-1} = (I+xq)$.
2° $q := y(I-xy)^{-1} \implies I + qx = I + yx(I-yx)^{-1} = (I-yx)^{-1}$ resp.
3° Let q be s.t. $\exists (I+qx)^{-1}$. $y := y_q := (I+qx)^{-1}q$. Then $q_{y_q} := y(I-xy)^{-1} = y(I+xq) = (I-qx)^{-1}q(I+xq) = (I-qx)^{-1}(I+qx)q = q$, hence each such q is determined by the y_q it determines, i.e., $q = q_{y_q}$, hence $q \mapsto y_q$ is an injection. Similarly, $\exists q := q_y := y(I-xy)^{-1} \implies y_{q_y} = (I+qx)^{-1}q = (I - yx)y(I-xy)^{-1} = y$, hence each such y is determined by the q_y it determines; therefore this correspondence is bijective.

(g1) Set $z = I$ in (g2) (and "$y = \pm y$").

(g2) Set $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} := \begin{bmatrix} z & y \\ \mp x & w \end{bmatrix} =: \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}^{-1}$ in (d1). The formula $B_{11} = A_{11}^{-1} + A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1}$ is obtained from (d1). Part (c1) implies that $\exists B_{11}^{-1} = A_{11} - A_{12} A_{22}^{-1} A_{21}$.

(h1) Assume that $xyz = y^{-1}$. Then $zyxyz = zyy^{-1} = z$, hence $zyxy = I$, hence $zyx = y^{-1}$. The converse is obtained analogously. □

Remark A.1.2 *Lemma A.1.1 is also valid in case*

(5.) $\mathcal{X}$ *is the collection of groups, and* $\mathcal{A}(X,Y)$ *is the set of group homomorphisms* $X \to Y$,

provided we replace the assumptions of the form "$\dim X = \dim Y < \infty$ *by "Inv*(X,Y)*", and "If* X_1 *or* X_2 *is finite-dimensional" in (c1) by "If Inv*(X_1,X_1) *or Inv*(X_2,X_2)*", where Inv*(X,Y) *is the assumption on* X *and* Y *that*

any $A \in \mathcal{A}(X,Y)$ *is left (resp. right) invertible iff it is invertible.* □

Except for the projection–imbedding claims, the proof of Lemma A.1.1 is based only on the properties listed below, hence its conclusions are true under a wider set of circumstances:

Remark A.1.3 *Let* $\mathcal{X}$ *be a set. Let* $\mathcal{A}(X,Y)$ *be a group with a zero* $0 = 0_{\mathcal{A}(X,Y)}$, *for all* $X,Y \in \mathcal{X}$, *and let for each* $X,Y,Z \in \mathcal{X}$ *there be an operation* $\mathcal{A}(X,Y) \times \mathcal{A}(Y,Z) \to \mathcal{A}(X,Z)$ *be defined in such a way that this operation is associative and distributive:*

$$(AB)C = A(BC) \text{ and } (A+A')(B+B') = AB + AB' + A'B + A'B' \tag{A.15}$$

for all $A,A' \in \mathcal{A}(X,Y)$, $B,B' \in \mathcal{A}(Y,Z)$, $C \in \mathcal{A}(X,Z)$*;* $\mathcal{A}(X)$ *is a ring with a unit* $I = I_{\mathcal{A}(X)}$ *(we allow for* $I = 0$*, that is, for* $\mathcal{A}(X) = \{0\}$*); and the rules*

$$A0 = 0 = 0B, \quad A = AI, \quad IB = B \qquad (A \in \mathcal{A}(X,Y),\ B \in \mathcal{A}(Y,Z)) \tag{A.16}$$

are obeyed.

As explained below, the structure $(\mathcal{A},\mathcal{X})$ *determines naturally another structure that also satisfies the above assumptions: If* $n,m,N \in \{1,2,3,\ldots\}$*,* $A_{ij},C_{ij} \in \mathcal{A}(X_j,Y_i)$ $B_{jk} \in \mathcal{A}(Y_i,Z_k)$ *(*$i = 1,\ldots,m,\ j = 1,\ldots,n,\ k = 1,\ldots,N$*), we denote by* A *the* $\mathcal{A}$ *matrix*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nm} \end{bmatrix}; \tag{A.17}$$

similarly for B *and* C*. For such representations, we use the standard matrix operation rules, that is,* $(AB)_{ij} := \sum_k A_{ik}B_{kj}$*,* $(A+C)_{ij} := A_{ij} + C_{ij}$*,* $(-A)_{ij} := -A_{ij}$*.*

For these $\mathcal{A}$ *and* $\mathcal{X}$*, all the conclusions of Lemma A.1.1 except (a4) hold if we make the Inv*(X,Y) *replacements of Remark A.1.2 and replace all expressions of the form* $\mathcal{A}(X_1 \times X_2, Y_1 \times Y_2)$*,* $X_1,X_2,Y_1,Y_2 \in \mathcal{X}$*, by the set of the corresponding* $\mathcal{A}$*-matrices.* □

In particular, Lemma A.1.1 applies when $\mathcal{A}(X)$ is a ring with a unit and $\mathcal{X} = \{X\}$ (here X need not stand for anything).

Notes

Many of the formulae of this section and some additional ones are often used in control theory; some of them and further formulae are presented in many matrix calculus textbooks.

A.2 Topological spaces

Noli turbare circulos meos!

— Archimedes (287–212 BC)

In this section, we very briefly introduce metric spaces and other topological spaces and present some well-known lemmas. For most of this monograph, basic knowledge on Hilbert and Banach spaces (Section A.3) is sufficient. Therefore, the reader may skip this section unless (s)he wishes to go through also the proofs of certain auxiliary results.

The details presented here are sufficient for most of our purposes, but a reader wishing to know more may consult any book on topological spaces; also many books on functional analysis (e.g., [Rud86] or [Rud73]) contain the basic theory of topological spaces.

Since any metric space is a topological space, it is adviseable to visualize the topological spaces as metric ones (or as $\mathbf{R}^2$) to get an intuitive picture on the general case (all concepts defined here are direct generalizations of those defined for metric spaces). Most spaces that we meet are metric (see, e.g., [Rud76] for the theory of metric spaces).

The most important nonmetrizable topologies (see Exercises 2.1 and 3.15 of [Rud73]) are the weak and weak* topologies of infinite-dimensional Banach spaces.

A *topology* on a set Q is a collection $\mathcal{T}$ of subsets of Q s.t. $\emptyset, Q \in \mathcal{T}$, and $\mathcal{T}$ is closed under finite interjections and arbitrary unions. We call the pair $(Q,\mathcal{T})$ (or just Q when there is no ambiguity about $\mathcal{T}$ or when we do not need to specify it) a *topological space*.

The elements of $\mathcal{T}$ are called *open* and their complements are called *closed*. If $E \subset Q$, then $E^o := \cup\{V \in \mathcal{T} \,\big|\, V \subset E\}$ is the *interior* and $\overline{E} := \cap\{F \,\big|\, F^c \in \mathcal{T}$ and $E \subset F\}$ is the *closure* of E. We call $\partial E := \overline{E} \cap \overline{E^c}$ the *boundary* of E. A set $K \subset Q$ is *compact* if $\mathcal{V} \subset \mathcal{T}$ and $K \subset \cup\mathcal{V}$ imply that $K \subset \cup\mathcal{V}'$ for some finite $\mathcal{V}' \subset \mathcal{V}$.

Let also $(Q_2,\mathcal{T}_2)$ be a topological space. Then $f : Q \to Q_2$ is *continuous* (i.e., $f \in \mathcal{C}(Q;Q_2)$) if $f^{-1}[V] \in \mathcal{T}$ for any open $V \in \mathcal{T}_2$. If f is a continuous bijection and also its inverse is continuous, then f is called a *homeomorphism*. The sets Q and Q_2 are called *homeomorphic* if there is a homeomorphism $Q \to Q_2$. The set $Q \times Q_2$ is usually equipped by its *product topology*, which is the smallest topology containing $\mathcal{T} \times \mathcal{T}_2$.

We equip any subset E of Q with the topology $\{V \cap E \,\big|\, V \in \mathcal{T}\}$ *inherited from* Q (unless something else is indicated).

A sequence $\{q_n\} \subset Q$ *converges to* $q \in Q$, i.e., $\lim_{n\to+\infty} q_n = q$, iff, for each open set $V \ni q$, there is $N_V \in \mathbf{N}$ s.t. $q_n \in V$ for all $n \ge N_V$. If this is the case, we also say that $\{q_n\}$ *converges in* Q.

If $\mathcal{T}$ and $\mathcal{T}'$ are topologies on Q and $\mathcal{T} \subset \mathcal{T}'$, then $\mathcal{T}$ is *weaker* than $\mathcal{T}'$ and $\mathcal{T}'$ is *stronger* than $\mathcal{T}$. It obviously follows that if T is a nonempty collection of topologies on Q and $\mathcal{T}_0$ is the weakest element of T, then $\mathcal{T}_0 = \cap_{\mathcal{T}\in T}\mathcal{T}$.

A *neighborhood* of $q_0 \in Q$ means an open set containing q_0. A point $q_0 \in Q$ is called a *limit point* of $E \subset Q$ if $q_0 \in \overline{E \setminus \{q_0\}}$, equivalently, if every neighborhood of q_0 contains an element of $E \setminus \{q_0\}$.

A set $E \subset Q$ is *disconnected* if there are nonempty sets $A, B \subset Q$ s.t. $E = A \cup B$ and $A \cap \overline{B} = \emptyset = \overline{A} \cap B$. Otherwise E is *connected*. An *interval* is a nonempty connected subset of $\mathbf{R}$.

Let Q be a set. A function $d : Q \times Q \to [0, +\infty)$ is called a *metric* if 1. $d(x,y) = 0 \Leftrightarrow x = y$, 2. $d(x,y) = d(y,x)$, and 3. $d(x,z) \le d(x,y) + d(y,z)$ for all $x, y, z \in Q$. A set equipped with a metric is called a *metric space* (thus, we call Q or (Q,d) a metric space if the above conditions are satisfied). A topological space (or a topology) is called *metrizable* if it is induced by some metric.

If d is a metric on a set Q, then we usually equip Q with the topology induced by d, which consists of arbitrary unions of open balls $D(q,r) := \{q' \in Q \,\big|\, d(q',q) < r\}$ ($q \in Q$, $r > 0$). It follows that d becomes continuous $Q \to \mathbf{R}$.

We recall from [Rud76] that if (Q,d) and (Q',d') are metric spaces, $q_0 \in Q$, $q_0' \in Q'$ and $f : Q \to Q'$, then $\lim_{q \to q_0} f(q) = q_0'$ iff $\lim_{n \to \infty} f(q_n) = q_0'$ whenever $q_n \to q_0$ (this is not true for all topological spaces Q, not even for all TVSs). A map $f : Q \to Q'$ satisfying $d(x,y) = d'(f(x), f(y))$ for all $x, y \in Q$ is called an *isometry* (or *isometric*). If $f : Q \to Q'$ is s.t. for all $\varepsilon > 0$ there is $\delta_\varepsilon > 0$ s.t. $d(x,y) < \delta_\varepsilon \Longrightarrow d'(f(x), f(y)) < \varepsilon$ for all $x, y \in Q$, then f is *uniformly continuous*.

Let (Q,d) be a metric space. A sequence $\{q_n\} \subset Q$ is a *Cauchy-sequence* in Q if for each $\varepsilon > 0$, there is $N_\varepsilon \in \mathbf{N}$ s.t. $d(q_n, q_m) < \varepsilon$ for all $n, m \ge N_\varepsilon$. It easy to show that any converging sequence is a Cauchy-sequence. If any Cauchy-sequence in Q converges in Q, then the metric space Q is called *complete*.

A compact subset of a metric space is closed and bounded; the converse holds for subsets of $\mathbf{R}^n$ (or of $\mathbf{C}^n$). If $K \subset Q$ is compact and $f \in \mathcal{C}(K, Q')$, then $f[K]$ is compact in Q'.

If Q is a metric space and $K, E \subset Q$, then we set $d(q,K) := \inf_{q' \in K} d(q,q')$, $d(E,K) := \inf_{q \in E,\ q' \in K} d(q,q')$; if K is compact and nonempty and $q \notin K$, then $d(q,K) = \min_{q' \in K} d(q,q') > 0$. If $a \in Q$, then, obviously, $d(a,E) = 0$ iff $a \in \overline{E}$.

Lemma A.2.1 *Let $\emptyset \ne K \subset V \subset Q$, where Q is a metric space, V is open and K is compact.*

- *(a) If $f : K \to \mathbf{R}$ is continuous, then $\min_{q \in Q} f(q)$ and $\max_{q \in Q} f(q)$ exist.*
- *(b) If $f : K \to Q'$ is continuous, where Q' is a metric space, then f is uniformly continuous, i.e., for all $\varepsilon > 0$ there is $\delta > 0$ s.t. $x, y \in K \& d(x,y) < \delta \Longrightarrow d(f(x), f(y)) < \varepsilon$.*
- *(c) If $V \ne Q$, then $d(K, V^c) = d(a, V^c) > 0$ for some $a \in K$.*
- *(d) If $V \ne Q = \mathbf{R}^n$, then there are $a \in K$ and $b \in V^c$ s.t. $d(a,b) = d(K, V^c) := \inf_{x \in K, y \in V^c} d(x,y)$.*

Proof: (a)&(b) These are Theorems 4.16 and 4.19 of [Rud76], respectively.

(c) Assume that $V \ne Q$. Then $d(\cdot, V^c)$ attains a minimum on K, by (a), hence $d(K, V^c) = d(a, V^c)$ for some $a \in K$. Since $a \notin V^c = \overline{V^c}$, we have $d(a, V^c) > 0$.

(d) Assume that $V \ne Q = \mathbf{R}^n$. Choose $x \in K$. Choose $R > 0$ s.t. $K \subset D_R$ and $D_R \cap V^c \ne \emptyset$, where $D_R := \{q \in Q \,\big|\, |q| < R\}$. The set $F := \overline{D_{3R}} \cap V^c$ is closed and bounded, hence compact. By (a), there is

$a \in K$ s.t. $d(a,F) = d(K,F)$. By (a), there is $b \in F'$ s.t. $d(a,b) = d(a,F)$. But $d(K,V^c) = \min\{d(K,F), d(K,V^c \setminus F)\} = d(K,F)$ (since $d(K,F) < 2R$ and $d(K,V^c \setminus F) > 2R$) and $d(K,F) = d(a,b)$. □

The following well-known fact is sometimes handy:

Lemma A.2.2 *Let $V \subset \mathbf{R}^n$ be open. Then V is the union of an at most countable number of disjoint open, connected sets.*

Thus, open $V \subset \mathbf{R}$ is the union of an at most countable number of disjoint open intervals.

Proof: Let $\{q_k\}_{k\in\mathbf{N}} \subset \mathbf{R}^n$ be dense (e.g., enumerate $\mathbf{Q}^n$). For each $k \in \mathbf{N}$, take $V_k := \emptyset$ if $q_k \notin V$ or if $q_k \in V_j$ for some $j < k$; otherwise let V_k be the connected component of V that contains q_k. Then $V = \cup_{k\in\mathbf{N}} V_k$, and the sets V_k are open (because, obviously, $\partial V_k \subset \partial V$). □

The following lemma often allows one to work on an open set with a compact closure instead of a general open set:

Lemma A.2.3 (Compact exhaustion of Ω) *Let $\Omega \subset \mathbf{R}^n$ be open. Set $K_k := \{q \in \Omega \,\big|\, |q| \le k \ \& \ d(q,\Omega^c) \ge 1/k\}$ ($k \in \mathbf{N}+1$). Then each K_k is a compact subset of Ω, $K_1 \subset K_2 \subset \cdots$, $\Omega = \cup_k K_k^o$, and each compact $K \subset \Omega$ is contained in some K_k^o.*

Note that each K_k is compact.

Proof: This quite obvious. If $K \subset \Omega$ is compact, then some finite subset of sets K_k^o contains K, hence some K_k contains K. □

Notes

All of this is well known, see any book on topology (e.g., [Bredon], [Kelley] or even [Rud86]) for more.

A.3 Hilbert and Banach spaces

Mathematicians often resort to something called Hilbert space, which is described as being n-dimensional. Like modern sex, any number can play.

— James Blish, "The Quincunx of Time"

In this section, we present certain standard definitions and useful facts on Hilbert and Banach spaces; see any text on functional analysis (e.g., [Rud86] or [Rud73]) for their basic properties (and for details for most facts presented below).

Recall that in this appendix, the scalar field of any vector space is assumed to be **K** (that is **R** or **C**). If $\mathbf{K} = \mathbf{R}$, then, naturally, conjugation $\alpha \mapsto \bar{\alpha}$ becomes the identity operator on **K**, conjugate-linear is the same as linear, and sesquilinear is the same as bilinear.

A set A is closed under a function if the function maps the elements of A into A.

Let $\mathcal{T}$ be a topology for a vector space X. If $\{x\}^c \in \mathcal{T}$ for each $x \in X$, and sum and scalar multiplication on X are continuous, then X (or $(X,\mathcal{T})$) is called a *topological vector space (TVS)*. Most important examples of TVSs are Banach spaces, and we need other TVSs only in some external references. See, e.g., [Rud73] for more on TVSs.

A *normed space* is a vector space X equipped with a function (*norm*) $\|\cdot\|$: $X \to [0,+\infty)$ satisfying $\|\alpha x\| = |\alpha|\|x\|$, $\|x+y\| \le \|x\|+\|y\|$ and $\|x\| = 0 \Rightarrow x = 0$ for all $x \in X$, $\alpha \in \mathbf{K}$. We often write $\|\cdot\|_X := \|\cdot\|$ to distinguish between the norms of different normed spaces.

An *inner product space* is a vector space H equipped with a function (*inner product*) $\langle\cdot,\cdot\rangle : H \times H \to \mathbf{K}$ satisfying $\langle y,x\rangle = \overline{\langle x,y\rangle}$, $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$, $\langle \alpha x,y\rangle = \alpha\langle x,y\rangle$, $\langle x,x\rangle \ge 0$ and $\langle x,x\rangle = 0 \Rightarrow x = 0$ for all $x,y,z \in H$, $\alpha \in \mathbf{K}$. We often write $\langle\cdot,\cdot\rangle_H := \langle\cdot,\cdot\rangle$ to distinguish between the inner products of different normed spaces.

We equip any inner product space with norm $\|x\| = \langle x,x\rangle^{1/2}$ (it follows that any inner product space is a normed space). We equip any normed space X by the metric $d(x,y) := \|x-y\|_B$ (it follows that any normed space is a metric space).

A complete normed space is called a *Banach space*. A complete inner product space is called a *Hilbert space* (in particular, any Hilbert space is a Banach space).

The space $\mathbf{K}^n$ ($n \in 1+\mathbf{N}$) is equipped with the canonical inner product $\langle x,y\rangle_{\mathbf{K}^n} := \sum_{k=1}^n x_k\overline{y_k}$.

Let X and Y be normed spaces. We set $\|x\|_X := +\infty$ for $x \notin X$. We equip $X \times Y$ with the product topology, i.e., with the norm $\|(x,y)\|^2_{X\times Y} := \|x\|^2_X + \|y\|^2_Y$ (or some equivalent norm, such as $\max\{\|x\|_X, \|y\|_Y\}$ or $(\|x\|^2_X + \|y\|^2_Y)^{1/2}$).

By $\mathcal{B}(X,Y)$ we denote the normed space of continuous (i.e., bounded) linear operators $L : X \to Y$ with norm $\|L\|_{\mathcal{B}(X,Y)} := \sup_{\|x\|_X \le 1} \|Lx\|_Y$, and vector operations $(\alpha L + \beta L')x := \alpha(Lx) + \beta(L'x)$ ($L, L' \in \mathcal{B}(X,Y)$, $\alpha,\beta \in \mathbf{K}$, $x \in X$). We usually write $Lx := L(x)$ when L is linear. One easily verifies that $\mathcal{B}(X,Y)$ is a Banach space iff Y is a Banach space.

The space $X^* := \mathcal{B}(X,\mathbf{K})$ is called the *dual space of Xdual space (X^*)* (see also Remark A.3.22), and we set $X^{**} := (X^*)^*$. We identify $x \in X$ and the element

$x^{**} : \Lambda \mapsto \Lambda x$ of X^{**}. If all elements of X^{**} are of this form, then X is called *reflexive* (and then X is isometrically isomorphic to X^{**}, hence a Banach space). Any Hilbert space is reflexive.

Let X and Y be normed spaces. To each $T \in \mathcal{B}(X,Y)$ corresponds a unique $T^{\mathrm{B}} \in \mathcal{B}(Y^*,X^*)$ s.t. $y^*(Tx) = (T^{\mathrm{B}}y^*)x$ for all $x \in X$ and $y^* \in Y^*$; moreover, $\|T^{\mathrm{B}}\| = \|T\|$. We call T^{B} the *Banach adjoint* of T.

Let X and Y be Hilbert spaces. To each $T \in \mathcal{B}(X,Y)$ corresponds a unique $T^{\mathrm{H}} \in \mathcal{B}(Y,X)$ s.t. $\langle Tx,y\rangle_Y = \langle x, T^{\mathrm{H}}y\rangle_X$ for all $x \in X$ and $y \in Y$; moreover, $\|T^{\mathrm{H}}\| = \|T\|$. We call T^{H} the *Hilbert adjoint* of T. In a Banach space context, T^* denotes T^{B}, whereas in a Hilbert space context, T^* denotes T^{H} (unless we use pivot spaces, see Definition A.3.23 and Lemma A.3.24).

Let H be a Hilbert space and let B be a Banach space. A set $E \subset H$ is *orthonormal* if $\langle x,y\rangle = 0$ whenever $x,y \in E$, $x \neq y$ and $\langle x,x\rangle = 1$ for all $x \in E$. If $\operatorname{span} E$ is dense in H, then E is an *orthonormal basis* of H. An operator $P \in \mathcal{B}(B)$ is a *projection* if $P^2 = P$ (here $P^2 := PP$). A projection $P \in \mathcal{B}(H)$ is an *orthogonal projection* if $P = P^*$ (equivalently, $\operatorname{Ran}(P) = \operatorname{Ker}(P)^\perp$, by Theorem 12.14 of [Rud73]).

By $\mathcal{BC}(X,Y)$ we denote the set of linear mappings $T : X \to Y$ that are *compact*, that is, such that $\overline{\{Tx \,|\, x \in X,\ \|x\| < 1\}}$ is compact. It follows that $\mathcal{BC}(X,Y)$ is a subspace of $\mathcal{B}(X,Y)$.

The *weak topology* of X is the weakest topology on X on which each $\Lambda \in X^*$ is continuous. The *weak* topology* of X^* is the weakest topology on X on which each of the maps $x : \Lambda \mapsto \Lambda x$ $(x \in X)$ is continuous. We do not use these two topologies except when we explicitly say so. (If X is finite-dimensional, then $X = X^*$ and the weak, weak* and original (normed) topologies of X and X^* coincide with the standard Euclidean topology of X. In general, weak and weak* topologies need not be even metrizable.) See, e.g., [Rud73] for further details.

We set $\mathcal{GB}(X,Y) := \{L \in \mathcal{B}(X,Y) \,|\, LT = I_Y \,\&\, TL = I_X \text{ for some } T \in \mathcal{B}(Y,X)\}$ (and we write $T := L^{-1}$ for T,L as above) hence $\mathcal{GB}(X)$ becomes the subgroup of invertible operators ($\mathcal{G}$ for "group"). It follows that $L \in \mathcal{GB}(X,Y)$ iff $L \in \mathcal{B}(X,Y)$ and L is onto and one-to-one (i.e., the inverse is necessarily bounded), by (part (c3)(ii) of) Lemma A.3.4(F1).

For $L \in \mathcal{B}(X,Y)$ we set $\sigma(L) := \{\zeta \in \mathbf{K} \,|\, \zeta - L \notin \mathcal{GB}(X,Y)\}$, $\rho(L) := \sup|\sigma(L)|$ (see Lemma A.3.3; recall that $\zeta - L := \zeta I - L$).

An element of $\mathcal{GB}(X,Y)$ is called a *(Banach) isomorphism* of X *onto* Y. Whenever the range of LX is closed in Y (hence a Banach space itself if Y is complete) and L is as isomorphism of X onto LX, then L is an isomorphism of X *into* Y. (Note the standard abuse of language: we do not require an isomorphism to be an *isometry*, that is, to satisfy $\|Lx\| = \|x\|$ for all $x \in X$; a more rigorous term would be "a topological vector space isomorphism".)

Thus, if X and Y are Hilbert spaces, then a map $L \in \mathcal{B}(X,Y)$ is an isometric isomorphism of X onto Y iff L is *unitary*, i.e., $L^*L = I = LL^*$ (note that we often use same I (resp. 0) for identity (resp. zero) mappings in different groups.

Two norms, say $\|\cdot\|$ and $\|\cdot\|'$, defined on a normed space X are called *equivalent* if they define same topology. This is the case iff there are $\varepsilon, M \in (0,\infty)$ s.t. $\varepsilon\|x\| \leq \|x\|' \leq M\|x\|$ for all $x \in X$.

Let H be a Hilbert space. By $\langle x,y\rangle_H$ (often just $\langle x,y\rangle$), we denote the (sesquilinear, i.e., $\langle \alpha x, \beta y\rangle = \alpha\bar{\beta}\langle x,y\rangle$) inner product $H \times H \to \mathbf{K}$. It is well known that $H^* = \{x \mapsto \langle x,y\rangle_H \mid y \in H\}$, hence H is reflexive. The mapping $\Lambda_x : y \mapsto \langle y,x\rangle_H$ is conjugate-linear (i.e., $\Lambda_{\alpha x} = \bar{\alpha}\Lambda_x$, $\Lambda_{x+y} = \Lambda_x + \Lambda_y$), isometric and onto, hence $\Lambda \in \mathcal{GC}(H,H^*)$ (an element of $\mathcal{GC}$ is called a *homeomorphism*).

Let $x,y \in H$, $A \subset H$ and $B \subset H$. If $\langle x,y\rangle = 0$, then we write $x \perp y$; if $a \perp b$ for all $a \in A$ and $b \in B$, then we write $A \perp B$. Obviously, $A^\perp := \{x \in H \mid x \perp A\}$ is a closed subspace of H.

Let $T,S \in \mathcal{B}(H)$. We call T *nonnegative* [*positive*] and write $T \geq 0$ [$T > 0$] if $\langle Tx,x\rangle \geq 0$ [> 0] for all $x \in H \setminus \{0\}$. By $T \geq S$ we mean that $T = T^*$, $S = S^*$, and $T - S \geq 0$. By $T \gg 0$ we mean that $T \geq \varepsilon I$ for some $\varepsilon > 0$; in this case we say that T is *uniformly positive* (note that $(x_k) \mapsto (x_k/k)$ is is positive and one-to-one but not uniformly positive on $\ell^2(\mathbf{N})$). If $T^*T \gg 0$, then we say that T is *coercive* (or bounded from below). These relations (among others) are studied within the following lemma.

Lemma A.3.1 (Hilbert spaces) *Let H, U and Y be complex Hilbert spaces (much of this holds for real ones too).*

(a1) Each orthonormal base of H has the same cardinality; we denote this cardinality by $\dim H$.[1]

(a2) H is isomorphic to $\ell^2(\mathcal{E})$ iff $\mathcal{E}$ is a set of cardinality $\dim H$.

(a3) If $T \in \mathcal{B}(U,H)$, then $\dim \overline{\mathrm{Ran}(T)} \leq \dim U$.

(a4) The following are equivalent:

(i) $\dim U \leq \dim H$;
(ii) $T^*T \gg 0$ for some $T \in \mathcal{B}(U,H)$;
(iii) $T^*T = I$ for some $T \in \mathcal{B}(U,H)$;
(iv) some $S \in \mathcal{B}(H,U)$ is onto.

(a5) The following are equivalent:

(i) $\dim U = \dim H$;
(ii) $\mathcal{GB}(U,H) \neq \emptyset$;
(iii) $T^*T = I_U$ and $TT^* = I_H$ for some $T \in \mathcal{B}(U,H)$.

(a6) Let $\dim U \geq \dim H < \infty$. Then $T \in \mathcal{B}(U,H)$ is invertible iff $T^*T \gg 0$.

*(b1)***[T $\gg$ 0]** *Let $T \in \mathcal{B}(H)$ and $E \in \mathcal{GB}(H)$. Then $T \gg 0 \overset{\text{def}}{\Leftrightarrow} T \geq \varepsilon I$ for some $\varepsilon > 0 \Leftrightarrow T = T^*$ & $\sigma(T) \subset (0,\infty) \Leftrightarrow T = P^2$ for some $P \gg 0 \Leftrightarrow T = X^2$ for some $X \in \mathcal{GB}(H) \Leftrightarrow E^*TE \gg 0 \Leftrightarrow T \in \mathcal{GB}$ & $T \geq 0 \Leftrightarrow T \in \mathcal{GB}$ & $T^{-1} \geq 0$.*

If $T \geq \varepsilon I$, then $\varepsilon^{-1} \geq T^{-1} \geq \|T\|^{-1} I \gg 0$.

*(b2)***[T $>$ 0]** *We have $T > 0$ iff $T \geq 0$ and* $\mathrm{Ker}(T) = \{0\}$.

[1]We also use the standard notation $\dim B < \infty$ [$\dim B = \infty$] to mean that a vector space B is [in]finite-dimensional; if B is a Hilbert space, then, obviously, $\dim B < \infty$ iff $\dim B$ is finite.

*(b3)***[T ≥ 0]** *Let $T \in \mathcal{B}(H)$. Then $T \geq 0$ iff $T = T^*$ and $\sigma(T) \subset [0,\infty)$.*

(b4) Let $T \geq 0$. Then there is a unique $T^{1/2} \geq 0$ s.t. $(T^{1/2})^2 = T$. Moreover, $T^{1/2} \gg 0$ (resp. > 0) iff $T \gg 0$ (resp. > 0). Furthermore, $ST = TS \Leftrightarrow ST^{1/2} = T^{1/2}S$ for all $S \in \mathcal{GB}$.

(b5) Let $T_n \geq T_{n+1} \geq A$ for all $n \in \mathbf{N}$ for some $A = A^* \in \mathcal{B}(H)$. Then there is $T \geq A$ s.t. $T_n x \to Tx$ for all $x \in H$.

*(b6)***[T = T*]** *Let $T \in \mathcal{B}(H)$. Then $T = T^*$ iff $T^*T = TT^*$ and $\sigma(T) \subset \mathbf{R}$.*

(b7) Assume that $S,T \in \mathcal{GB}(H)$ and $ST = TS$. If $T \geq S \gg 0$, then $S^{-1} \geq T^{-1} \gg 0$; if $T > S \gg 0$, then $S^{-1} > T^{-1} \gg 0$; if $T \gg S \gg 0$, then $S^{-1} \gg T^{-1} \gg 0$.

(b8) If $T \in \mathcal{B}(H)$, $\varepsilon > 0$, and $|\langle Tx,x\rangle| \geq \varepsilon\|x\|^2$ for all $x \in H$, then $T \in \mathcal{GB}(H)$.

(b9) If $I \geq T \gg 0$, then $\|I - T\| < 1$.

*(c1)***[R*R ≫ 0]** *Let $R \in \mathcal{B}(U,H)$. The following are equivalent:*

- *(i) $\|Rx\| \geq \varepsilon\|x\|$ for all x for some $\varepsilon > 0$, i.e., R is* coercive *("uniformly bounded from below");*
- *(ii) $R^*R \geq \varepsilon^2 I$ for some $\varepsilon > 0$;*
- *(iii) $\mathrm{Ran}(R^*) = U$, i.e., R^* is* onto*;*
- *(iv) $\mathrm{Ran}(R)$ is closed and $\mathrm{Ker}(R) = \{0\}$;*
- *(v) there is $L \in \mathcal{B}(H,U)$ s.t. $LR = I$;*
- *(vi) there is some closed subspace $H_1 \subset H$ and some $S \in \mathcal{B}(H_1,H)$ s.t. $\begin{bmatrix} R & S \end{bmatrix} \in \mathcal{GB}(U \times H_1, H)$;*
- *(vii) $X^*X = R^*R$ for some $X \gg 0$;*
- *(viii) $P \gg 0 \Longrightarrow R^*PR \gg 0$;*
- *(viii') $\|G - I_H\| < 1 \Longrightarrow R^*GR \in \mathcal{GB}(U)$;*
- *(ix) $\|RX\|_{\mathcal{B}(*,H)} \geq \varepsilon\|X\|_{\mathcal{B}(*,*)}$ for some $\varepsilon > 0$ whenever X is linear (in particular, X is bounded iff RX is bounded);*
- *(x) $R \in \mathcal{GB}(U, \mathrm{Ran}(R))$;*
- *(xi) There is $\varepsilon > 0$ s.t. for all $x \in H \setminus \{0\}$ there is $y \in H \setminus \{0\}$ s.t. $\langle y, Rx\rangle \geq \varepsilon\|x\|\|y\|$.*

Moreover, if (i) holds, then the following hold:

- *(1) $\|(R^*R)^{-1}\| \leq \varepsilon^{-2}$ and $\|(R^*R)^{-1}R^*\| \leq \varepsilon^{-1}$*
- *(1') if $R \in \mathcal{GB}$, then $\|R^{-1}\| \leq \varepsilon^{-1}$.*
- *(2) For any $r > 0$ we can choose H_1 and S in (vi) so that $\|\begin{bmatrix} R & S \end{bmatrix}\| \leq \max\{\|R\|, r\}$ and $\|\begin{bmatrix} R & S \end{bmatrix}^{-1}\| \leq \max\{\varepsilon^{-1}, r^{-1}\}$. If also $\dim U < \infty$ and $H = U \times U'$, then we can take $S \in \mathcal{B}(U',H)$ without affecting the above norms.*
- *(3) If $R^*R = I$, then we can have $\begin{bmatrix} R & S \end{bmatrix}$ unitary (see (e3)).*

However $R^*R \gg 0$ & $G \in \mathcal{GB} \not\Rightarrow R^*GR \in \mathcal{GB}$ in general (cf. (viii'); even $\sigma(G) \subset [0,+\infty)$ is not sufficient).

(c2) Let $M \in \mathcal{B}(\mathbf{K}^m)$, $N \in \mathcal{B}(\mathbf{K}^m,\mathbf{K}^n)$. There is $L \in \mathcal{B}(\mathbf{K}^n,\mathbf{K}^m)$ such that $M + LN$ is invertible iff $M^*M + N^*N > 0$.

(c3) Let $R \in \mathcal{B}(U,H)$*. Then the following are equivalent:*

(i) $R \in \mathcal{GB}(U,H)$ *(i.e., R is invertible)*
(ii) R is injective and onto;
(iii) R is coercive and has a dense range;
(iv) $R^*R > 0$ *and* $RR^* \gg 0$*;*
(v) $R^*R \gg 0$ *and* $RR^* \gg 0$*;*
(vi) $R^* \in \mathcal{GB}(H,U)$.

If $\dim U = \dim H < \infty$*, then one more equivalent condition is that R is one-to-one (equivalently,* $\det R \neq 0$ *or R is onto).*

(c4) Let $R = R^* \in \mathcal{B}(H)$*. Then* $R \in \mathcal{GB}(H) \Leftrightarrow R^*R \gg 0$.

(c6) Let $R \in \mathcal{B}(H)$ *and* $R^*R = RR^*$*. Then* $\|R\| = \sup_{\|x\|=1} |\langle x, Rx\rangle| = \rho(R)$.

(c7) Let $R \in \mathcal{B}(H,U)$*. Then* $\mathrm{Ker}(R^*) = \mathrm{Ran}(R)^{\perp}$.

(c8) Let $R \in \mathcal{B}(H,U)$*. Then* $\mathrm{Ker}(R^*R) = \mathrm{Ker}(R)$ *and* $\overline{\mathrm{Ran}(R)} = \overline{\mathrm{Ran}(RR^*)}$.

(c9) Let $R \in \mathcal{B}(U,H)$*. Then the following are equivalent:*

(i) R is one-to-one (i.e., $\mathrm{Ker}(R) = \{0\}$*);*
(ii) R^* *has a dense range;*
(iii) $R^*R > 0$

If $\dim U < \infty$*, then (i)–(iii) hold iff* $R^*R \gg 0$ *(cf. (c1)).*

(c10) Let $A \subset H$*. Then* $A^{\perp}$ *is a closed subspace of H and* $(A^{\perp})^{\perp} = \overline{\mathrm{span}(A)}$.

(c11) $\sup_{|\alpha|=1} \|\alpha x + y\| \geq \|x\|^2 + \|y\|^2$ *(*$x, y \in H$*).*

In (d)–(f) we assume that $A \in \mathcal{B}(U,Y)$, $B \in \mathcal{B}(U,H)$*. In (d)–(e2) we assume that* $\gamma > 0$ *(the* TI_ω *claims concern cases when U and Y are* L^2_ω *spaces for some* $\omega \in \mathbf{R} \cup \{\infty\}$*).*

(d) We have $\gamma^2 - B^*B \geq 0 \Leftrightarrow \|B\| \leq \gamma$*. Analogously,* $\gamma^2 - B^*B \gg 0 \Leftrightarrow \|B\| < \gamma$.

(e1) $\|Ax\|^2 - \gamma^2\|Bx\|^2 \leq 0$ *for all* $x \in H \Leftrightarrow \gamma^2 B^*B \geq A^*A \Leftrightarrow A = LB$ *for some* $\|L\| \leq \gamma$
(if $A, B \in \mathrm{TI}_\omega$*, then we can take* $L \in \mathrm{TI}_\omega$*).*

(e2) The following are equivalent:

(i) $\gamma^2 B^*B - A^*A \gg 0$*;*
(i') $\|Ax\|^2 - \gamma^2\|Bx\|^2 \leq -\varepsilon\|x\|^2$ *for all* $x \in H$*;*
(i") $\gamma^2 B^*B \gg A^*A \geq 0$*;*
(ii) $A = LB$ *for some* $\|L\| < \gamma$ *and* $B^*B \gg 0$*;*
(iii) $B^*B \gg 0$ *and* $\|AB^{-1}_{\mathrm{left}}\| < \gamma]$*, where* $B^{-1}_{\mathrm{left}} := (B^*B)^{-1}B^*$.

If $B \in \mathcal{GB}$*, then (i) (hence (i)–(iii)) is also equivalent to* $\|AB^{-1}\| < \gamma$*. If (i) holds and* $\dim U \geq \dim H < \infty$*, then* $B \in \mathcal{GB}$.

In (ii) we can take $L = AB^{-1}_{\mathrm{left}}$*; this way we get* $L \in \mathrm{TI}_\omega$*, if* $B, A \in \mathrm{TI}_\omega$.

(e3) *Let* $B^*B = I$. *Then* $\overline{B} := \left[\begin{smallmatrix} B & 0 \\ 0 & I \end{smallmatrix}\right] \in \mathcal{GB}(U \times H_1, H)$ *is an* unitary extension *of* B, *where* $H_1 := \mathrm{Ran}(B)^\perp \subset H$.

(f) *Suppose that* $\dim U < \infty$. *Then* $A = LB$ *for some* $L \in \mathcal{B}(H,Y) \Leftrightarrow \mathrm{Ker}(B) \subset \mathrm{Ker}(A) \Leftrightarrow \mathrm{Ran}(A^*) \subset \mathrm{Ran}(B^*) \Leftrightarrow A^* = B^*L^*$ *for some* $L^* \in \mathcal{B}(F,H) \Leftrightarrow$ *(e1) holds for some* $\gamma > 0$.

(g1) *Let* $A, B \in \mathcal{B}(H)$. *If* $\langle x, Ax\rangle = \langle x, Bx\rangle$ *for all* $x \in H$, *then* $A = B$.

(g2) *Let* $A_t \in \mathcal{B}(H)$ $(t \in \mathbf{R}_+)$. *If* $\lim_{t\to+\infty}\langle x, A_t x\rangle = 0$ *for all* $x \in H$, *then* $\lim_{t\to+\infty}\langle y, A_t x\rangle = 0$ *for all* $x, y \in H$.

(g3) *Let* X *be a vector space. Let* $A, B : X \to U$, *and* $C, D : X \to Y$ *be linear. If* $\langle Ax, Bx\rangle = \langle Cx, Dx\rangle$ *for all* $x \in X$, *then* $\langle Ax, Bz\rangle = \langle Cx, Dz\rangle$ *for all* $x, z \in X$.

In (h1)–(k2) we assume that $x, y, x_n, y_n \in H$ *for all* $n \in \mathbf{Z}_+$, $T_n, T \in \mathcal{B}(H,Y)$ *for all* $n \in \mathbf{Z}_+$, *and* $S_n, S \in \mathcal{B}(Y,U)$ *for all* $n \in \mathbf{Z}_+$. *We write* $x_n \rightharpoonup x$ *if* $\langle x_n, z\rangle \to \langle x, z\rangle$ *for all* $z \in H$ *(i.e., if* $x_n \to x$ *weakly) (here always* $n, m, k \to +\infty$*).*

We say that T_n *converges to* T uniformly *(resp.* strongly, weakly*), if* $\|T_n - T\| \to 0$ *(resp.* $T_n x \to Tx$ *for all* $x \in H$, $T_n x \rightharpoonup Tx$ *for all* $x \in H$*).*

(h1) *If* $T_n \to T$ *weakly, then* $\{T_n\}$ *is uniformly bounded.*

(h2) *If* $T_n x$ *converges for all* $x \in H$, *then the limiting operator* $\widehat{T} : x \mapsto \lim_n T_n x$ *satisfies* $\widehat{T} \in \mathcal{B}(H,Y)$ *and* $\|\widehat{T}\| \le \liminf_n \|T_n\| < \infty$.

(h3) *If* $\langle y, T_n x\rangle_Y$ *converges for all* $x \in H$, $y \in Y$, *then there is* $\widehat{T} \in \mathcal{B}(H,Y)$ *s.t.* $T_n \to T$ *weakly,* $\|\widehat{T}\| \le \liminf_n \|T_n\| \le \sup_n \|T_n\| < \infty$.

(i1) *If* $x_n \rightharpoonup x$, *then* $\{x_n\}$ *is uniformly bounded and* $\|x\| \le \liminf_n \|x_n\| < \infty$.

(i2) *If* $x_n \to x$ *and* $y_m \rightharpoonup y$, *then* $\langle x_n, y_m\rangle \to \langle x, y\rangle$.

(i3) *If* $\{\langle x_n, y\rangle\}$ *converges for all* $y \in H$, *then* $\{x_n\}$ *converges weakly.*

(i4) *If(f)* $x_n \rightharpoonup x$ *and* $\|x_n\| \to \|x\|$, *then* $x_n \to x$.

(j1) $S_n \to S$ *weakly iff* $S_n^* \to S^*$ *weakly.*

(j2) *If* $S_n \to S$ *weakly and* $T_n \to T$ *strongly, then* $S_n T_n \to ST$ *weakly (but even for* $U = H$ *we may have* $T_n S_n \not\to TS$*).*

(j3) *If* $S_n \to S$ *strongly,* $T_m \to T$ *strongly, and* $x_k \to x$ *strongly, then* $S_n T_m \to ST$ *strongly and* $S_n T_m x_k \to STx$ *strongly.*

(j4) *Let* $T_n \in \mathcal{GB}(H,Y)$ *for all* n, $T_n \to T$ *strongly and* $T_n^{-1} \to P$ *strongly for some* $P \in \mathcal{B}(Y,H)$. *Then* $T \in \mathcal{GB}(H,Y)$ *and* $T^{-1} = P$.

(j5) *Conversely, let* $T_n \in \mathcal{GB}(H,Y)$ *for all* n, *let* $T_n \to T$ *strongly, and let* $\{T_n^{-1}\}$ *be uniformly bounded. Then* $T \in \mathcal{GB}$ *iff* $T_n^{-1} \to P$ *strongly for some* $P \in \mathcal{B}(Y,H)$.

(k1) *Let* $\dim Y < \infty$ *and* $T \in \mathcal{B}(U,Y)$. *Then* $T_n \to T$ *strongly iff* $T_n \to T$ *uniformly.*

(k2) *Let* $\dim U < \infty$ *and* $T \in \mathcal{B}(U,Y)$. *Then* $T_n \to T$ *strongly iff* $T_n \to T$ *weakly.*

In (p1)–(q) we assume that $\left[\begin{smallmatrix} A & B \\ B^* & D \end{smallmatrix}\right] \in \mathcal{B}(H_1 \times H_2)$, *where* H_1 *and* H_2 *are Hilbert spaces.*

(p1) Let $AX = B$. Then $\begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \geq 0$ iff $A \geq 0$ & $D - X^*AX \geq 0$. Moreover, $\begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \gg 0$ iff $A \gg 0$ & $D - X^*AX \gg 0$.

(p2) Let $\exists A^{-1}$. Then $\begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \geq 0$ iff $A \geq 0$ & $D - B^*A^{-1}B \geq 0$.

(p3) Let $\dim H_2 < \infty$. Then $\begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \geq 0$ iff for some $X \in \mathcal{B}(H_2, H_1)$ we have $AX = B$, $D - X^*AX \geq 0$ and $A \geq 0$.

(p4) We have $\begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \gg 0$ iff $A \gg 0$ & $D - B^*A^{-1}B \gg 0$.

(q) Let $P \in \mathcal{B}(U, H_1)$, $Q \in \mathcal{B}(U, H_2)$ and $D \leq 0$. If $T := \begin{bmatrix} P^* & Q^* \end{bmatrix} \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \begin{bmatrix} P \\ Q \end{bmatrix} > 0$ $[\gg 0]$, then $P^*P > 0$ $[\gg 0]$.

If $\varepsilon > 0$ and $T \geq \varepsilon I$, then $P^*P \geq \delta I$, where δ only depends on ε, $\|Q\|$, $\|A\|$ and $\|B\|$.

If $\varepsilon > 0$, $D \leq -\varepsilon I$ and $T \geq 0$, then $P^*P \geq \delta Q^*Q$ and $Q = LP$ for some $L \in \mathcal{B}(H_1, H_2)$ s.t. $\|L\|_{\mathcal{B}} \leq \delta^{-1/2}$, where δ only depends on ε, $\|A\|$ and $\|B\|$.

(s) Let $T \in \mathcal{B}(U, Y)$, let $U_1 \subset U$ be a finite-dimensional subspace, and let $\dim T[U_1] = \dim U_1 < \infty$. Then T is of the form $T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \in \mathcal{B}(U_1 \times U_1^\perp, Y_1 \times Y_1^\perp)$, where $Y_1 := T[U_1]$, and $T_{11} \in \mathcal{GB}(U_1, Y_1)$. Moreover, T is (resp. right-, left-)invertible iff T_{22} is (resp. right-, left-)invertible.

(P) Let H_1 be a closed subspace of H. Let P be the orthogonal projection of H onto $H_1^\perp$. Let $x \in H$. Then Px is the unique element of minimum norm on $x + H_1$.

As a curiosity, we remark that (c1) demonstrates that finding a left inverse (v), a complement (vi) or a spectral factor (vii) is easy, if we do not have to worry about causality (in contrast to the Corona and spectral factorization theorems).

Results concerning convergence of sequences (such as Monotone Convergence Theorem or (h1)–(k2) above) are applicable for limits of functions between any metric (or "first countable") spaces (recall from Theorem 4.2 of [Rud76] that $f(t) \to q$ as $t \to T$ iff $f(t_n) \to q$ for each sequence $\{t_n\}$ converging to T (with $t_n \neq T$ for all n)). We shall use this fact without further mention.

Proof of Lemma A.3.1: (We often apply Banach adjoint results for Hilbert adjoints, see Remark A.3.20 for the justification.)

(a1)&(a2) These follow from [Rud86, 4.19].

(a3) If $\dim U < \infty$, these claims are easy to prove, so assume $\dim U = \infty$ and let $\{u_a\}_{a \in A}$ be an orthonormal base of U. Then the cardinality of $Q := \{\sum_{k=1}^n q_k u_{a_k} \mid q_k \in \mathbf{Q} + i\mathbf{Q},\ a_k \in A \text{ for all } k\}$ is $A =: \dim U$ and Q is dense in U (obviously the cardinality of any dense set is at least that of A).

Obviously, TQ is dense in $\overline{\mathrm{Ran}(T)}$, hence $\dim \overline{\mathrm{Ran}(T)}$ is at the cardinality of Q, i.e., $\dim U$.

(a4) "(iv)⇒(i)" follows from (a3). "(iv)⇔(ii)" is [Rud73, Theorem 4.15]. "(ii)⇒(iii)" Follows by taking $T' := T(T^*T)^{-1/2}$ (cf. (b)). "(iii)⇒(ii)" is trivial. "(i)⇒(iv)" Let $\{u_a\}_{a \in A}$ and $\{h_a\}_{a \in B}$ be bases of U and H, respectively, and $A \subset B$. Set $S \sum_{a \in B} \alpha_a h_a := \sum_{a \in A} \alpha_a u_a$.

(a5) "(i)⇒(iii)&(ii)" Construct S as in "(i)⇒(iv)" above, with $A = B$. The converses follow from (a2).

(a6) Let $T^*T \gg 0$. From (a4) we get that $\dim U = \dim H$. Because T is injective, it is invertible. The converse is trivial.

(b1)&(b3)&(b6) These follow from [Rud73, Chapter 12] and straightforward computations.

(b2) Let $T \geq 0$. Assume that $\langle x, Tx\rangle = 0$. Then $\|T^{1/2}x\|^2 = \langle T^{1/2}x, T^{1/2}x\rangle = 0$ (see (b4)), hence then $Tx = T^{1/2}T^{1/2}x = 0$. Consequently, $x \in \mathrm{Ker}(T)$; the converse is trivial.

(b4) This follows from Theorem 6.2.10 of [Aupetit] (alternatively, from Chapters 11–12 of [Rud73]) and (c3)&(c9).

(b5) (The convergence need not be uniform; e.g., $\pi_{[n,n+1)} \geq \pi_{[n+1,\infty)} \to 0$ strongly but not uniformly on $\ell^2(\mathbf{N})$.)

W.l.o.g. we assume that $A = 0$ (use $T_n \mapsto T_n - A$). Set $S(x) := \lim_n \langle x, T_n x\rangle \geq 0$ $(x \in H)$. Now $S(x+y) - S(x) - S(y) = \lim_n 2\,\mathrm{Re}\langle y, T_n x\rangle$ exists for all $x, y \in H$. Apply this to iy to see that $R(y,x) := \lim_n \langle y, T_n x\rangle$ exists for all $x, y \in H$. Obviously, R is sesquilinear and $\|R(y,x)\| \leq \|T_1\|((\|x\| + \|y\|) + \|x\| + \|y\|)$, hence $R(y,x) = \langle y, Tx\rangle$ for some $T \in \mathcal{B}(H)$, by Theorem 12.8 of [Rud73].

Now $\|(T_n - T)^{1/2}x\|^2 = \langle x, (T_n - T)x\rangle \to 0$, hence $\|(T_n - T)x\| \to 0$, for all $x \in H$, since $0 \leq T_n - T \in \mathcal{B}(H)$. Thus, $T_n x \to Tx$ strongly.

(b7) By (b1), $S^{-1}, T^{-1} \gg 0$. Apply twice (b4) to obtain $S^{1/2}T^{1/2} = T^{1/2}S^{1/2}$. It follows that $T^{-1/2}S^{1/2} = S^{1/2}T^{-1/2}$. Let $x \in H$ be arbitrary, and set $y := S^{-1/2}x$. Then

$$\begin{aligned} \langle x, T^{-1}x\rangle &= \langle S^{1/2}y, T^{-1}S^{1/2}y\rangle = \langle T^{-1/2}y, ST^{-1/2}y\rangle \\ &\leq \langle T^{-1/2}y, TT^{-1/2}y\rangle = \langle y, y\rangle = \langle x, S^{-1}x\rangle. \end{aligned} \tag{A.18}$$

Thus, $T^{-1} \leq S^{-1}$. Analogously, if $T > S$, then $T^{-1} < S^{-1}$; if $T \gg S$, then $T^{-1} \ll S^{-1}$.

(b8) Now $\|Tx\| \geq \varepsilon\|x\|$ and $\|T^*x\| \geq \varepsilon\|x\|$ for all $x \in H$, hence this follows from (c3)(i)&(iv). (The converse of (b8) is obviously not true.)

(b9) If $I \geq T \gg \varepsilon I$, then $\langle (I-T)x, x\rangle \leq (1-\varepsilon)\|x\|^2$ for all $x \in H$, hence $\|I - T\| \leq 1 - \varepsilon$, by (c6).

(c1) The equivalence of (i)–(viii) is obtained as follows:

"(i)⇔(ii)": $\|Rx\|^2 = \langle x, R^*Rx\rangle$. "(i)⇔(iii)": See [Rud73, 4.15]. "(i)⇒(iv)": Clearly $\mathrm{Ker}(R) = \{0\}$. If $\{Rx_n\}$ is a Cauchy-sequence, then so is $\{x_n\}$. "(iv)⇒(x)": See [Rud73, 2.12b]. "(x)⇒(i)": See [Rud73, 2.12c]. "(i)⇒(v)": Take $L := (R^*R)^{-1}R^*$ (note: $\|L\| \leq \|R\|/\varepsilon$). "(v)⇒(i)": $\|Rx\| \geq \|x\|/\|L\|$ for all $x \in U$. "(vii)⇔(ii)": See [Rud73, 12.33] ($X = X^* \gg 0$ is unique). "(viii)⇒(ii)": Take $P = I$. "(ii)⇒(viii)": $\langle x, R^*PRx\rangle \geq \varepsilon_P\|Rx\| \geq \varepsilon_P\varepsilon_R\|x\|$ for all x. "(i)⇒(xi)": Take $y = Rx$. "(xi)⇒(i)": Obviously.

"(vi)⇒(v)": If $\begin{bmatrix} L \\ M \end{bmatrix}\begin{bmatrix} R & S \end{bmatrix} = I_{H\times H_1}$, then $LR = I_H$.

"(iii)⇒(vi) & (2)": Let $H_2 := \mathrm{Ran}(R)$, $H_1 := H_2^\perp$, and write $R =: \begin{bmatrix} T \\ 0 \end{bmatrix} \in \mathcal{B}(U, H_2 \times H_1)$ (i.e., $T := P_{H_2}R$), so that $T \in \mathcal{GB}(U, H_2)$, because it is 1-1 and onto.

Choose some $r > 0$. Then $S = \begin{bmatrix} 0 \\ rI \end{bmatrix}$ complements R, and the inverse is $\begin{bmatrix} T^{-1} & 0 \\ 0 & r^{-1}I \end{bmatrix}$. The inequalities in (2) follow as in (C1).

If $\dim U < \infty$ and $H = U \times U'$, then $\dim H_2 = \dim U$ implies that $\dim H_1 = \dim U'$, so we may choose some isometric isomorphism $I' \in \mathcal{GB}(U', H_1)$ and replace S by SI' to obtain the rest of (2).

”(ix)⇔(i)”: Let H' be a Hilbert space, let $X : H' \to H$ be linear, and let (i) hold. Then $\|RXy\| \geq \varepsilon \|Xy\|$ for all $y \in H'$, hence then $\|RX\| \geq \varepsilon \|X\|$. For the converse, let $x \in H$ be given, take some $H' \neq \{0\}$, $y_0 \in H' \setminus \{0\}$, $\Lambda \in H'^*$ with $\Lambda y_0 = \|y_0\|$, $\|\Lambda\| = 1$ [Rud73, 3.3Cor], and define $Xy := x\Lambda y$. Then $\|X\| = \|x\|$, $\|RX\| = \|Rx\|\|\Lambda\| = \|Rx\|$, hence then $\|Rx\| \geq \varepsilon \|X\| = \varepsilon \|x\|$, i.e., (i) holds.

“(viii’)⇒(ii)”: Set $G := I_H$. “(x)⇒(viii’)”: Set $G' := P^*GP = I - P^*FP$, where $F := I - G$ and $P \in \mathcal{GB}(H, H_R)$ is the orthogonal projection onto $H_R := \mathrm{Ran}(R)$. Since $\|P^*FP\| \leq \|F\| < 1$, we have $G'_{|(}H_R) \in \mathcal{GB}(H_R)$, hence $R^*GR = R^*G'R \in \mathcal{GB}(U)$ (since $R \in \mathcal{GB}(H, H_R)$).

(Note: condition $\sigma(G) \subset \mathbf{R}^+$ would not be sufficient in (viii’): if $G' = \left[\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right]$, $G = \left[\begin{smallmatrix} G' & 0 \\ 0 & I \end{smallmatrix}\right] \in \mathcal{B}(H)$ and $H := \ell^2 =: U$, then $\sigma(G) = \{1\}$ but $R^*GR(1,0,0,\ldots) = 0$, where $R : (a,b,c,\ldots) \mapsto (a,-a,b,c,\ldots)$, so that $R^*R \geq I \gg 0$.)

(1)&(1’) From (ii) and (b1) we obtain $(R^*R)^{-1} \leq \varepsilon^{-2}$. Now $\langle R(R^*R)^{-1}x, R(R^*R)^{-1}x\rangle = \langle (R^*R)^{-1}x, x\rangle \leq \varepsilon^{-2}\|x\|^2$, hence $\|R(R^*R)^{-1}\| \leq \varepsilon^{-1}$. If $R \in \mathcal{GB}$, then $(R^*R)^{-1}R^* = R^{-1}$.

(2) This was shown in “(iii)⇒(vi) & (2)”.

(3) Let $r = 1$ in “(iii)⇒(vi) & (2)”.

The final remark was justified before “(1)&(1’)” above.

(c2) By (c1), $M^*M + N^*N > 0$ is necessary (of course, $\mathrm{Ker}(\left[\begin{smallmatrix} M \\ N \end{smallmatrix}\right]) = \{0\}$). Assume that $M^*M + N^*N > 0$. Set $U_2 := \mathrm{Ker}(M)$, $U_1 := U_2^\perp$. Let $x_1, \ldots, x_k$ be a base of U_2, let $x_{k+1}, \ldots, x_m$ be a base of U_1, and let $y_1, \ldots, y_k$ be a base of $(MU_1)^\perp$. Let $S\sum_{j=1}^m \alpha_j x_j := \sum_{j=1}^k \alpha_j y_j$. Then $M + S$ maps $\{x_1, \ldots, x_m\}$ to a base of $\mathbf{K}^m$, hence it is invertible. Now $\mathrm{Ker}(N) \subset U_1 = U_2^\perp = \mathrm{Ker}(S)$, hence, by (f), $S = LN$ for some $L \in \mathcal{B}(\mathbf{K}^n, \mathbf{K}^m)$.

(c3) 1° *(i)–(vi):* This equivalence follows easily from (c1).

2° *(i)⇔(vii):* If $R \in \mathcal{GB}(H)$, we can take $y = Rx$. Conversely, $\langle y, Rx\rangle \geq \varepsilon\|x\|\|y\|$ implies that $\|Rx\| \geq \varepsilon\|x\|$, hence $R^*R \gg 0$, by (c1), and analogously $RR^* \gg 0$, so that $R \in \mathcal{GB}(H)$, by “(v)⇔(i)”.

3° *The last claim:* This is given in almost any matrix calculus textbook.

(c4) This follows from (c3).

(c6)&(c7) These are Theorems 12.25, 11.28(b) and 12.10 of [Rud73] (slightly modified).

(c8) If $x \in \mathrm{Ker}(R^*R)$, then $\|Rx\|^2 = \langle x, R^*Rx\rangle = 0$, hence then $x \in \mathrm{Ker}(R)$. Thus, $\mathrm{Ker}(R^*R) = \mathrm{Ker}(R)$. Consequently, $\mathrm{Ran}(R^*R)^\perp = \mathrm{Ran}(R^*)^\perp$, by (c7), hence $\overline{\mathrm{Ran}(R^*)} = (\mathrm{Ran}(R^*)^\perp)^\perp = \overline{\mathrm{Ran}(RR^*)}$. Because $R^{**} = R$, also the latter claim holds.

(c9) By (c8), we have (i)⇔(ii). But (i) holds iff $\|Rx\|^2 = \langle Rx, Rx\rangle > 0$ for all $x \in U$, i.e., iff (iii) holds.

(c10) This is an easy exercise.

(c11) This is obvious (and this is not true for, e.g., $H = \mathrm{L}^\infty(\mathbf{R})$).

(d) The first claim follows from $\gamma^2 - B^*B \geq 0 \Leftrightarrow \gamma^2\|x\|^2 - \|Bx\|^2 \geq 0$ for all $x \in H$. The second follows by replacing γ by some $\gamma - \varepsilon$.

(e1) (This is from [RR, Lemma 1.14].) Assume $\|Ax\|^2 \le \gamma^2\|Bx\|^2$. Then we can define $L_0 \in \mathcal{B}(H_1,Y)$ (where $H_1 := \overline{B[U]} \subset H$) by $L_0(Bx) := Ax$ to get $\|L_0\| \le \gamma$. Let P be the orthonormal projection $H \to H_1$ and define $L := L_0P \in \mathcal{B}(H,Y)$ to get $A = LB$ and $\|L\| = \|L_0\| \le \gamma$. The other direction is straightforward.

If $B,A \in \mathrm{TI}_\omega$, then $\tau_r H_1 = H_1$ for all $r \in \mathbf{R}$ ($\tau_r Bx = B\tau_r x$) and $L_0 \in \mathrm{TI}_\omega$, hence then $L \in \mathrm{TI}_\infty$ too (if $P : H \to H_1$ is orthonormal, then $L = L_0P \in \mathrm{TI}_\omega$) (N.B. Even if $A,B \in \mathrm{TIC}_\omega$, we may have to take $L \notin \mathrm{TIC}_\infty$, e.g., $A = I \wedge B = \tau_{-1} \implies L = \tau_1$. Note also that we may have that $B = 0 = C = L$, i.e., that $\nexists B_{\mathrm{left}}^{-1}$).

(e2) "(i)⇔(i')⇔(i")" and "(iii)⇒(ii)" are obvious. "(ii)⇒(i)" follows from $A^*A = B^*L^*LB \ll B^*B\gamma^2$, which holds by (c1)(viii).

"(i')⇒(iii)": Assume (i') and set $S := (B^*B)^{-1}B^*$. Then $\|ASy\|^2 - \gamma^2\|y\|^2 \le -\varepsilon\|y\|^2$ for all $y \in \mathrm{Ran}(B)$, i.e., $\|ASy\| \le (\gamma - \varepsilon')\|y\|$ for all $y \in \mathrm{Ran}(B)$. Because $Sy = 0$ for all $y \in \mathrm{Ran}(B)^\perp = \mathrm{Ker}(B^*)$ and $\mathrm{Ran}(B)$ is closed, this holds for all $y \in H$,

If $B \in \mathcal{GB}$, then $B_{\mathrm{left}}^{-1} = L = B^{-1}$. If (i) holds and $\dim U \ge \dim H < \infty$, then $B \in \mathcal{GB}$, by (a6).

(e3) See the proof of (c1)(3).

(f) The proof goes as that of (e1), except that L_0 is necessarily continuous, because it is linear and $\dim H_1 < \infty$ (note that now $\mathrm{Ker}(B) = \mathrm{Ran}(B^*)^\perp$).

(g2) Now $2\,\mathrm{Re}\langle y, A_t x\rangle = \langle (x+y), A_t(x+y)\rangle - \langle x, A_t x\rangle - \langle y, A_t y\rangle \to 0$. If $\mathbf{K} = \mathbf{C}$, we apply this to y and iy to obtain (g2).

(g1)&(g3) The proofs are analogous to that of (g2).

(h1) This (i.e., that $\|T_n\| \le M$ for all n for some $M < \infty$) follows from Theorems 2.6 and 3.18 of [Rud73].

(h2) By Theorem 2.8 of [Rud73], $\widehat{T} \in \mathcal{B}(H,U)$. The bound is obtained from $\|\widehat{T}x\| \le \liminf_n \|T_n\|\|x\| = (\liminf_n \|T_n\|)\|x\|$ for all $x \in H$. By (h1), $\liminf_{n\to+\infty} \|T_n\| < \infty$.

(h3) By (i3), $\widehat{T}x := \text{w-lim}_n T_n x \in Y$ exists for all $x \in H$. Obviously, $\widehat{T} : H \to Y$ is linear. By Theorem 3.18 of [Rud73], $\{T_n x\}$ is bounded for each $x \in H$, hence $M := \sup_n \|T_n\| < \infty$, by Theorem 2.6 of [Rud73]. The inequalities follow from this, hence $\widehat{T} \in \mathcal{B}(H,Y)$ and $T_n \rightharpoonup T$.

(i0) *If $T_n \to T$ strongly and $x_m \to x$, then $T_n x_m \to Tx$*: we have $T_n x_m - Tx = T_n(x_m - x) + (T_n - T)x \to 0$, because $\{T_n\}$ is uniformly bounded, by (h1).

(i1)&(i3) Set $T_n := \langle \cdot, x_n\rangle$ and apply (h1)&(h2) (recall that any $T \in \mathcal{B}(H,\mathbf{K})$ is of form $\langle \cdot, x\rangle_H$ for some $x \in H$).

(i2) As above, $\langle x_n, y_n\rangle - \langle x, y\rangle = \langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle \to 0$, because $\{y_n\}$ is uniformly bounded (alternatively, apply (i0) for $T_n := \langle \cdot, y_n\rangle$, $T := \langle \cdot, y\rangle \in \mathcal{B}(H,\mathbf{K})$).

(i4) This holds because $\langle x - x_n, x - x_n\rangle = \|x\|^2 + \|x_n\|^2 - 2\,\mathrm{Re}\langle x, x_n\rangle \to 0$.

(j1) This is obvious.

(j2) Now $\langle z, S_n T_n x\rangle = \langle S_n^* z, T_n x\rangle \to \langle S^* z, Tx\rangle = \langle z, STx\rangle$ for all x, z, by (i2) and (j1). (But $T_n S_n = I \to I \ne 0 = TS$ for $T_n := \tau(n) \to 0$, $S_n := \tau(-n) \to 0$ on $\ell^2(\mathbf{N}) =: H =: Y =: U$.)

(j3) The first claim follows from (i0), the second from the first and (i0).

(j4) By (j3), $x = T_n^{-1}T_n x \to PTx$ for all $x \in H$, and $y = T_n T_n^{-1} y \to TPy$ for all $y \in Y$, hence $\exists T^{-1} = P$.

(If $T_n \to T$ strongly and $T_n^{-1} \to P$ weakly, then T and P need not be invertible (and if either is, then both are (by (j2), we have $PT = I$, as above, hence $P, T \in \mathcal{GB}$) and $T_n^{-1} \to P$ strongly, by (h1) and (j5)). An example of this is given by $T_n x := (x_{n+1}, x_0, x_1, \cdots, x_n, x_{n+2}, x_{n+3}, \cdots) \in \mathcal{B}(\ell^2(\mathbf{N}))$.)

(j5) "Only if" is given in (j4), so we only prove "if": Let $T \in \mathcal{GB}$. Then $T_n^{-1}y - T^{-1}y = T_n^{-1}(T - T_n)T^{-1}y \to 0$ for all $y \in Y$.

(In fact, if $T, T_n \in \mathcal{GB}$ and $T_n \to T$ strongly, then $T_n^{-1} \to T^{-1}$ iff $\{T_n^{-1}\}$ is uniformly bounded (this is not the case in general (take $T_n = I - 2^{-1/n^3}\tau(n)$ to obtain $\|T_n^{-1}(j^{-2/3})_{j \in \mathbf{Z}}\| \to +\infty$)), by (j5) and (h1).)

(k1) Cf. the proof of (k2).

(k2) Set $m := \dim U$, so that $T = \begin{bmatrix} T^1 & T^2 & \cdots & T^m \end{bmatrix}$. Let $x \in U$. Then $T_n x \to Tx$ (strongly) $\Leftrightarrow T_n^k x \to T^k x$ for all $k = 1, \ldots, m \Leftrightarrow T_n x \rightharpoonup Tx$.

(p1) $\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^* \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & D - X^*AX \end{bmatrix}$, which is ≥ 0 [$\gg 0$] iff $A, D - X^*AX \geq 0$ [$\gg 0$].

(p2) Set $X := A^{-1}B$ in (p1) to obtain $D - X^*AX = D - B^*(A^{-1})^*AA^{-1}B = D - B^*A^{-1}B$.

(p3) "$\Longleftarrow$": from (p1). "$\Longrightarrow$": Obviously $\operatorname{Ker}(A) \subset \operatorname{Ker}(B^*)$, hence $AX = B$ for some linear (hence bounded, because $\dim H_2 < \infty$) X. The rest follows from (p1).

(p4) See the proofs of (p2) and (p1).

(q) (W.l.o.g. we assume that $U \neq \{0\}$.) 1° *Case* $T > 0$*:* Set $M := \max\{\|Q\|, \|A\|, \|B\|, 1\}$. If $x \in B$ is s.t. $Px = 0$, then $\langle x, Tx \rangle = \langle Qx, DQx \rangle \leq 0$; thus, $T > 0$ implies that $\operatorname{Ker}(P) = \{0\}$, i.e., that $P^*P > 0$.

2° *Case* $T \geq \varepsilon I$*:* Assume that $T \geq \varepsilon I$, $\varepsilon > 0$. Choose $\delta > 0$ s.t. $M\delta^2 + 2\delta M^2 \leq \varepsilon$. If $x \in B$ is s.t. $\|Px\| < \delta$, then $\langle x, Tx \rangle < \delta^2 M + 2\delta M^2 + 0 \leq \varepsilon$ (since $T = P^*AP + P^*BQ + Q^*B^*P + Q^*D^*Q$), a contradiction; hence $P^*P \geq \delta^2 I$.

3° *Case* $D \leq -\varepsilon I$*:* Assume that $D \leq -\varepsilon I$. Then, for any $x \in U$ s.t. $\|x\|_U = 1$, we must have

$$0 \leq |\langle x, Tx \rangle| \leq Mp^2 + 2Mpq - \varepsilon q^2, \tag{A.19}$$

where $p := \|Px\|$, $q := \|Qx\|$, $M := \max\{\|A\|, \|B\|\}$. Thus, if $p = 0$, then $q = 0$, and if $p > 0$, then $0 \leq M + 2Mr - \varepsilon r^2$, where $r := q/p \geq 0$, so that $r \leq \delta^{-1/2}$ for some $\delta := \delta_{M,\varepsilon} > 0$.

Therefore, $p \geq q/r \geq \delta^{1/2} q$, or $\langle Px, Px \rangle \geq \delta \langle Qx, Qx \rangle$, for any $x \in U$. The claim on L follows from (e1).

(s) $\dim T[U_1] = \dim U_1$ means that T is coercive on U_1, i.e., that $T_{11}^* T_{11} \gg 0$. If P_2 is the orthogonal projection of Y onto $Y_1^\perp$, then $P_2 T = 0$ on U_1, hence T is of the form claimed in the lemma. The last claim follows from Lemma A.1.1(b).

(P) The minimum of $\|x - z\|$ ($z \in H_1$) is obtained at $z = (x - Px)$, by Theorem 4.11 of [Rud86]. □

We now show that each $T = T^* \in \mathcal{B}(H)$ can be written as $T = T_+ - T_-$, where $T_\pm \geq 0$:

Lemma A.3.2 ($\mathbf{f}(\mathbf{T})$) *Let $T \in \mathcal{B}(H)$ and $T^*T = TT^*$, where H is a complex Hilbert space. Then there is a (canonical) algebra homomorphism $f \mapsto f(T)$. from the set of all bounded Borel functions on $\sigma(T)$ to $\mathcal{B}(H)$. Moreover, this algebra homomorphism satisfies (here f and g are bounded Borel functions):*

(a) $1(T) = I$, $I(T) = T$, $f(T)g(T) = fg(T)$ *and* $\bar{f}(T) = f(T)^*$, *hence* $f(T)^* f(T) = f(T) f(T)^* = |f|^2(T)$.

(b) $\|f(T)\| \le \sup|f|$ *(if* $f \in \mathcal{C}(\sigma(T))$*, then* $\|f(T)\| = \sup|f|$*).*

(c1) If $f \in \mathcal{C}(\sigma(T))$*, then* $\sigma(f(T)) = f(\sigma(T))$.

(c2) We have $\mathrm{Ker}(T) = \chi_{\{0\}}(T)$.

(d) If $S \in \mathcal{B}(H)$ *and* $ST = TS$*, then* $Sf(T) = f(T)S$.

(e1) If $fg = 0$*, then there are closed subspaces* $H_\pm \subset H$ *s.t.* $\mathrm{Ker}(f(T)), \mathrm{Ran}(f(T)) \subset H_+$, $\mathrm{Ker}(g(T)), \mathrm{Ran}(g(T)) \subset H_-$, $H = H_+ + H_-$ *and* $H_+ \cap H_- = \emptyset$.

(e2) If $T = T^*$*, then* $f(T) = f(T)^*$*, and we can have* $H_+ = H_-^\perp$ *in (e1).*

(f1) If $T = T^*$*, then there are orthogonal projections* $P_+, P_0, P_- \in \mathcal{B}(H)$ *s.t.* $I = P_+ + P_0 + P_-$, $T_\pm := TP_\pm = P_\pm T = P_\pm T P_\pm$ *satisfy* $\pm T_\pm > 0$ *on* $H_\pm := \mathrm{Ran}(P_\pm)$, $T = T_+ + T_-$*. Consequently,* $H_0 := \mathrm{Ran}(P_0) = \mathrm{Ker}(T)$*, and*

$$T = \begin{bmatrix} \widetilde{P}_+ & \widetilde{P}_0 & \widetilde{P}_- \end{bmatrix} \begin{bmatrix} \widetilde{T}_+ & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \widetilde{T}_- \end{bmatrix} \begin{bmatrix} \widetilde{P}_+ \\ \widetilde{P}_0 \\ \widetilde{P}_- \end{bmatrix}, \qquad \text{(A.20)}$$

where $\widetilde{P}_\pm \in \mathcal{B}(H, H_\pm)$, $\widetilde{P}_0 \in \mathcal{B}(H, H_0)$ *and* $\widetilde{T}_\pm \in \mathcal{B}(H_\pm)$ *have the same values on their domains as the corresponding operators without tildes (note that* $J := \begin{bmatrix} P_+ & P_0 & P_- \end{bmatrix} = J^{-*} \in \mathcal{GB}(H)$*, i.e., J is unitary).*

(f2) If $T = T^* \in \mathcal{GB}(H)$*, then* $P_0 = 0$ *and* $\widetilde{T}_+ \gg 0 \gg \widetilde{T}_-$ *in (f1); thus, then* $T = \begin{bmatrix} \widetilde{P}_+ & \widetilde{P}_- \end{bmatrix} \begin{bmatrix} \widetilde{T}_+ & 0 \\ 0 & \widetilde{T}_- \end{bmatrix} \begin{bmatrix} \widetilde{P}_+ \\ \widetilde{P}_- \end{bmatrix}$

Note that if $H = \mathrm{L}^2(\mathbf{R}; U)$ and $T \in \mathrm{TI}(U)$, then $f(T) \in \mathrm{TI}(U)$, by (d).

Proof: We obtain the initial claims, (a) and (b) from Theorem 12.24 of [Rud73]. Claim (c1) follows from Theorems 13.27(c) and 12.22(b) of [Rud73], (c2) from Theorem 12.29(a) of [Rud73], and (d) from 12.24 of [Rud73].

(e1) Set $E := \{z \in \sigma(T) \,|\, f(z) \neq 0\}$, $P_+ := \chi_E(T)$, $P_- := I - P_+ = \chi_{E^c}(T)$, $H_\pm := \mathrm{Ran}(P_\pm)$. Then $P_\pm = P_\pm^2$ (i.e., $P_\pm$ are projections), $P_+P_- = 0 = P_-P_+$, $P_+ + P_- = I$, $f = P_+ f P_+$ and $g = P_- g P_-$, by (a). Therefore, $\mathrm{Ran}(P_\pm) = \mathrm{Ker}(P_\mp)$, $H_+ + H_- = H$, $H_+ \cap H_- = \emptyset$ and $H_\pm$ are closed (see Section 5.15 of [Rud73]).

(e2) By Lemma A.3.1(b6), we have $\sigma(T) \subset \mathbf{R}$, hence $f = \bar{f}$, so that $f(T)^* = f(T)$. In (e1) we have $P_\pm = P_\pm^*$ (i.e., $P_\pm$ is orthogonal), hence $H_+ = H_-^\perp$, by Theorem 12.14(c) of [Rud73].

(f1) (The statement means that $\widetilde{T}_\pm x := T_+ x := Tx$ for all $x \in H_+$ etc.) Set $P_\pm := \chi_{\pm(0,+\infty)}(T)$, $P_0 := \chi_{\{0\}}(T)$, so that $T_\pm = (s \mapsto s\chi_{\pm(0,+\infty)}(s))(T)$ etc. Now we obtain the claims easily from (c1), (c2), (e1) and (e2) (e.g., by (c1) we have $\pm T_\pm \geq 0$ on H and $0 \neq Tx = P_\pm T_\pm P_\pm x$ for $x \in H_\pm$, hence $\pm T_\pm > 0$ on $H_\pm$).

(f2) If $T \in \mathcal{GB}$, then $J^{-1}TJ \in \mathcal{GB}$, hence then $\widetilde{T}_\pm \in \mathcal{GB}(H_\pm)$, hence $\pm\widetilde{T}_\pm \gg 0$, by Lemma A.3.1(b1). □

By a *Banach algebra* we mean a complex Banach space A equipped with a multiplication $A \times A \to A$ (not necessarily commutative) and possessing a *unit* $I = I_A$ that satisfy $xI = Ix = x$, $x(yz) = (xy)z$, $(x+y)z = xz+yz$, $x(y+z) = xy+xz$, $\alpha(xy) = (\alpha x)y = x(\alpha y)$, $\|xy\| \leq \|x\|\|y\|$ and $\|I\| = 1$ for all $x,y,z \in A$, $\alpha \in \mathbf{C}$. (In the literature, Banach algebras are not always required to possess a unit.)

Some examples of Banach algebras are $\mathcal{B}(B)$, $\mathrm{TI}_\omega(B)$, $\mathrm{TIC}_\omega(B)$ or $\mathrm{H}^\infty(\Omega;B)$ for any complex Banach space B, real number ω and open set $\Omega \subset \mathbf{C}$.

Let A be a Banach algebra. We define $\mathcal{G}A := \{x \in A \mid xy = I = yx$ for some $y \in A\}$. For $x \in A$ we define the *spectrum* $\sigma(x) := \{\zeta \in \mathbf{C} \mid \zeta - x \notin \mathcal{G}A\}$ and the *spectral radius* $\rho(x) := \sup|\sigma(x)|$ of x; by $\zeta - x$ we mean $\zeta I - x$, where I is the identity operator on A. Note that these definitions coincide with those made earlier for $\mathcal{B}(B)$. We also set $x^0 := I$, $e^x := \sum_{n=0}^\infty x^n/n!$ for $x \in A$.

Lemma A.3.3 (Banach algebras) *Let A be a Banach algebra and $x,y,h \in A$. Then we have the following:*

(A0) If $x \in \mathcal{G}A$ and $\|x\| < 1$, then $\exists (I-x)^{-1} = \sum_{k=0}^\infty x^k$.

(A1) Let $x \in \mathcal{G}A$, and set $M := \|x^{-1}\|$. If $\|h\| < 1/2M$, then $x+h \in \mathcal{G}A$ and

$$\|(x+h)^{-1} - x^{-1}\| < 2M^3\|h\|^2 + M^2\|h\| < M. \tag{A.21}$$

(A2) The set $\mathcal{G}A$ of invertible elements is open in A, and $x \mapsto x^{-1}$ is a continuous bijection $\mathcal{G}A \to \mathcal{G}A$.

(A3) If $\{x_n\} \subset \mathcal{G}A$ and $x_n \to x \notin A \setminus \mathcal{G}A$, as $n \to \infty$, then $\|x_n^{-1}\| \to \infty$.

(A4) Let $x \in A \setminus \mathcal{G}A$. If $M > 0$, then there is $\delta > 0$ s.t. $y \in \mathcal{G}A$ & $\|y - x\| < \delta \Rightarrow \|y^{-1}\| > M$.

(r1) $\rho(x) = \lim_{n\to\infty}\|x^n\|^{1/n} = \inf_{n\geq 1}\|x^n\|^{1/n} = \max|\sigma(x)| \leq \|x\|$.

(s1) $\sigma(x) \subset \mathbf{C}$ *is compact and nonempty.*

(s2) $\sigma(xy) \cap \{0\} = \sigma(yx) \cup \{0\}$ *and* $\rho(xy) = \rho(yx)$.

(s3) If p is a nonconstant polynomial, then $\sigma(p(x)) = \{p(\zeta) \mid \zeta \in \sigma(x)\}$.

(s4) If $x \in \mathcal{G}A$, then $\sigma(x^{-1}) = \{\zeta^{-1} \mid \zeta \in \sigma(x)\}$.

Proof: (A0)–(A4) See Theorems 10.7, 10.11, 10.12 and 10.17 of [Rud73]. (A4) follows from (A3).

(r1)&(s1)&(s4) See Theorems 10.13 and 10.28 of [Rud73].

(s2)&(s3) See Theorems 3.1.2 and 3.2.4 of [Aupetit]. □

We remind the reader that a B^*-*algebra* (also called as a C^*-algebra), to which we sometimes refer, is a Banach algebra A with an *involution* $(\cdot)^* : A \to A$ satisfying $(x+y)^* = x^* + y^*$, $(\alpha x)^* = \widehat{\alpha} x^*$, $(xy)^* = y^*x^*$, $x^{**} = x$, and $\|xx^*\| = \|x\|^2$

for all $x,y \in A$, $\alpha \in \mathbf{C}$. Thus, by a *B^*-algebra isomorphism* $T : A \mapsto A'$ we mean that T is an algebra isomorphism and a Banach space isomorphism, and that $(Tx)^* = Tx^*$ for all $x \in A$.

The Banach algebras $\mathcal{B}(H)$, $\mathrm{TI}(H)$, $\mathcal{C}_{\mathrm{b}}(\Omega;\mathcal{B}(H))$ and $\mathrm{L}^\infty(Q;\mathcal{B}(H))$ are B^*-algebras whenever H is a complex Hilbert space, Ω is a topological space and Q is as on p. 907. See [Rud73] for more on B^*-algebras.

Let $x \in B$ and $\Lambda \in B^*$, where B is a Banach space. Then we write $\langle \Lambda, x\rangle_{\langle B^*,B\rangle} := \Lambda x =: \langle x, \Lambda\rangle_{\langle B,B^*\rangle}$ for the bilinear pairing between B and B^* (except that in Hilbert space contexts we use conjugate-linear scalar multiplication and hence we must set $\langle x, \Lambda\rangle_{\langle B,B^*\rangle} := \Lambda x =: \overline{\langle \Lambda, x\rangle_{\langle B^*,B\rangle}}$, so that pairing is still linear in its first argument but no more linear (but conjugate-linear) in its second argument (unless $\mathbf{K} = \mathbf{R}$), so that the pairing coincides with the inner product for Hilbert spaces; see Remark A.3.22 for details).

Lemma A.3.4 (Banach spaces) *Let B_1, B_2, B_3 and B_4 be Banach spaces and denote by $\mathcal{B}(B_i,B_j)$ the continuous and by $\mathcal{BC}(B_i,B_j)$ the compact linear mappings $B_i \to B_j$. Let $\mathcal{GB}(B_i,B_j)$ be the set of invertible elements of $\mathcal{B}(B_i,B_j)$ ($i,j = 1,2,3,4$).*

(B1) $\mathcal{BC}(B_1,B_2)$ is a Banach space.

Let $S \in \mathcal{B}(B_1,B_2)$, $K \in \mathcal{BC}(B_2,B_3)$, and $T \in \mathcal{B}(B_3,B_4)$. Then $SK, KT \in \mathcal{BC}$ (in particular, $\mathcal{BC}(B_1)$ is an ideal of $\mathcal{B}$). Moreover, $K \in \mathcal{BC} \Leftrightarrow K^* \in \mathcal{BC}$.

If $\dim B_1 = \infty$, then $\mathcal{BC}(B_1,B_2) \cap \mathcal{GB}(B_1,B_2) = \emptyset$. If $K[B_2]$ is closed, then $\dim K[B_2] < \infty$. If $\dim B_1 < \infty$ or $\dim B_2 < \infty$, then $\mathcal{BC}(B_1,B_2) = \mathcal{B}(B_1,B_2)$.

If $\{S_n\} \subset \mathcal{B}(B_1,B_2)$, $\dim S_n[B_1] < \infty$ for all n, and $S_n \to S$ as $n \to \infty$, then $S \in \mathcal{BC}(B_1,B_2)$. Conversely, if B_2 is a Hilbert space, then $\mathcal{BC}(B_1,B_2)$ is the closure of finite-dimensional operators.

(B2) Let U and Y be Hilbert spaces and $K \in \mathcal{BC}(U,Y)$. Then there are sequences $\{u_n\}_{n=1}^\infty \subset U$ and $\{y_n\}_{n=1}^\infty \subset Y$ s.t. when P_n [P_n'] is the orthogonal projection of U [Y] onto $\mathrm{span}(u_1,\ldots,u_n)$ [$\mathrm{span}(y_1,\ldots,y_n)$], we have $P_n'KP_n \to K$, as $n \to \infty$.

If $U = Y$, we can choose the sequence $\{u_n\} \subset U$ so that $P_nKP_n \to K$, as $n \to \infty$.

(B3) Let $C \in \mathcal{BC}(B_1,B_2)$ and $L \in \mathcal{B}(B_1,B_2)$. Then $L(I-C) = I$ iff $(I-C)L = I$. Moreover, $I - L \in \mathcal{BC}$, if $L(I-C) = I$.

(B4) The presentation $G+K$ of an operator $G+K \in \mathcal{GB}+\mathcal{BC}(B_1,B_2) := \{G+K \mid G \in \mathcal{GB}(B_1,B_2),\ K \in \mathcal{BC}(B_1,B_2)\}$ is unique unless $\dim B_1 = \dim B_2 < \infty$.

Moreover, $G+K$ is left-invertible iff it is right-invertible. If $G+K$ is invertible, then $(G+K)^{-1} = G^{-1} + K'$ for some $K' \in \mathcal{BC}$.

The class $\mathcal{GB}+\mathcal{BC}$ is closed w.r.t. composition and adjungation; in particular, $\mathcal{GB}+\mathcal{BC}(B_1)$ is a group. Obviously, $\mathcal{GB} \subset \mathcal{GB}+\mathcal{BC}$.

Each $A \in \mathcal{GB}+\mathcal{BC}(B_1,B_2)$ is a Fredholm operator, i.e., $\mathrm{Ran}(A)$ is closed, $\dim \mathrm{Ker}(A) < \infty$, and $\dim B_1/\,\mathrm{Ran}(A) < \infty$.

(B5) Let B_1 be complex. The set $\alpha I + K \in \mathbf{C}I + \mathcal{BC}(B_1)$ is a closed subalgebra of the Banach algebra $\mathcal{BC}(B_1)$, and this subalgebra is also closed w.r.t. inverses and adjoints. The presentation of an element of $\mathbf{C}I + \mathcal{BC}(B_1)$ is unique unless $\dim B_1 < \infty$.

Furthermore, $\alpha I + K$ is left-invertible iff it is right-invertible, and if $\alpha I + K$ is invertible, then $(\alpha I + K)^{-1} = \alpha^{-1} I + K'$ for some $K' \in \mathcal{BC}$.

If $\dim B_1 = \infty$ and $K \in \mathcal{BC}(B_1)$, then $\|\lambda I + K\| \geq |\lambda|$ for all $\lambda \in \mathbf{C}$.

(C1) Let $\begin{bmatrix} S & T \\ U & V \end{bmatrix} \in \mathcal{B}(B_1 \times B_2, B_3 \times B_4)$. Then $\| \begin{bmatrix} S & 0 \\ 0 & V \end{bmatrix} \| = \max(\|S\|, \|V\|)$, $\| \begin{bmatrix} S & T \end{bmatrix} \| \leq (\|S\|^2 + \|T\|^2)^{1/2}$, and $\| \begin{bmatrix} S \\ U \end{bmatrix} \| \leq (\|S\|^2 + \|U\|^2)^{1/2}$; more generally $\|(L_{ij})_{i=1,\ldots,n;\, j=1,\ldots,m}\| \leq \|(\|L_{ij}\|)_{i=1,\ldots,n;\, j=1,\ldots,m}\|_{\mathbf{K}^{n\times m}}$.

(D1) Let $R \in \mathcal{B}(B_1, B_2)$ and $\|Rx\| \geq \varepsilon \|x\|$ for all $x \in X$. Then R is one-to-one and its range is closed, in particular, $R \in \mathcal{GB}$ iff its range is dense. Moreover, if $R \in \mathcal{GB}$, then $\|R^{-1}\| \leq 1/\varepsilon$.

(E1) **(Closed Graph Theorem)** *Let $T : B_1 \to B_2$ be linear. Then $T \in \mathcal{B}(B_1, B_2)$ iff $x_n \to 0 \,\&\, Tx_n \to y \Rightarrow y = 0$ (for all $\{x_n\} \subset B_1$, $y \in B_2$).*

(F1) Claims (j3)–(k2) and (c3)(i)–(iii)&(vi) of Lemma A.3.1 hold for Banach spaces too.

(G1) Let $A : \mathrm{Dom}(A) \to B_2$ be linear, where $\mathrm{Dom}(A)$ is a subspace of B_1. Equip $\mathrm{Dom}(A)$ with norm $\|x\|_A^2 := \|x\|_{B_1}^2 + \|Ax\|_{B_2}^2$. Then $\mathrm{Dom}(A)$ is a normed space, and $A \in \mathcal{B}(\mathrm{Dom}(A), B_2)$. If B_1 and B_2 are inner product spaces, then so is $\mathrm{Dom}(A)$, with inner product $\langle x, y\rangle_A := \langle x, y\rangle_{B_1} + \langle Ax, Ay\rangle_{B_2}$. Moreover, $\mathrm{Dom}(A)$ is complete iff A is closed.

We recall from [Rud73] that A in (G1) is called closed *iff $x_n \to x$ and $Ax_n \to y$ imply that $x \in \mathrm{Dom}(A)$ and $Ax = y$ (this holds if $\{(x, Ax) \mid x \in \mathrm{Dom}(A)\}$ is a closed subset of $B_1 \times B_2$).*

If $B_1 = B_2$ and B_1 is complex, then we also define $\sigma(A)^c := \{\lambda \in \mathbf{C} \mid \exists (\lambda - A)^{-1} \in \mathcal{B}(B_1)\}$ (that is, $\lambda \notin \sigma(A)$ iff there is $T \in \mathcal{B}(B_1)$ s.t. $(\lambda I - A)T = I_{B_1}$ and $T(\lambda I - A) = I_{\mathrm{Dom}(A)}$).

(G2) Let $A \in \mathcal{B}(B_1, B_3)$. Let $B_2 \subset B_3$, continuously. Set $\mathrm{Dom}(A) := \{x \in B_1 \mid Ax \in B_2\}$. Then $A_{|\mathrm{Dom}(A)} : \mathrm{Dom}(A) \to B_2$ is closed and (G1) applies.

(G3) If A is as in (G1), $B_1 = B_2$ is complex and $\sigma(A)^c \neq \emptyset$, then A is closed and $x \mapsto \|(\alpha - A)x\|_{B_1}$ is an equivalent norm on $\mathrm{Dom}(A)$.

(H1) Let $F : \mathbf{R}_+ \to \mathcal{B}(B_1, B_2)$, $M < \infty$ be s.t. $\|F(t)\| \leq M$ for all $t \in \mathbf{R}_+$, and let the span of $X \subset B_1$ be dense. If $F(t)x \to 0$, as $t \to +\infty$, for all $x \in X$, then $F(t)x \to 0$ for all $x \in B_1$.

(H2) Let $F : \mathbf{R}_+ \to \mathcal{B}(B_1, B_2)$, $M < \infty$ be s.t. $\|F(t)\| \leq M$ for all $t \in \mathbf{R}_+$. If $K \subset B_1$ is compact and $F(t)x \to 0$, as $t \to +\infty$, for all $x \in K$, then $F(t)x \to 0$ uniformly for $x \in K$.

(I1) Let $B_1 \subset B_2$ be a subspace, $\Lambda \in B_2^*$, $x \in B_1$ and $\Lambda x \neq 0$. Then $B_1 = B_1 \cap \mathrm{Ker}(\Lambda) + \mathbf{K}x$.

(J1) **(Bilinear)** *Let $S : B_1 \times B_2 \to B_3$ be bilinear. Then S is continuous iff there is $M < \infty$ s.t.*

$$\|S(x,y)\|_{B_3} \le M\|x\|_{B_1}\|y\|_{B_2} \quad \text{for all } x \in B_1,\ y \in B_2. \tag{A.22}$$

Assume that this is the case. Then there is a unique $T \in \mathcal{B}(B_2, \mathcal{B}(B_1, B_3))$ s.t. $S(x,y) = (Ty)x$ for all $x \in B_1$, $y \in B_2$. Moreover, $\|T\|_{\mathcal{B}(B_2,\mathcal{B}(B_1,B_3))}$ is the minimal number M satisfying (A.22).

(K1) If $x \in B$, then there is $\Lambda \in B^*$ s.t. $\|\Lambda\| \le 1$ and $\Lambda x = \|x\|_B$.

(L1) If $\sum_{n\in\mathbf{N}} \|x_n\|_{B_1} < \infty$, then $\sum_{n\in\mathbf{N}} x_n$ converges absolutely, and $\|\sum_{n\in\mathbf{N}} x_n\|_{B_1} \le \sum_{n\in\mathbf{N}} \|x_n\|_{B_1}$.

(M1) Let $T \in \mathcal{B}(B_1, B_2)$. If B_1 is reflexive, then $T = T^{**}$.

(N1) Let $T \in \mathcal{B}(B_1, B_2)$. Then $\mathrm{Ran}(T)$ is dense in B_2 iff T^* is one-to-one.

(N2) If B_1 is reflexive, then $T \in \mathcal{B}(B_1, B_2)$ is one-to-one iff $\mathrm{Ran}(T^*)$ is dense in B_1^*.

(N3) Let $T \in \mathcal{B}(B_1, B_2)$. Then $T[B_1] = B_2$ iff $\|T^*y^*\|_{B_1^*} \ge \varepsilon\|y^*\|_{B_2^*}$ for some $\varepsilon > 0$ and all $y^* \in B_2^*$.

(N4) **($\|\mathbf{Tx}\| \ge \varepsilon\|\mathbf{x}\|$)** *Let $T \in \mathcal{B}(B_1, B_2)$. Then the following are equivalent:*

- *(i) $\|Tx\| \ge \varepsilon\|x\|$ for all x for some $\varepsilon > 0$, i.e., T is* coercive *("uniformly bounded from below");*
- *(iii) $\mathrm{Ran}(T^*) = B_1^*$, i.e., T^* is* onto*;*
- *(iv) $\mathrm{Ran}(T)$ is closed and $\mathrm{Ker}(T) = \{0\}$;*
- *(ix) $\|TX\|_{\mathcal{B}(*,H)} \ge \varepsilon\|X\|_{\mathcal{B}(*,*)}$ for some $\varepsilon > 0$ whenever X is linear (in particular, X is bounded iff TX is bounded);*
- *(x) $T \in \mathcal{GB}(U, \mathrm{Ran}(T))$;*
- *(xi) There is $\varepsilon > 0$ s.t. for all $x \in B_1 \setminus \{0\}$ there is $y^* \in B_2^* \setminus \{0\}$ s.t. $\langle y^*, Tx\rangle_{\langle B_2^*, B_2\rangle} \ge \varepsilon\|x\|\|y^*\|$.*

We can replace B_2^* by any of its norming subsets in (xi).

(N5) **($\mathbf{T \in \mathcal{GB}(U, H)}$)** *Let $T \in \mathcal{B}(U, H)$. Then the following are equivalent:*

- *(i) $T \in \mathcal{GB}(U, H)$ (i.e., T is invertible)*
- *(ii) T is one-to-one and onto;*
- *(iii) T is coercive and has a dense range;*
- *(vi) $T^* \in \mathcal{GB}$.*

Moreover, each of (i)–(xi) implies that $(T^*)^{-1} = (T^{-1})^*$.

(O1) **(Uniform Boundedness Principle)** *Let $\mathcal{A} \subset \mathcal{B}(B, B_2)$. If $\{\|Ax\|_{B_2}\}_{A\in\mathcal{A}}$ is bounded for each $x \in B$, then there is $M < \infty$ s.t. $\|A\|_{\mathcal{B}(B,B_2)} \le M$ for all $A \in \mathcal{A}$.*

(P1) **(Completion)** *If X is a normed space, then there is a Banach space $\overline{X}$ (the* completion *of X) s.t. X is a dense subspace of $\overline{X}$ (with the same topology). If X is an inner product space, then $\overline{X}$ is a Hilbert space.*

(Q1) **($\dim X < \infty$)** *If X is a normed space and $n := \dim X < \infty$, then X is isomorphic to $\mathbf{K}^n$ (with equivalent norms).*

(R1) The Banach space B_1 is reflexive iff B_1^* is reflexive.

(R2) If B_1^* is reflexive, then B_1 is reflexive.

(R3) The Banach space B_1 is reflexive and separable iff B_1^* is reflexive and separable.

(S1) **(Subspaces)** *Let X be a closed subspace of a normed space Y. If Y is complete (resp. reflexive, separable), then so is X.*

By (J1), the space of bilinear continuous mappings $B_1 \times B_2 \to B_3$ with norm $\inf\{M \,|\, M$ satisfies (A.22)$\}$ is a Banach space isometrically isomorphic to $\mathcal{B}(B_2, \mathcal{B}(B_1, B_3))$ (and to $\mathcal{B}(B_1, \mathcal{B}(B_2, B_3))$).

Proof: (B1) All these claims are given in Chapter 4 of [Rud73] (Theorems 4.18 and 4.19 in particular) or easily deducible from them. Although the results of [Rud73] concern the Banach space adjoint $(\cdot)^{\mathrm{B}}$ only, (e.g., $K \in \mathcal{BC}(B_1, B_2) \Leftrightarrow K^{\mathrm{B}} \in \mathcal{BC}(B_2^*, B_1^*)$), they are applicable for the Hilbert space adjoint $(\cdot)^{\mathrm{H}}$ too (when the Banach spaces concerned happen to be Hilbert spaces), because H is isomorphic to H^* as a real Hilbert space [Rud73, 12.5] (note that each complex Hilbert space is also a real Hilbert space), hence $K^{\mathrm{H}} \in \mathcal{BC}(B_2, B_1) \Leftrightarrow K^{\mathrm{B}} \in \mathcal{BC}(B_2^*, B_1^*)$, $K^{\mathrm{H}} \in \mathcal{GB}(B_2, B_1) \Leftrightarrow K^{\mathrm{B}} \in \mathcal{GB}(B_2^*, B_1^*)$ etc. The last claim can be proved as the first one in (B2).

(B2) 1° Let $F := \{u \in U \,|\, \|u\| \le 1\}$. For each $n \in \mathbf{N}$, choose $m_n \in \mathbf{N}$, $y_{n,1}, y_{n,2}, \ldots, y_{n,m_n} \in K[F]$ s.t. the balls $\{y \in Y \,|\, \|y - y_{n,k}\| < 1/n\}$ cover $K[F]$. If P_n'' is the orthogonal projection of Y onto $\mathrm{span}(y_{n,1}, y_{n,2}, \ldots, y_{n,m_n})$, then clearly $\|y - P_n'' y\| < 1/n$ for each $y \in K[F]$, i.e., $\|Ku - P_n'' Ku\| < 1/n$ whenever $\|u\| \le 1$. Let $y_1', y_2', \ldots$ be the sequence $y_{1,1}, y_{1,2}, \ldots, y_{1,m_1}, y_{2,1}, \ldots$ so that $\|(I - P_n')K\| \to 0$, as $n \to \infty$.

2° Choose $\{u_k\} \subset U$ analogously for $K^* \in \mathcal{BC}(Y, U)$. Then $\|K(I - P_n)\| = \|(I - P_n)K^*\| \to 0$, hence $\|K - P_n' K P_n\| = \|(I - P_n')K + P_n'[K(I - P_n)]\| \to 0$, as $n \to \infty$, as required.

3° The sequence $u_1, y_1, u_2, y_2, \ldots$ satisfies the requirements of the last claim, because the projections P_n'' thus obtained satisfy $\mathrm{Ran}(I - P_{2n}'') = \mathrm{Ker}(P_{2n}'') \subset \mathrm{Ker}(P_n) \cap \mathrm{Ker}(P_n')$

(B3) See [Rud73, Exercise 4.9].

(B4) 1° If (there is) $G \in \mathcal{GB}(B_1, B_2)$, then $G[B_1] = B_2$, and B_1 and B_2 have the same dimension ($\in \mathbf{N} \cup \{\infty\}$).

If $K \in \mathcal{BC}(B_1, B_2)$, then [Rud73, Theorem 4.18(a)] implies that $\dim K[B_1] < \infty$; hence then $\mathcal{GB} \cap \mathcal{BC} = \{0\}$ unless $\dim B_1 = \dim B_2 < \infty$. If $G + K = G' + K' \in \mathcal{GB} + \mathcal{BC}$, then $\mathcal{GB} \ni G - G' = K' - K \in \mathcal{BC}$, hence the uniqueness.

2° Set $C := -G^{-1}K \in \mathcal{BC}$. By (B3) and the invertibility of G, we have $I = S(G + K) = SG(I - C) \Leftrightarrow (I - C)SG = I \Leftrightarrow G(I - C)S = I$ for any $S \in \mathcal{B}$, hence $G + K$ is right-invertible iff it is left-invertible.

Let now $\exists (G + K)^{-1} = S \in \mathcal{B}$. Set $(I - C') := (I - C)^{-1}$ to obtain $C' \in \mathcal{BC}$, by (B3). Then $S = (I - C)^{-1} G^{-1} = G^{-1} - C' G^{-1} \in \mathcal{BGC}$.

3° Obviously, $G^* + K^* \in \mathcal{GB} + \mathcal{BC}$ and $(G+K)(G'+K') = GG' + (GK' + KG' + KK') \in \mathcal{GB} + \mathcal{BC}$ when $G+K \in \mathcal{GB} + \mathcal{BC}(B_1, B_2)$ and $G'+K' \in \mathcal{GB} + \mathcal{BC}(B_2, B_3)$.

4° Fredholm: If $A = I + K$, $K \in \mathcal{BC}$, then the claim follows from Theorems 4.23 and 4.25 of[Rud73]. If $G + K' \in \mathcal{GB} + \mathcal{BC}(B_1, B_2)$, then $G + K'$ is a Fredholm operator, because so is $A := I + K'G^{-1}$ (by what we have noted above) and we have $G + K' = AG$.

(B5) Most claims follows easily from (B4). For the norm inequality, let $K \in \mathcal{BC}(B_1)$ and $\dim B_1 = \infty$. If we had $\|Kx\| \geq \varepsilon\|x\|$ for all $x \in B_1$ for some $\varepsilon > 0$, then K^* would be onto [Rud73, 4.15] and compact, which is impossible [Rud73, 4.18(b)].

If $\|Kx\| < \varepsilon\|x\|$, then $\|\lambda x + Kx\| \geq (|\lambda| - \varepsilon)\|x\|$; because this kind of an x exists for an arbitrary $\varepsilon > 0$, we have $\|\lambda I + K\| \geq |\lambda|$.

It follows that if $\{\lambda_n I + K_n\}$ is a Cauchy sequence, then so is $\{\lambda_n\}$, hence then $\lambda_n \to \lambda$ for some $\lambda \in \mathbf{C}$, which in turn implies that also $\{K_n\}$ is a Cauchy sequence, thus converging to some $K \in \mathcal{BC}$. (If, instead, $\dim B_1 < \infty$, then $I + \mathcal{BC} = \mathcal{BC} = \mathcal{B}$.)

(C1) This is easy: if $x \in B_1 \times \ldots \times B_m$ and we replace each term by its norm in the vector Lx, we get a product of the form $\mathbf{K}^{n\times m} \cdot \mathbf{K}^m$, from which the upper bound is obtained (the converse does not hold, e.g., $\| \begin{bmatrix} A & B \end{bmatrix} \| = 1 < \sqrt{2} = \| \begin{bmatrix} 1 & 1 \end{bmatrix} \|$, when $A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$).

We have used here the 2-norm $\|(b, b')\|^2 := \|b\|^2 + \|b'\|^2 = |(\|b\|, \|b'\|)|_{\mathbf{K}^2}$ (which is always compatible with the product topology and the canonical one in case of Hilbert spaces); of course the results are different for other norms. Also the space $\mathbf{K}^{n\times m}$ is equipped with its canonical ($\mathcal{B}(\mathbf{K}^m, \mathbf{K}^n)$) norm (i.e., $\|A\|_{\mathbf{K}^{n\times m}} := \sup_{\|x\| \leq 1} \|Ax\|$; if $\mathbf{K} = \mathbf{C}$, then one can show that $\|A\|_{\mathbf{K}^{n\times m}} = \max \sigma(A^*A)^{1/2}$).

(D1) The first claim is straightforward, the equivalence follows from the open mapping theorem [Rud73, 2.12(b)], and the inequality from the formula $\|x\| = \|RR^{-1}x\| \geq \varepsilon\|R^{-1}x\|$ for all $x \in H$.

(E1) “Only if” is obvious; “if” follows from Theorem 2.15 of [Rud73].

(F1) The same proofs apply (see the references in the proof of (h1)).

(G1) The claims are quite obvious (note that $\{x_n\}$ is a Cauchy-sequence in $\mathrm{Dom}(A)$ iff $\{x_n\}$ and $\{Ax_n\}$ converge to some $x \in B_1$ and $y \in B_2$, respectively).

(G2) Let $\{x_n\} \subset \mathrm{Dom}(A)$, $x_n \to x$ in B_1, and $Ax_n \to y$ in B_2. Then $Ax_n \to y$ in B_3 too, by also $Ax_n \to Ax$ in B_3, hence $Ax = y$ and $x \in \mathrm{Dom}(A)$.

(G3) One easily verifies the first claim. If $\alpha \in \sigma(A)^c$ and $\{x_n\}$ is a Cauchy-sequence in $\mathrm{Dom}(A)$, then $(\alpha - A)x_n \to y$ for some $y \in B_1$, hence $x_n \to (\alpha - A)^{-1}y =: x$, because $(\alpha - A)^{-1} \in \mathcal{B}(B_1)$. Consequently, then $x \in (\alpha - A)^{-1}[B_1] = \mathrm{Dom}(A)$ and $(\alpha - A)x = y$, so that $x_n \to x$ in $\mathrm{Dom}(A)$. Thus, $\mathrm{Dom}(A)$ is complete.

(H1) Let $x \in B_1$. Replace X by its span. Choose $x' \in X$ s.t. $\|x - x'\| < \varepsilon/2M$. If $\|F(t)x'\| < \varepsilon/2$, then $\|F(t)x\| < \varepsilon$.

(H2) Let $\varepsilon > 0$ be given. Set $\delta := \varepsilon/2M$. Then there are $n \in \mathbf{N}$ and $x_1, \ldots, x_n \in K$ s.t. $K \subset \cup_{k=1}^n D(x_k, \delta)$. Choose $T_1, \ldots, T_n < \infty$ s.t. $\|F(t)x_k\| < \varepsilon/2$ for all $t > T_k$, for all $k = 1, \ldots, n$.

If $x\in K$ and $t>\max\{T_1,\dots,T_n\}$, choose k s.t. $\|x-x_k\|<\delta$ to observe that $\|F(t)x\|\le\|F(t)x_k\|+\|F(t)(x-x_k)\|<\varepsilon/2+M\delta=\varepsilon$.

(I1) Obviously, $B_1\cap\mathrm{Ker}(\Lambda)+\mathbf{K}x\subset B_1$. For the converse, assume that $z\in B_1$. Choose $\alpha\in\mathbf{K}$ s.t. $\Lambda z=\alpha\Lambda x$. Then $\alpha x-z\in\mathrm{Ker}(\Lambda)\cap B_1$.

(J1) 1° Assume that S is continuous. Then there are $\varepsilon,\varepsilon'>0$ s.t. $S[D^1_\varepsilon\times D^2_{\varepsilon'}]\subset D^3_1$, where $D^k_r:=\{x\in B^k\,\big|\,\|x\|<r\}$. By bilinearity, we can take $M:=(\min\{\varepsilon,\varepsilon'\})^{-2}$.

2° Assume that $M<\infty$ satisfies (A.22). (If $M=0$, then $S=0$, hence we assume that $M>0$.) Then, for any $\varepsilon>0$, we have $S[D^1_\delta\times D^2_\delta]\subset D^3_\varepsilon$, where $\delta:=(M/\varepsilon)^{-1/2}$. Thus, S is continuous at zero. By bilinearity, we have

$$\begin{aligned}S(x+x',y+y')-S(x,y)&=S(x+x',y+y')-S(x+x',y)+S(x+x',y)-S(x,y)\\&=S(x+x',y')+S(x',y)\end{aligned}\tag{A.23}$$

for all $x,x'\in B_1$, $y,y'\in B_2$. Given $(x,y)\in B_1\times B_2$ and $\varepsilon>0$, it obviously follows that $\|S(x+x',y+y')-S(x,y)\|_{B_3}<\varepsilon$ for $\|x'\|,\|y'\|<\delta:=\varepsilon/2M(R+1)$, where $R:=\max\{\|x\|,\|y\|\}$. Therefore, S is continuous.

3° Assume that $M<\infty$ satisfies (A.22). For each $y\in Y$, the map $T_y:x\mapsto S(x,y)$ is in $\mathcal{B}(B_1,B_3)$, and $T:y\mapsto T_y$ is linear and $\|T_y\|_{\mathcal{B}(B_1,B_3)}\le M\|y\|_{B_2}$, hence T is as in (J1) and $\|T\|\le M$.

We have $\|S(x,y)\|_{B_3}\le\|Ty\|\|x\|\le\|T\|\|y\|\|x\|$, thus $\|T\|$ satisfies (A.22) in place of M; because $\|T\|\le M$ for any other such number M, $\|T\|$ is the infimal one.

(K1) See Corollary 3.3 of [Rud73].

(L1) This is an easy exercise.

(M1) For all $x\in B_1=B_1^{**}$ and $y\in B_2^*$, we have

$$\langle y^*,Tx\rangle_{\langle B_2^*,B_2\rangle}=\langle T^*y^*,x\rangle_{\langle B_1^*,B_1\rangle}=\langle T^*y^*,x\rangle_{\langle B_1^*,B_1^{**}\rangle}=\langle y^*,T^{**}x\rangle_{\langle B_2^*,B_2^{**}\rangle}.\tag{A.24}$$

Thus, $T^{**}x\in B_2^{**}$ is the element $y^*\mapsto\langle y^*,Tx\rangle_{B_2^*,B_2}$ of B_2^{**}, which is identified to $Tx\in B_2$. I.e., the range of $T^{**}\in\mathcal{B}(B_1^{**},B_2^{**})=\mathcal{B}(B_1,B_2^{**})$ is contained in the closed subspace B_2 of B_2^{**}.

(N1) This is Corollary 4.12(b) of [Rud73].

(N2) By (N1), $\mathrm{Ran}(T^*)$ is dense iff T^{**} is one-to-one. But $T^{**}=T$, by (M1).

(N3) This is Theorem 4.15 of [Rud73].

(N4) One easily verifies the implication "(i)⇒(iv)". We obtain "(ix)⇔(i)⇐(xi)" and "(iv)⇒(x)⇒(i)" as in Lemma A.3.1(c3), "(i)⇔(iii)" from Theorems 4.14 and 4.12(c) of [Rud73], and "(i)⇒(xi) from (K1).

(N7) "(iii)⇔(i)⇒(ii)" follows from (N4)(x), "(ii)⇒(vi)" from (N3) and Theorems 4.12(c)&4.14 (since then T^* is coercive and onto) "(vi)⇒(iii)" from (N3)&(N4)(iii).

Finally, if $y\in B_2$ and $y^*\in B_2^*$, then $\langle(T^{-1})^*T^*y^*,y\rangle=\langle y^*,TT^{-1}y\rangle=\langle y^*,y\rangle$, hence $(T^{-1})^*T^*=I$, hence $(T^{-1})^*=(T^*)^{-1}$.

(O1) See, e.g., [Rud86], Theorem 5.8.

(P1) This is well-known (any metric space has a completion — a complete metric space in which the space is complete, by Exercise 24 of [Rud76]; one easily verifies that vector operations, norm and inner product can be extended to the completion).

(Q1) By Theorem 1.21 of [Rud73], every vector isomorphism (i.e., linear bijection) from $\mathbf{K}^n$ to a TVS is a homeomorphism, hence a TVS isomorphism.

(R1)&(S1) See Exercice 4.1(d)&(f) of [Rud73].

(R2) See [Adams, Theorem 1.14].

(R3) This follows from (R1) and (R2). □

Sometimes one needs concepts such as "self-adjoint" also for spaces other than inner product spaces. Such extended concepts are treated below (see Remark A.3.22 for $(\cdot)^{\mathrm{d}}$):

Lemma A.3.5 *Let $J,S \in \mathcal{B}(B,B^{\mathrm{d}})$, where B is a Banach space. Then (we use below sesquilinear pairing)*

(a) If $\langle x,Jx\rangle = \langle x,Sx\rangle$ for all $x \in B$, then $J = S$.

(b) We have $\overline{\langle y,Jx\rangle}_{\langle B^{\mathrm{dd}},B^{\mathrm{d}}\rangle} = \langle Jx,y\rangle_{\langle B^{\mathrm{d}},B^{\mathrm{dd}}\rangle} = \langle x,J^{\mathrm{d}}y\rangle_{\langle B,B^{\mathrm{d}}\rangle}$ for all $x,y \in B$.

(c1) The following are equivalent:

(i) $J = J^{\mathrm{d}}|_B$;

(ii) $\langle y,Jx\rangle_{\langle B,B^{\mathrm{d}}\rangle} = \overline{\langle x,Jy\rangle_{\langle B,B^{\mathrm{d}}\rangle}}$ ($=: \langle Jy,x\rangle_{\langle B^{\mathrm{d}},B^{\mathrm{dd}}\rangle}$) for all $x,y \in B$;

(iii) $\langle x,Jx\rangle = \overline{\langle x,Jx\rangle}$ for all $x \in B$.

(Thus, "$J \geq 0$" (i.e., $\langle x,Jx\rangle \geq 0$ for all $x \in B$) implies that $J = J^{\mathrm{d}}|_B$.)

(c2) Assume that $J = J^{\mathrm{d}}|_B$. Then T is invertible iff T is onto (iff T is coercive and B is reflexive).

(d) Then the following are equivalent:

(i) "$J \gg 0$" (i.e., $\langle x,Jx\rangle_{\langle B,B^{\mathrm{d}}\rangle} \geq \varepsilon\|x\|^2$ for some $\varepsilon > 0$ and all $x \in B$);

(ii) "$J \geq 0$" and J is coercive.

Moreover, in either case $J = J^{\mathrm{d}}|_B$ (hence $J \in \mathcal{GB}(B,B^*)$ iff B is reflexive).

Proof: (We apply $()^{\mathrm{B}}$ results freely for $()^{\mathrm{d}}$; see Remark A.3.20 for justification. Note that J,S are linear to B^{d}, hence not to B^{B} (unless $\mathbf{K} = \mathbf{R}$ or $B = \{0\}$); cf. Remark A.3.22.)

(a) The proof of Theorem 12.7 of [Rud73] shows that $\langle x,(J-S)y\rangle = 0$ for all $x,y \in B$, hence $(J-S)y = 0 \in B^{\mathrm{d}}$ for all $y \in B$, hence $J - S = 0$.

(b) This is obvious (see Remark A.3.22).

(c1) We obtain "(i)⇔(iii)" from (a) and (b). Since $\langle y,J^{\mathrm{d}}x\rangle := \langle Jy,x\rangle$, we have "(i)⇔(ii)"

(c2) The map J is onto (i.e., $J[B] = B^{\mathrm{d}}$) iff $J^{\mathrm{d}} \in \mathcal{B}(B^{\mathrm{dd}},B^{d})$ is coercive, by Lemma A.3.4(N3). In either case, $J = J^{\mathrm{d}}|_B$ is coercive, hence $J,J^{\mathrm{d}} \in \mathcal{GB}$, by (N5)(iii)&(i)&(iv). Thus, $J[B] = B^{\mathrm{d}} = J^{\mathrm{d}}[B^{\mathrm{dd}}]$, hence $B = B^{\mathrm{dd}}$, i.e., B is reflexive.

Finally, if(f) B is reflexive, then $J^{\mathrm{d}} = J^{\mathrm{d}}|_B = J$, hence then coercivity implies "onto", by (N3).

(d) (Note that (i) or (ii) implies that J is an isomorphism onto its range, by Lemma A.3.4(N4).)

Trivially, (i) implies (ii) and (ii) implies that (c1)(iii) holds. Assume then (ii), so that $\|Jx\| \geq \delta\|x\|$ for all $x \in X$ and some $\delta > 0$. Set $\varepsilon := \delta^2/4\|J\|$.

Let $x \in B$ be arbitrary. There is $y \in B$ s.t. $\|y\| = 1$ and $\langle y, Jx\rangle \geq \delta\|x\|/2$. By (c1)(ii) we have

$$0 \leq \langle (rx+y), J(rx+y)\rangle = \langle y, Jy\rangle + 2r\,\mathrm{Re}\,\langle x, Jy\rangle + r^2\langle x, Jx\rangle \quad (r \in \mathbf{R}). \tag{A.25}$$

Thus, (A.25) is a real polynomial with at most one root, hence $(2\,\mathrm{Re}\langle x, Jy\rangle)^2 - 4\langle x, Jx\rangle\langle y, Jy\rangle \leq 0$, hence

$$\langle x, Jx\rangle \geq \mathrm{Re}\,\langle x, Jy\rangle/\langle y, Jy\rangle \geq (\delta/2)^2\|x\|^2/\|J\| = \varepsilon\|x\|^2. \tag{A.26}$$

□

We end this section with some auxiliary lemmas. We often use the following lemma to show the continuity of a continuous mapping under a stronger norm:

Lemma A.3.6 ($TX_1 \subset X_2 \Rightarrow T \in \mathcal{B}(X_1, X_2)$) *Assume that X_1 and X_2 are Banach spaces, X_3 is a TVS, and $X_2 \underset{c}{\subset} X_3$. If $T \in \mathcal{B}(X_1, X_3)$ and $TX_1 \subset X_2$, then $T \in \mathcal{B}(X_1, X_2)$.*

Here $X_2 \underset{c}{\subset} X_3$ means that $X_2 \subset X_3$ continuously, i.e., there is $M < \infty$ s.t. $\|x\|_{X_3} \leq \|x\|_{X_2}$ for all $x \in X_2$. A continuous inclusion is called an *embedding* (or an imbedding). E.g., $\mathrm{L}^\infty([0,1]) \underset{c}{\subset} \mathrm{L}^1([0,1])$. (Sometimes one means by an embedding any continuous injective linear mapping; the name *inclusion* requires injectivity and some sense of unique (canonical) identification.)

It follows that $X_1 \underset{c}{\subset} X_2 \underset{c}{\subset} X_3 \implies X_1 \underset{c}{\subset} X_3$ and that $X_1 \underset{c}{\subset} X_2$ & $T \in \mathcal{B}(X_2, X_3) \implies T \in \mathcal{B}(X_1, X_3)$.

Proof of Lemma A.3.6: Let $x_n \to x$ in X_1 and $Tx_n \to 0$ in X_2. Then $Tx_n \to 0$ in X_3, hence $Tx = 0$, hence $T \in \mathcal{B}(X_1, X_2)$, by Lemma A.3.4(E1). □

By setting $T := I$ and applying Lemma A.3.6, we obtain the following:

Corollary A.3.7 (Inclusions are continuous) *Assume that X_1 and X_2 are Banach spaces and X_3 is a TVS. If $X_1 \underset{c}{\subset} X_3$, $X_2 \underset{c}{\subset} X_3$ and $X_1 \subset X_2$, then $X_1 \underset{c}{\subset} X_2$.*

In particular, if $X_1 = X_2$ as sets and $X_1, X_2 \underset{c}{\subset} X_3$, then X_1 and X_2 have equivalent norms (i.e., the same topology). □

Sometimes we wish to apply the above corollary with $X_3 = \mathrm{L}^p_{\mathrm{loc}}(\Omega; B)$, where $\Omega \subset \mathbf{R}^n$ is open or $\Omega \subset \mathbf{Z}^n$ (in the latter case $\mathrm{L}^p_{\mathrm{loc}}(\Omega; B)$ becomes the set of all functions $\Omega \to B$ with the topology of pointwise convergence). For that purpose, we note that $\mathrm{L}^p_{\mathrm{loc}}$ is a TVS (even a Fréchet space) with the topology induced by the seminorms $\|\cdot\|_{\mathrm{L}^p(K;B)}$ ($K \subset \Omega$, K is compact); see Appendix B.3 (or Chapter 1 of [Rud73]) for details.

Definition A.3.8 (Norming set) *A set $C \subset B^*$ is a* norming set (for B) *(or a determining set) if $\|x\|_B = \sup_{\Lambda \in C} |\Lambda x|$ for each $x \in B$. A subspace $X \subset B^*$ is a* norming subspace (for B) *if it contains a norming set.*

Note that $\|\Lambda\| \leq 1$ for each $\Lambda \in C$ above; obviously, we can redefine C s.t. $\|\Lambda\| = 1$ for each $\Lambda \in C$ unless $B = \{0\}$.

We may take X to be B^* or a dense subspace of B^*; if $B = B_2^*$, then we may take $X = B_2$.

Lemma A.3.9 *Let B be separable. Then B and B^* possess countable norming sets.* □

(This is Theorem 2.8.5 of [HP].)

Lemma A.3.10 (Extension by density) *Let X be a normed space and Y a Banach space. Let X_0 be a dense subspace of X (with the same norm) and let $T \in \mathcal{B}(X_0, Y)$. Then there is a unique $\overline{T} \in \mathcal{B}(X, Y)$ s.t. $\overline{T}x = Tx$ for all $x \in X_0$. Moreover, $\|\overline{T}\| = \|T\|$.* □

The lemma is approximately Proposition 2.3.1 of [Rauch]; the proof is straightforward (if $\{x_n\} \subset X_0$ is a Cauchy sequence, then so is $\{Tx_n\}$, etc.), and we omit it. Note that one can identify X_0^* and X^* (as well as $\mathcal{B}(X_0, Y)$ and $\mathcal{B}(X, Y)$) by the above lemma.

Lemma A.3.11 (Norm-preserving extension) *Let $T \in \mathcal{B}(X, B_2)$, where X is a subspace of B. If $B_2 = \mathbf{K}^n$ or B is a Hilbert space, then T has an extension $S \in \mathcal{B}(B, B_2)$ s.t. $\|S\| = \|T\|$.*

(Note, in contrast, that $I \in \mathcal{B}(\mathrm{H}^1, \mathrm{H}^1)$ has no extension in $\mathcal{B}(\mathrm{L}^1, \mathrm{H}^1)$, by p. 154 of [Hoffman]. However, whenever $B_2 = \mathbf{K}^n$, we can extend all n components of T to get a continuous extension of T (possibly with a greater norm).)

Proof: If $B_2 = \mathbf{K}$, any Hahn–Banach extension of T will do. If B is a Hilbert space, then we can extend T to $S \in \mathcal{B}(\overline{X}, B_2)$ by Lemma A.3.10, and set $S = 0$ on $X^\perp$. □

Not all continuous linear mappings can be extended (e.g., $\mathrm{H}^1(\mathbf{D}$ is not complemented in $\mathrm{L}^1(\partial\mathbf{D})$, by p. 130 of [Rud73]):

Lemma A.3.12 *Let X be a subspace of B. Then the following are equivalent:*

(i) For each Banach space B_2 and operator $T \in \mathcal{B}(X, B_2)$ there is an extension $S \in \mathcal{B}(B, B_2)$ of T;

(ii) $\overline{X}$ is complemented in B;

(iii) there is a continuous projection $B \to \overline{X}$.

We say that a closed subspace $M \subset B$ is *complemented in B* if there is a closed subspace $N \subset B$ s.t. $B = M + N$ and $M \cap N = \{0\}$.

Proof: By Lemma A.3.10, we assume that X is closed, w.l.o.g.

(ii)⇔(iii) This is Theorem 5.16 of [Rud73].

(i)⇔(iii) If (iii) holds and T is given, set $S = TP$, where $P \in \mathcal{B}(B,X)$ is a projection, to establish (i). Conversely, if (i) holds, there is a continuous extension of $I \in \mathcal{B}(X,X)$ to an operator $P \in \mathcal{B}(B,X)$, which must satisfy $PP = IP = P$. □

Lemma A.3.13 ($\mathcal{B}(Y,Z) \underset{c}{\subset} \mathcal{B}(X,Z)$) *Let X, Y and Z be normed spaces. If $X \underset{c}{\subset} Y$ densely, then $\mathcal{B}(Y,Z) \underset{c}{\subset} \mathcal{B}(X,Z)$.*

Trivially, if $Y \underset{c}{\subset} Z$, then $\mathcal{B}(X,Y) \underset{c}{\subset} \mathcal{B}(X,Z)$.

Proof: Choose $M < \infty$ s.t. $\|x\|_Y \le M\|x\|_X$ $(x \in X)$. Let $T \in \mathcal{B}(Y,Z)$. Then $T_{|X} \in \mathcal{B}(X,Z)$ determines T uniquely by density, hence $\mathcal{B}(Y,Z) \subset \mathcal{B}(X,Z)$. Moreover,

$$\|Tx\|_Z \le \|T\|_{\mathcal{B}(Y,Z)}\|x\|_Y \le M\|T\|_{\mathcal{B}(Y,Z)}\|x\|_X \quad (x \in X), \tag{A.27}$$

hence $\|T\|_{\mathcal{B}(X,Z)} \le M\|T\|_{\mathcal{B}(Y,Z)}$, i.e., the inclusion is continuous. □

Lemma A.3.14 (Bounds for a projection) *Let $Y \neq X$ be a closed subspace of the Banach space X and $\varepsilon > 0$. Then there is $x \in X$ s.t. $\|x\| = 1$ and $d(x,Y) := \inf_{y\in Y}\|x-y\| > (1+\varepsilon)^{-1}$. Consequently, $\|P_1\| < 1+\varepsilon$ and $\|P_0\| < 2+\varepsilon$, where $P_0 : (y+\alpha x) \to y$ and $P_1 : (y+\alpha x) \to x$, for $y \in Y$ and $\alpha \in \mathbf{K}$.*

Moreover, there is $\Lambda \in X^*$ s.t. $\|\Lambda\| = 1$, $\Lambda = 0$ on Y, and $\|\Lambda x\| > (1+\varepsilon)^{-1}$.

Proof: The existence of x follows from [Rud73, Lemma 4.22].

Choose $\delta < \varepsilon$ s.t. $d(x,Y) > (1+\delta)^{-1}$. For $\alpha \neq 0$ we have

$$\|y+\alpha x\| = |\alpha|\|y/\alpha + x\| > (1+\delta)^{-1}|\alpha| \tag{A.28}$$

and $\|P_1(y+\alpha x)\| = \|\alpha x\| = |\alpha| < (1+\delta)\|y+\alpha x\|$, hence $\|P_1\| \le 1+\delta < 1+\varepsilon$. Consequently, $\|P_0\| = \|I - P_1\| < 1+1+\varepsilon$.

Set $\Lambda'\alpha x := \alpha$ for $\alpha \in \mathbf{K}$, so that $\Lambda' = 1$. Set $\Lambda'' := \|P_0\|^{-1}\Lambda' P_0 \in X_0^*$, so that $\|\Lambda''\| = 1$ and $\Lambda'' x = \|P_0\|^{-1} > (1+\varepsilon)^{-1}$. Now Λ can be taken to be a Hahn–Banach extension of Λ''. □

We now show how norms can be induced and coinduced:

Lemma A.3.15 (Induction) *Let X be a vector space, Y a normed space and $T \in \mathrm{Hom}(X,Y)$ s.t. $\mathrm{Ker}(T) = \{0\}$. Set $\|x\|_X := \|Tx\|_Y$ for each $x \in X$.*

Then X becomes a normed space, $T \in \mathcal{B}(X,Y)$, and T is an isometry. If Y is an inner product space, then so is X.

Assume that $X' := X$ is a Banach space under some norm $\|\cdot\|_X'$, and that $T \in \mathcal{B}(X',Y)$ and $\|Tx\|_Y \ge \varepsilon\|x\|_X'$ for all $x \in X$ and some $\varepsilon > 0$. If Y is complete (resp. reflexive), then so is X, and then X is isometrically isomorphic to a closed subspace of Y.

Proof: 1° Obviously, $\|\cdot\|_X$ is a norm on X and T is an isometry (hence $T \in \mathcal{B}(X,Y)$). If $\langle\cdot,\cdot\rangle$ is an inner product on Y, then $\langle T\cdot, T\cdot\rangle$ is obviously an inner product on X and $\langle Tx, Tx\rangle = \|x\|_X^2$ for all $x \in X$.

Remark: If $X \subset Y$ and we set $T = I$, then $\|\cdot\|_X = \|\cdot\|_{Y|_X}$; in particular, X need not be complete even if Y were complete.

2° *When Y is complete:* If $\{x_n\}$ is Cauchy in X, then $\{Tx_n\}$ is Cauchy in Y, hence $\{x_n\}$ is Cauchy in X', hence then $x_n \to x$ in X' for some $x \in X$. By continuity, $Tx_n \to Tx$ in Y, hence $x_n \to x$ in X.

It follows that T is an isometric isomorphism of X to a closed subspace of Y.

3° *When Y is reflexive:* Being reflexive, Y is complete, hence so is X. Therefore, X is reflexive, by Lemma A.3.4(S1) and 2°. □

Lemma A.3.16 (Coinduction) *Let X and Y be normed spaces and $T \in \mathcal{B}(X,Y)$. Set $Z := T[X]$. Then $\|z\|_Z := \inf\{\|x\|_X \mid x \in X,\ z = Tx\}$ is a norm on Z. With this norm, we have $\|T\|_{\mathcal{B}(X,Z)} \le 1$ and $\|\cdot\|_Y \le \|T\|_{\mathcal{B}(X,Y)}\|\cdot\|_Z$. If X is complete or a Hilbert space, then so is Z.*

Proof: Because the space $N := \mathrm{Ker}(T) \subset X$ is closed, the quotient space X/N with norm $\|x+N\|_{X/N} := \inf\|x+N\|_X$ is a normed space, and it is complete if X is complete, by Theorem 1.41 of [Rud73].

Set $S(x+N) := Tx$ for $x \in X$. Then $S : X/N \to Z$ is obviously a linear bijection, and $\|S(x+N)\|_Z = \|x+N\|_{X/N}$ for all $x \in X$. It follows that $\|\cdot\|_Z$ is a norm and it makes S an isometric isomorphism. In particular, if X is complete, then Z is complete.

For each $x \in X$, we have

$$\|Tx\|_Y = \inf_{Tx'=Tx}\|Tx'\|_Y \le \|T\|_{\mathcal{B}(X,Y)} \inf_{Tx'=Tx}\|x'\|_X = \|T\|_{\mathcal{B}(X,Y)}\|Tx\|_Z, \quad \text{(A.29)}$$

hence $\|\cdot\|_Y \le \|T\|_{\mathcal{B}(X,Y)}\|\cdot\|_Z$. Trivially, $\|Tx\|_Z \le \|x\|_X$ for $x \in X$.

Assume now that X is a Hilbert space. Define $P \in \mathcal{B}(N^\perp, X/N)$ by $Px := x+N$. Then $\|Px\|_{X/N} = \|x\|_H$ for all $x \in N^\perp$. It follows that X/N is a Hilbert space (because so is $N^\perp$ and P is an isometric isomorphism of $N^\perp$ onto X/N), hence so is Z (with $\langle z, z'\rangle_Z := \langle x, x'\rangle_X$, where $\{x\} = T^{-1}(z) \cap N^\perp$ and $\{x'\} = T^{-1}(z') \cap N^\perp$). □

One sometimes wishes to define spaces such as $\mathrm{L}^1(\mathbf{R}) + \mathrm{L}^\infty(\mathbf{R})$; a necessary condition is, of course, that both original spaces lie in a single vector space. To guarantee that the resulting space is a normed space, we require a bit more.

Normed spaces X and Y are called *sum-compatible*, or (X,Y) is called a *sum-compatible pair*, if $X \underset{c}{\subset} Z$, $Y \underset{c}{\subset} Z$ for some TVS Z.

Lemma A.3.17 ($X+Y$) *Let X and Y be sum-compatible normed spaces. Then $X+Y$ is a normed space with norm $\|z\|_{X+Y} := \inf_{z=x+y}(\|x\|_X + \|y\|_Y)$.*

This operation is commutative and associative: $X+Y = Y+X$ and $(X+Y)+X_2 = X+Y(+X_2)$, isometrically, when X_2 is sum-compatible with X and with Y. Moreover, $X \underset{c}{\subset} X+Y$ and $Y \underset{c}{\subset} X+Y$.

If X and Y are Hilbert spaces, then so is $X+Y$ (with an equivalent norm). If X and Y are complete, then so is $X+Y$.

The following are equivalent:

(i) X and Y are sum-compatible;

(ii) X and Y are vector subspaces of some third vector space, and $\|\cdot\|_{X+Y}$ is a norm on $X+Y$.

Indeed, (ii) implies that we can take $Z = X+Y$ in (i), and the above lemma provides the converse (i)⇒(ii). Thus, $X+Y$ is well defined iff X and Y are compatible.

Proof: This is contained in Lemma 2.3.1 of [BL], except the claims on commutativity, associativity and embeddings, which are obvious, and the claim on Hilbert spaces, which we prove below:

One easily verifies that $T(z,z') := z+z'$ defines an element of $T \in \mathcal{B}(X\times Y, X+Y)$. The norm of $X+Y = [T(X\times Y)]$ is clearly that coinduced by T (see Lemma A.3.16), hence $X+Y$ is a Hilbert space, by Lemma A.3.16. □

Lemma A.3.18 ($\boldsymbol{X\cap Y}$) *Let X and Y be normed spaces that are subspaces of some vector space Z. Then $X\cap Y$ is a normed space with norm $\|x\|_{X\cap Y} :=$* $\max(\|x\|_X, \|x\|_Y)$.

This operation is commutative and associative: $X\cap Y = Y\cap X$ and $(X\cap Y)\cap X_2 = X\cap(Y\cap X_2)$, isometrically, when $X_2 \subset Z$. Moreover, $X\cap Y \underset{c}{\subset} X$ and $X\cap Y \underset{c}{\subset} Y$.

If X and Y are inner product spaces, then so is $X\cap Y$ (with inner product $\langle x,y\rangle_X + \langle x,y\rangle_Y$ and the corresponding norm $\|x\| := (\|x\|_X^2 + \|x\|_Y^2)^{1/2}$, equivalent to $\max\{\|x\|_X, \|x\|_Y\}$*).*

If $x = y$ whenever $x_n \to x$ in X and $x_n \to y$ in Y (this holds whenever Z is a TVS and $X \underset{c}{\subset} Z$ and $Y \underset{c}{\subset} Z$), and X and Y are complete, then $X\cap Y$ is complete.

Thus, also the spaces $X := \mathrm{L}^p(Q;B)$ and $Y := \mathrm{L}^q(Q;B)$ of Definition B.3.1 will do when $p,q \in [1,\infty]$ (if limits in L^p and L^q must be equal a.e., hence equal as elements of $\mathrm{L}^p \cap \mathrm{L}^q$, by Theorem B.3.2.

Proof: All these claims are quite obvious (cf. Lemma 2.3.1 of [BL]). (The inclusions to Z make sum, scalar multiplication and $X\cap Y$ well defined.) □

Lemma A.3.19 *Let X_1 and X_2 be sum-compatible normed spaces, and let Y be a normed space.*

(a) If $X_1 \underset{c}{\subset} X_2$, then $X_1 \cap X_2 = X_1$ and $X_1 + X_2 = X_2$ as TVSs (i.e., with equivalent norms).

(b) If $X_0 \subset X_1 \cap X_2$ is dense in X_1 and in X_2, then X_0 is dense in $X_1 + X_2$.

(c1) If $X_k \underset{c}{\subset} Y$ ($k = 1,2$), then $X_1 + X_2 \underset{c}{\subset} Y$.

(c2) If $Y \underset{c}{\subset} X_k$ $(k = 1, 2)$, then $Y \underset{c}{\subset} X_1 \cap X_2$.

Note that the assumptions of (a) and (c1) imply the sum-compatibility of X_1 and X_2.

Proof: (a) Now $\|\cdot\|_{X_2} \le M\|\cdot\|_{X_1}$ on X_1 for some $M \in [1,\infty)$, hence $\|x\|_{X_1} \le \|x\|_{X_1 \cap X_2} \le M\|x\|_{X_1}$. Obviously, $\|z\|_{X_1+X_2} \le \|z\|_{X_2}$. But if $z = x + y$, then $\|z\|_{X_2} \le M\|x\|_{X_1} + \|y\|_{X_2}$. Therefore, $\|z\|_{X_2} \le M\|z\|_{X_1+X_2}$.

(b) Let $x + y \in X_1 + X_2$. Then $\|(x+y) - (x'+y')\|_{X_1+X_2} \le \|x - x'\|_{X_1} + \|y - y'\|_{X_2}$ for any $x', y' \in X_0$.

(c1)&(c2) These are quite obvious too. □

We deduce that sum-compatibility is *not* an equivalence relation, because it is not transitive: Set $Y := \mathrm{span}(\{x_n\}_{n \in \mathbf{N}}) = \mathrm{c_c}(\mathbf{N})$, where $x_n := ne_0 + e_n$ $(n \in \mathbf{N})$, with the norm $\|\sum_k \alpha_k x_k\|_Y := \sum_k |\alpha_k|^2$ (thus Y becomes isometrically isomorphic a subspace of $\ell^2(\mathbf{N})$). Set $X := \ell^2(\mathbf{N})$. Then $V := \mathrm{span}(\{e_7\})$ is sum-compatible with X and Y, but X is not sum-compatible with Y: The function $\|\cdot\|_{X+Y}$ is not a norm on $X+Y$, because $\|e_0\|_{X+Y} = n^{-1}\|x_n - e_n\|_{X+Y} \le 2/n$ for all $n \in \mathbf{N}+1$, i.e., $\|e_0\| = 0$, although $e_0 \ne 0$. (One can also replace Y by its completion; then all three spaces become Hilbert spaces.)

Remark A.3.20 ($(\mathbf{K} = \mathbf{C}) \mapsto (\mathbf{K} = \mathbf{R})$) *Complex normed spaces can be considered as real normed spaces having same topological properties, as shown in Lemma A.3.21. It follows that several results on real normed spaces can be applied to complex normed spaces; in particular, we obtain results for duals and adjoints (both $()^{\mathrm{d}}$ and $()^{\mathrm{B}}$ (and $()^{\mathrm{H}}$), see Remark A.3.22) on complex normed spaces from results for duals and adjoints given on real normed spaces.*

We apply this fact without further mention (this has been done in Lemmas A.3.1 and A.3.5). □

Here $()^{\mathrm{B}}$ (resp. $()^{\mathrm{d}}$) refers to standard (resp. "conjugate-linear") duals and adjoints in Banach spaces; $()^{\mathrm{H}}$ to the adjoints and duals in Hilbert spaces (thus $()^{\mathrm{H}}$ coincides with $()^{\mathrm{d}}$ on Hilbert spaces when we identify a Hilbert space and its dual).

Analogously, a complex vector space is also a real vector space. Next we record the invariance of certain properties under this modification:

Lemma A.3.21 ($X_{\mathbf{R}}$) *Let X be a complex vector space. Let $X_{\mathbf{R}}$ be X with scalars restricted to $\mathbf{R}$. Then $X_{\mathbf{R}}$ is a vector space.*

Assume, in addition, that X and Y are TVSs. Then $X_{\mathbf{R}}$ is a TVS, $X_{\mathbf{R}}^* = \mathrm{Re}\, X^* = \mathrm{Re}\, X^{\mathrm{d}}$ and $\mathcal{B}(X,Y) \subset \mathcal{B}(X_{\mathbf{R}}, Y_{\mathbf{R}})$. If X is normed, then the same norm is a norm on X too and $\|\mathrm{Re}\,\Lambda\| = \|\Lambda\|$ for all $\Lambda \in X^*$, in particular, $X_{\mathbf{R}}^* = \mathrm{Re}\, X^* = \mathrm{Re}\, X^{\mathrm{d}}$ isometrically. If X and Y are normed, then $\mathcal{B}(X,Y) \subset \mathcal{B}(X_{\mathbf{R}}, Y_{\mathbf{R}})$ linearly and isometrically; thus, if $T \in \mathcal{B}(X,Y)$, then $T_{\mathbf{R}} := T \in \mathcal{B}(X_{\mathbf{R}}, Y_{\mathbf{R}})$ and $\|T_{\mathbf{R}}\|_{\mathcal{B}(X_{\mathbf{R}};Y_{\mathbf{R}})} = \|T\|_{\mathcal{B}(X;Y)}$; moreover, then $(T^*)_{\mathbf{R}} = (T^{\mathrm{d}})_{\mathbf{R}}$ (i.e., $T^* = T^{\mathrm{d}}$ as elements of $\mathcal{B}(Y_{\mathbf{R}}^*, X_{\mathbf{R}}^*)$).

Moreover, X is complete (resp. metrizable, separable, normable, reflexive) iff $X_{\mathbf{R}}$ is. If $\langle\cdot,\cdot\rangle$ is an inner product in X, then $\mathrm{Re}\langle\cdot,\cdot\rangle$ is an inner product in

$X_{\mathbf{R}}$. *Finally,* $(X_{\mathbf{R}})_{\mathbf{R}} = X_{\mathbf{R}}$ *and* $\mathbf{C}_{\mathbf{R}} = \mathbf{R}^2$ *as TVSs, but* $\mathcal{B}(\mathbf{C}_{\mathbf{R}}) = \mathbf{R}^{2\times 2}$ *whereas* $\mathcal{B}(\mathbf{C}) = \mathbf{C}$.

Analogously, if Y is a real vector space (resp. [separable] TVS, normed space, Banach space, Hilbert space), then so is $Y_{\mathbf{C}} := Y + iY$ (with the topology and addition of Y^2 and natural scalar multiplication $(i(y+iy') := -y' + iy)$), but we do not need this.

We note that if X is normed, then $(X_{\mathbf{R}})^* = (X^*)_{\mathbf{R}} = (X^{\mathrm{d}})_{\mathbf{R}}$ as Banach spaces (if we identify Λ and $\operatorname{Re}\Lambda$); analogously, $X_{\mathbf{R}}^{**} = (X^*)_{\mathbf{R}}^* = (X^{**})_{\mathbf{R}}$, which establishes the reflexivity claim.

Proof of Lemma A.3.21: It is not hard to verify the above claims, some of which are given in Section 3.1 of [Rud73].

By $X_{\mathbf{R}}^* = \operatorname{Re} X^* = \operatorname{Re} X^{\mathrm{d}}$ we mean that $X_{\mathbf{R}}^* = \{\operatorname{Re}\Lambda \,\big|\, \Lambda \in X^* =: X^{\mathrm{B}}\} = \{\operatorname{Re}\Lambda \,\big|\, \Lambda \in X^{\mathrm{d}}\}$; cf. Remark A.3.22 ($X$ need not be normed; the identity map between vector spaces X^{B} and X^{d} is nevertheless a conjugate-linear bijection, hence the identity map between vector spaces $\operatorname{Re} X^{\mathrm{B}}$ and $\operatorname{Re} X^{\mathrm{d}}$ is a linear bijection, i.e., a vector space isomorphism (an isometric Banach space isomorphism if X is normed)). □

As noted above, the Hilbert space adjoint $T^{\mathrm{H}} \in \mathcal{B}(Y,H)$ ("T^*") of a bounded linear operator $T \in \mathcal{B}(H,Y)$ is identical to the Banach space adjoint $T^{\mathrm{B}} \in \mathcal{B}(Y^*,H^*)$ ("T^*") when we identify the Hilbert spaces with their duals; however, this identification does not preserve scalar multiplication (in the nontrivial case with $\mathbf{K} = \mathbf{C}$): $(\alpha T)^{\mathrm{H}} = \bar{\alpha} T^{\mathrm{H}}$, whereas $(\alpha T)^{\mathrm{B}} = \alpha T^{\mathrm{B}}$.

But this follows from the fact that multiplication by α in Y corresponds to multiplication by $\bar{\alpha}$ in Y^*; by defining the multiplication in Y^* by the latter, we can make these two concepts identical:

Remark A.3.22 (Conjugate-linear dual X^{d} (or X^*)) *Let X and Y be normed spaces. Usually, the dual $X^* =: X^{\mathrm{B}}$ is equipped with scalar multiplication $(\alpha\Lambda)x := \alpha(\Lambda x)$ ($\alpha \in \mathbf{K}$, $\Lambda \in X^*$, $x \in X$). However, X^* becomes a Banach space (denoted by X^{d}) also with the "conjugate-linear" scalar multiplication $(\alpha\Lambda)x := \bar{\alpha}(\Lambda x)$, i.e.,*

$$(\bar{\alpha}\Lambda)x =: \langle x, \bar{\alpha}\Lambda\rangle_{\langle X,X^{\mathrm{d}}\rangle} := \langle \alpha x, \Lambda\rangle_{\langle X,X^{\mathrm{d}}\rangle} = \alpha\langle x,\Lambda\rangle_{\langle X,X^{\mathrm{d}}\rangle} := \alpha(\Lambda x) =: \alpha\overline{\langle \Lambda, x\rangle_{\langle X^{\mathrm{d}},X\rangle}} \tag{A.30}$$

for all $x \in X$, $\Lambda \in X^*$, $\alpha \in \mathbf{K}$ (note that we have set $\langle \Lambda, x\rangle_{\langle X^{\mathrm{d}},X\rangle} := \overline{\Lambda x} = \overline{\langle x,\Lambda\rangle_{\langle X,X^{\mathrm{d}}\rangle}}$).

Then the identity mapping $I := I_{X^{\mathrm{d}}} : X^{\mathrm{d}} \to X^{\mathrm{B}}$ is a conjugate-linear (i.e., $I(\Lambda + \Lambda') = I\Lambda + I\Lambda'$, but $I\alpha\Lambda = \bar{\alpha}I\Lambda$) isometry of X^{d} onto X^{B}. In particular, all set-theoretic, topological and metric properties of X^{B} and X^{d} are identical. Moreover, if $T \in \mathcal{B}(X,Y)$ and we set

$$\langle x, T^{\mathrm{d}} y^{\mathrm{d}}\rangle_{\langle X,X^{\mathrm{d}}\rangle} := \langle Tx, y^{\mathrm{d}}\rangle_{\langle Y,Y^{\mathrm{d}}\rangle} \quad (x \in X,\ y^{\mathrm{d}} \in Y^{\mathrm{d}}), \tag{A.31}$$

then $T^{\mathrm{d}} \in \mathcal{B}(Y^{\mathrm{d}},X^{\mathrm{d}})$, $\|T^*\|_{\mathcal{B}(Y^{\mathrm{B}},X^{\mathrm{B}})} = \|T^{\mathrm{d}}\|_{\mathcal{B}(Y^{\mathrm{d}},X^{\mathrm{d}})} = \|T\|_{\mathcal{B}(Y,X)}$ and $T^{\mathrm{d}} = I_{X^{\mathrm{d}}}^* T^* I_{Y^{\mathrm{d}}}$. Thus, $\mathcal{B}(Y^{\mathrm{B}},X^{\mathrm{B}})$ and $\mathcal{B}(Y^{\mathrm{d}},X^{\mathrm{d}})$ are isometrically isomorphic.

The conjugate-linear adjoint satisfies algebraic laws $(ST)^{\mathrm{d}} = T^{\mathrm{d}}S^{\mathrm{d}}$ *and* $(\alpha S + \beta T)^{\mathrm{d}} = \bar{\alpha}S^{\mathrm{d}} + \bar{\beta}T^{\mathrm{d}}$.

The map $x \mapsto (\Lambda \mapsto \overline{\Lambda x})$ *is a (canonical) linear isometry* $X \to X^{\mathrm{dd}} := (X^{\mathrm{d}})^{\mathrm{d}}$ *(cf. the canonical linear isometry* $X \ni x \mapsto (\Lambda \mapsto \Lambda x) \in X^{\mathrm{BB}} := (X^{\mathrm{B}})^{\mathrm{B}}$*). If* X *is reflexive, then* $X^{\mathrm{dd}} := (X^{\mathrm{d}})^{\mathrm{d}}$ *is isometrically isomorphic to* X *(and we identify the two; analogously, we identify* X^{BB} *and* X *through the latter isometry when we use "linear duals").*

If X *is a Hilbert space, then* X^{d} *becomes isometrically isomorphic to* X *(and we identify the two). In particular, if* X *and* Y *are Hilbert spaces and* $T \in \mathcal{B}(X,Y)$*, then* T^{d} *becomes the Hilbert space adjoint of* T. □

Thus, $T^{\mathrm{d}} = T^{\mathrm{B}}$ except that their (otherwise equal) domain and range spaces have differently defined scalar multiplication.

The name "conjugate-linear dual" is somewhat misleading: we stress that also X^{d} is an ordinary Banach space over $\mathbf{K}$; in particular, the scalar multiplication (and addition) satisfies standard vector space axioms. The choice $(\alpha\Lambda)x := \alpha\Lambda x$ is certainly not the only one to provide a vector space structure for X^*, rather it is an unnatural one for any complex Hilbert space.

Note that the pairing $X \times X^{\mathrm{d}} \to \mathbf{K}$ becomes sesquilinear, i.e., it is linear in its first and conjugate-linear in its second argument; the same applies to $X^{\mathrm{d}} \times X \to \mathbf{K}$ (whereas $X \times X^* \to \mathbf{K}$ and $X^* \times X \to \mathbf{K}$ are bilinear).

In Hilbert space context we use Hilbert adjoints instead of Banach adjoints (i.e., $T^* := T^{\mathrm{d}}$ when T maps between Hilbert (or inner product) spaces; otherwise $T^* := T^{\mathrm{B}}$), following the standard convention. Usually, $X^* := X^{\mathrm{B}}$ (for normed spaces X), but in pivot space context we use sesquilinear pairings, i.e., $X^* := X^{\mathrm{d}}$; see Definition A.3.23 for details (this should be clear from the context). Of course, there is no difference when $\mathbf{K} = \mathbf{R}$.

Another common convention in infinite-dimensional control theory is to use pivot spaces (see Definition A.3.23), to have the pairing between spaces and their duals coincide to the inner product (or its extension) in a *pivot space*, and hence dependent only on the elements to be paired, not on the spaces.

To illustrate this, let $f \in \mathrm{L}^2_\omega(\mathbf{R})$, $g \in \mathrm{L}^2_{-\omega}(\mathbf{R}) = \mathrm{L}^2_\omega(\mathbf{R})^*$. Then we use $\mathrm{L}^2(\mathbf{R})$ as the *pivot space* by setting

$$\langle f, g\rangle_{\langle \mathrm{L}^2_\omega, \mathrm{L}^2_{-\omega}\rangle} := \int_{\mathbf{R}} \langle f(t), g(t)\rangle_{\mathbf{K}}\, dt := \int_{\mathbf{R}} \overline{g(t)}\, f(t)\, dt; \tag{A.32}$$

in particular, $\langle f, \alpha g\rangle = \bar{\alpha}\langle f, g\rangle$ (thus, αg corresponds to $\bar{\alpha}\Lambda$, where Λ is the corresponding element in the linear dual of L^2_ω). If $f \in \mathrm{L}^2_\omega \cap \mathrm{L}^2$ and $g \in \mathrm{L}^2_{-\omega} \cap \mathrm{L}^2$, then it follows that $\langle f, g\rangle_{\langle \mathrm{L}^2_\omega, \mathrm{L}^2_{-\omega}\rangle} = \langle f, g\rangle_{\mathrm{L}^2}$ (here the right-hand-side refers to inner product; or, equivalently, to the pairing $\mathrm{L}^2 \times (\mathrm{L}^2)^{\mathrm{d}} \to \mathbf{K}$).

Moreover, if by Y^{B} we denote the linear dual (i.e., the dual space in the ordinary sense) of $Y := \mathrm{L}^2_\omega$, then Y^{B} can be identified with $\mathrm{L}^2_{-\omega}$, by $Y \times Y^{\mathrm{B}} \ni (f, g) \mapsto \int_{\mathbf{R}} fg\, dm$. Thus, then the canonical identification $Y^{\mathrm{B}} \to Y^* := Y^{\mathrm{d}}$ becomes $Y^{\mathrm{B}} \ni g \mapsto \bar{g} \in Y^*$, whereas the identification $J : Y^* \to Y$ (corresponding to the inner product in Y instead of L^2) is given by $Jg := e^{2\omega\cdot} g$. Both these identifications are bijective isometries, but the former is conjugate-linear whereas the latter is linear (hence an isometric isomorphism), and neither is the isometric isomorphism (w.r.t. to the pivot space) illustrated above and defined below.

Thus, we take adjoints and duals w.r.t. a pivot space instead of initial and range spaces, hence only the pivot space remains the dual of itself. Formally this goes as follows:

Definition A.3.23 (Pivot space) *Let H be a Hilbert space. When we say that we use H as a* pivot space*, we mean the following:*

We use "conjugate-linear duals" (see Remark A.3.22) and Banach adjoints instead of Hilbert adjoints for any appropriate spaces.

If X is a normed space s.t. $X \cap H$ is dense in X, and if $y \in H$ is s.t. the operator $\langle \cdot, y \rangle_H : X \cap H \ni x \mapsto \langle x, y \rangle_H \in \mathbf{K}$ is continuous in the $\|\cdot\|_X$ norm, i.e.,

$$|\langle x, y \rangle_H| \le M_y \|x\|_X \quad (x \in X), \tag{A.33}$$

then we identify y with the unique continuous extension of $\langle \cdot, y \rangle_H$ in X^* (thus, then $y \in H \cap X^*$).

Thus, $H \cap X^*$ becomes identified with a dense subspace of X^*, and $\langle x, y \rangle_{\langle X, X^* \rangle} = \langle x, y \rangle_H$ for all $x \in X \cap H$, $y \in H \cap X^*$.

Some aspects of pivot spaces are treated in [Keu] with further details, but [Keu] only mentions the cases $X \subset H$ and $H \subset X$, hence it would require one to treat L^2_ω by parts (on $\mathbf{R}_\pm$).

Lemma A.3.24 *Definition A.3.23 is well posed.*

Let X and H be as in Definition A.3.23. Then $H^* = H$ (this is an isometric isomorphism, in particular, $\langle x, y \rangle_{\langle H, H^* \rangle} = \langle x, y \rangle_H$ for all $x, y \in H$), and the identification $H \cap X^* \to X^*$ is linear and one-to-one.

Assume that Y is a normed space and $H \underset{c}{\subset} Y$ densely. Then $Y^* \underset{c}{\subset} H$ densely and $H \underset{c}{\subset} Y \underset{c}{\subset} Y^{**}$.

Assume that $Z \underset{c}{\subset} H$ is a normed space and dense in H. Then $H \underset{c}{\subset} Z^*$. If Z is reflexive, then H is dense in Z^* and Z becomes identified to Z^{**} in the pivot space sense (with the canonical identification), thus, then $Z = Z^{**} \underset{c}{\subset} H \underset{c}{\subset} Z^* = Z^{***}$.

Assume that also $V \underset{c}{\subset} H$ is a dense (in H) normed space, and $T \in \mathcal{B}(H) \cap \mathcal{B}(V, Z)$. Then $T^* \in \mathcal{B}(Z^*, V^*) \cap \mathcal{B}(H)$. If $T = T^* \in \mathcal{B}(H) \cap \mathcal{B}(V, Z)$, then $T = T^* \in \mathcal{B}(V, Z) \cap \mathcal{B}(Z^*, V^*)$.

If, in addition, $Z \subset V$ densely, then $H \underset{c}{\subset} V^* \underset{c}{\subset} Z^*$.

By Lemma A.3.13, we have for any normed space N that if $T \in \mathcal{B}(N, Z)$, then $T^* \in \mathcal{B}(Z^*, N^*) \underset{c}{\subset} \mathcal{B}(H, N^*)$; if $T \in \mathcal{B}(Y, N)$, then $T^* \in \mathcal{B}(N^*, Y^*) \underset{c}{\subset} \mathcal{B}(N^*, H)$.

Note that for H, nothing has changed, but the duals of its proper (Hilbert) subspaces are no longer identified with their duals, although the two spaces are still isometrically isomorphic, by Remark A.3.22.

Following the standard convention, we shall use L^2 (resp. H) as the pivot space when taking adjoints in L^2_ω (resp. in $\mathrm{Dom}(A)$ or $\mathrm{Dom}(A^*)$, where A generates a C_0-semigroup in H). This is illustrated in (6.2) and Definition 6.1.17.

Finally we note that given a fixed pivot space H and its dense subspaces $B_1, B_2, V_1, V_2 \underset{c}{\subset} H$, and elements $w \in B_1 \cap B_2$, $w^* \in B_1^* \cap B_2^*$, we have $\langle w, w^* \rangle_{B_1,B_1^*} = \langle w, w^* \rangle_{B_2,B_2^*}$, i.e., the pairing is independent of the spaces chosen (as long as the pairing is defined). Therefore, the adjoints of an operator $F \in \mathcal{B}(B_1,V_1) \cap \mathcal{B}(B_2,V_2)$ (w.r.t. the pivot space H) coincide on $V_1^* \cap V_2^*$ (i.e., $\langle w, F^*v^* \rangle_{\langle B_1,B_1^* \rangle} := \langle Fw, v^* \rangle_{V_k,V_k^*} = \langle w, F^*v^* \rangle_{\langle B_2,B_2^* \rangle}$ for $k = 1,2$ and $v^* \in V_1^* \cap V_2^*$).

Proof of Lemma A.3.24: 1° By density, the identification of "$H \cap X^*$" to X^* is well defined; in particular, this identification is injective and linear (recall that $X^* = X^{\mathrm{d}}$).

2° *Case $H \underset{c}{\subset} Y$ densely:* By density, the canonical embedding $Y^* \subset H^* = H$ one-to-one. Obviously, it is linear and continuous. Because the identity (inclusion) operator $I \in \mathcal{B}(H,Y)$ is one-to-one (recall that inclusions are required to be one-to-one), the range of the canonical embedding $I^* \in \mathcal{B}(Y^*,H^*) = \mathcal{B}(Y^*,H)$ is dense in H, by Lemma A.3.4(N2) (replace Y by its completion; here $I^*\Lambda = \Lambda_{|H}$ for $\Lambda \in Y^*$; note that $\langle Ih, \Lambda \rangle_{\langle Y,Y^* \rangle} = \langle h, I^*\Lambda \rangle_{\langle H,H^* \rangle} = \langle h, I^*\Lambda \rangle_H$ for $h \in H,\ \Lambda \in Y^*$).

Consequently, $H \underset{c}{\subset} Y^{**}$, by Lemma A.3.13. By definition, $\langle Ih, y^* \rangle_{\langle Y,Y^* \rangle} = \langle h, I^*y^* \rangle_H = \langle Jh, y^* \rangle_{\langle Y,Y^* \rangle}$ for all $h \in H,\ y^* \in Y^*$, where J is the (identity) inclusion $H \to Y^{**}$ (of Definition A.3.23). Therefore, $Ih = Jh$ ($h \in H$), hence the completions Y and Y^{**} of H become equal, and the identification JI^{-1} extends to the linear isometry $T : Y \to Y^{**}$ that satisfies $\langle Iy, y^* \rangle_{\langle Y^{**},Y^* \rangle} = \langle y, y^* \rangle_{\langle Y,Y^* \rangle}$ for all $y \in Y,\ y^* \in Y^*$. Thus, the identification $H \to Y$ is the same whether we consider Y as itself or as a subspace of Y^{**}.

3° *Case $Z \underset{c}{\subset} H$ densely:* By Lemma A.3.13, we have $H \underset{c}{\subset} Z^*$. Assume that Z is reflexive. Then H is dense in Z^*, by Lemma A.3.4(N2), hence $Z^{**} \underset{c}{\subset} H$, by 1°. But the elements of Z^{**} are those $h \in H$, for which $|\langle z^*, h \rangle_H| \leq M \|z^*\|_{Z^*}$ for all $z^* \in H$ (i.e., those for which $\langle \cdot, h \rangle_H : H \to \mathbf{K}$ extends continuously to $Z^* \to \mathbf{K}$); reflexivity implies that this holds only for $h \in Z$, hence $Z = Z^{**}$ (and $\langle z^*, z \rangle_{\langle Z^*,Z^{**} \rangle} = \langle z^*, z \rangle_{\langle Z^*,Z \rangle}$ for all $z^* \in H$, hence for all $z^* \in Z^*$, by density). By 2°, we have $Z^* = Z^{***}$.

4° *Case $T \in \mathcal{B}(H) \cap \mathcal{B}(V,Z)$:* (By Lemma A.3.6, this is the case iff $T \in \mathcal{B}(H)$ is s.t. $T[V] \subset Z$.) If $z^* \in H$ and $x_0 \in V$, then

$$\langle x_0, T^*z^* \rangle_{\langle V,V^* \rangle} = \langle Tx_0, z^* \rangle_{\langle Z,Z^* \rangle} = \langle Tx_0, z^* \rangle_H = \langle x_0, T^{\mathrm{H}}z^* \rangle_H, \tag{A.34}$$

where $T^{\mathrm{H}} \in \mathcal{B}(H)$ (resp. $T^* \in \mathcal{B}(Z^*,V^*)$) denotes the adjoint of $T \in \mathcal{B}(H)$ (resp. $T \in \mathcal{B}(V,Z)$), hence then $T^*z^* = T^{\mathrm{H}}z^* \in H$. Thus, $T^*_{|H} = T^{\mathrm{H}}$, so that we can write $T^* \in \mathcal{B}(Z^*,V^*) \cap \mathcal{B}(H)$. The $T = T^*$ claim is obvious.

5° *$H \underset{c}{\subset} V^* \underset{c}{\subset} Z^*$:* By Lemma A.3.13, we have $\mathcal{B}(V,\mathbf{K}) \underset{c}{\subset} \mathcal{B}(Z,\mathbf{K})$. We have shown above that $H \underset{c}{\subset} V^*$. □

Notes

As indicated in the proofs, most results of this section are well known at least to some extent. Sum-compatibility, $X \cap Y$ and $X + Y$ are defined in [BL], which also contains parts of the corresponding lemmas.

For further results on Hilbert and Banach spaces or Banach algebras, see any textbook on functional analysis; standard references include [Yosida], [Rud73], [Rud86] and [HP].

A.4 C_0-Semigroups

Could Hamlet have been written by committee, or the Mona Lisa painted by a club? Creative ideas don't spring from groups. They spring from individuals.

— Alfred Whitney Griswold (1906–1963)

In this section, we present basic facts on strongly continous (C_0) semigroups. Throughout the section, the Hilbert and Banach spaces are assumed to be complex.

The solution of $x' = Ax$, $x(0) = x_0$ is $\mathrm{e}^{At}x_0$ for $x_0 \in B$, $A \in \mathcal{B}(B)$ and a Banach space B. For more general control systems, the operator A need not be bounded, hence the (C_0-)semigroup "$t \mapsto \mathrm{e}^{At}$" of a control system is usually more complicated:

Definition A.4.1 (Semigroups) *A* C_0-semigroup *on a complex Banach space* B *is a function* $\mathbb{A} : [0,\infty) \to \mathcal{B}(B)$ *having the following properties:*

$$\mathbb{A}(t+s) = \mathbb{A}(t)\mathbb{A}(s) \quad \text{for } t,s \geq 0; \tag{A.35}$$

$$\mathbb{A}(0) = I; \tag{A.36}$$

$$\|\mathbb{A}(t)x_0 - x_0\| \to 0 \quad \text{as } t \to 0+ \quad \forall x_0 \in B. \tag{A.37}$$

The (infinitesimal) generator A *of* $\mathbb{A}$ *is defined by*

$$Ax := \lim_{t\to 0+} \frac{1}{t}\left(\mathbb{A}(t) - I\right)x; \tag{A.38}$$

$\mathrm{Dom}(A) \subset B$ *is the set of* $x \in B$ *for which the limit* Ax *exists.*

If $s \in \mathbf{C}$ *is s.t.* $s - A := sI - A$ *maps* $\mathrm{Dom}(A)$ *one-to-one onto* B*, and the* resolvent $(s-A)^{-1}$ *is bounded* ($(s-A)^{-1} \in \mathcal{B}(B)$)*, then* $s \in \sigma(A)^c$*; this is the definition of the* spectrum $\sigma(A)$ *of* A.

We also define the growth rate $\omega_A := \inf_{t>0}\left(t^{-1}\log\|\mathbb{A}(t)\|\right)$ *of* $\mathbb{A}$.

Let $\omega \in \mathbf{R}$. *We call* $\mathbb{A}$ ω-stable *if* $t \mapsto \|\mathrm{e}^{-\omega t}\mathbb{A}(t)\|$ *is bounded on* $\mathbf{R}_+$. *We call* $\mathbb{A}$ strongly (resp. weakly) ω-stable *if* $\mathrm{e}^{-\omega t}\mathbb{A}(t) \to 0$ *strongly (resp. weakly) as* $t \to +\infty$. *We call* $\mathbb{A}$ exponentially stable *if* $\mathbb{A}$ *is* ω*-stable for some* $\omega < 0$. *By* [strongly/weakly] stable *we mean* [strongly/weakly] 0-stable.

The "C_0" condition (A.37) ("strongly continuous at 0") is equivalent to strong continuity (see (a1) below). In this monograph, all semigroups will be C_0-semigroups, hence we usually call them just semigroups. Sometimes we write $\mathbb{A}^t$ for $\mathbb{A}(t)$.

Note that A is the strong right derivative of $\mathbb{A}$ at 0, and recall that "$\mathbb{A}(t) \to 0$ strongly (resp. weakly)" means that $\mathbb{A}(t)x \to 0$ for all $x \in B$ ($\Lambda\mathbb{A}(t)x \to 0$ for all $x \in B$ and $\Lambda \in B^*$).

We need a number of facts on C_0-semigroups:

Lemma A.4.2 *Assume that* B *and* B_2 *are complex Banach spaces,* $\mathbb{A}$ *is a* C_0*-semigroup on* B*,* A *be the generator of* $\mathbb{A}$ *and* $t \geq 0$. *Then the following hold:*

(a1) $\mathbb{A}(\cdot)x \in \mathcal{C}([0,\infty);B)$ *for all* $x \in X$.

(a2) $A \in \mathcal{B}(B)$ *iff* $\mathbb{A}$ *is uniformly continuous (i.e., iff* $\mathbb{A} \in \mathcal{C}([0,\infty);\mathcal{B}(B))$*). If* $A \in \mathcal{B}(B)$*, then* $\mathbb{A}(t) = \mathrm{e}^{At}$ *(*$t \ge 0$*).*

(a3) $A \in \mathcal{B}(B)$ *iff* $\mathbb{A}$ *is (uniformly) Bochner-measurable (see Definition B.2.1).*

(a4) $\|\mathbb{A}(\cdot)\|$ *and* $\mathbb{A}f$ *are measurable for any measurable* $f : \mathbf{R}_+ \to B$*.*

(b) $\cap_{n\in\mathbf{N}} \mathrm{Dom}(A^n)$ *is dense in B.*

(c1) $\exists(\mathbb{A}x)'(t) = A\mathbb{A}(t)x = \mathbb{A}(t)Ax$ *for all* $x \in \mathrm{Dom}(A)$*.*

(c2) $\int_0^t \mathbb{A}(r)x\,dr \in \mathrm{Dom}(A)$ *and* $A\int_0^t \mathbb{A}(r)x\,dr = \mathbb{A}(t)x - x$ *for all* $x \in H$*. In particular,* $\int_0^t \mathbb{A}(r)\cdot dr \in \mathcal{B}(H,\mathrm{Dom}(A))$*.*

(c3) $\mathbb{A}(t)x = \lim_{s\to+\infty} \mathrm{e}^{tsA(s-A)^{-1}}x$ *for all* $x \in H$*.*

(c4) $\mathbb{A}^t[\mathrm{Dom}(A^n)] \subset \mathrm{Dom}(A^n)$ *for all* $n \in \mathbf{N}$*.*

(c5) $\mathbb{A}x \in \mathcal{C}^k(\mathbf{R}_+;\mathrm{Dom}(A^{n-k}))$ *and* $\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathbb{A}x = A^k\mathbb{A}x = \mathbb{A}A^kx$ *(*$n \in \mathbf{N}$*,* $k = 0,1,\ldots,n$*,* $x \in \mathrm{Dom}(A^n)$*).*

(d) The set $\sigma(A)$ *is closed, and* $\mathrm{Re}\,s > \omega_A \Rightarrow s \in \sigma(A)^c$*.*

(e1) $\omega_A = \lim_{t\to+\infty}\big(t^{-1}\log\|\mathbb{A}(t)\|\big) < \infty$*.*

(e2) If $\omega > \omega_A$*, then there is* $M_\omega > 0$ *s.t.* $\|\mathbb{A}(t)\| \le M_\omega \mathrm{e}^{\omega t}$ *for all* $t \ge 0$*.*

(f) If B is a Hilbert space, then $\mathbb{A}^*$ *is a* C_0*-semigroup on B and* A^* *is its generator.*

(g1) If $T \in \mathcal{B}(B)$*, then* $A + T$ *is a* C_0*-semigroup on B.*

(g2) The generator of $\mathrm{e}^{\alpha\cdot}\mathbb{A}(\cdot)$ *is* $\alpha I + A$ *with* $\mathrm{Dom}(\alpha + A) = \mathrm{Dom}(A)$*, and* $\omega_{\alpha I+A} = \alpha + \omega_A$*.*

(h1) If $T \in \mathcal{GB}(B,B_2)$*, then* $\widetilde{\mathbb{A}} := T\mathbb{A}T^{-1}$ *is a* C_0*-semigroup on* B_2*, and its generator* $\widetilde{A}$ *is given by* $\widetilde{A} = TAT^{-1}$*,* $\mathrm{Dom}(\widetilde{A}) = T\,\mathrm{Dom}(A)$*.*

Moreover, $M^{-1}\|\mathbb{A}\| \le \|\widetilde{\mathbb{A}}\| \le M\|\mathbb{A}\|$*, where* $M := \|T\|\|T^{-1}\|$*, hence* $\omega_{\widetilde{A}} = \omega_A$*. For* $\omega \in \mathbf{R}$*, the semigroup* $\widetilde{\mathbb{A}}$ *is [strongly (resp. weakly)]* ω*-stable iff* $\mathbb{A}$ *is.*

(h2) Let $R \in \mathcal{B}(B_2,B)$ *and* $L \in \mathcal{B}(B,B_2)$*. Then* $\mathbb{A}_{LR} := L\mathbb{A}R$ *is a* C_0*-semigroup on* B_2 *iff* $LR = I$ *and* $P\mathbb{A}^tP\mathbb{A}^sP = P\mathbb{A}^{t+s}P$ *for all* $t,s > 0$*, where* $P := RL \in \mathcal{B}(B)$*.*

Assume this. Then $\omega_{A_{LR}} \le \omega_A$ *and* $(s - A_{LR})^{-1} = L(s-A)^{-1}R$ *for* $s \in \mathbf{C}^+_{\omega_A}$*. Moreover, if* $y \in B_2$ *and* $Ry \in \mathrm{Dom}(A)$*, then* $y \in \mathrm{Dom}(A_{LR})$ *and* $A_{LR}y = LARy$*.*

(i) If also $\widetilde{\mathbb{A}}$ *is a* C_0*-semigroup on B and* $A \subset \widetilde{A}$*, then* $A = \widetilde{A}$*.*

By (c3), $\mathbb{A}$ is uniquely determined by A (the converse is trivial).

By (a3), we must not write $\big(\int_0^t \mathbb{A}(r)\,dt\big)\,x$ in (c2); see Appendix B for (Bochner) measurability and integration; alternatively, the strong integration theory of Section F.2 should be used.

By (d), we have $\omega_A \ge \sup\mathrm{Re}\,\sigma(A)$. For differentiable semigroups we have $\omega_A = \sup\mathrm{Re}\,\sigma(A)$, but this is not the case in general (see Example 5.14 of [CZ] for a counter-example).

Proof of Lemma A.4.2: The canonical references on semigroups are [Pazy] and [HP], but [Rud73], [CZ], [Prüss93] and [Sbook] also contain many of the above results. We only treat here the least known ones.

For (a3)&(a4) we note that on pp. 304–306 of [HP] it is stated that a uniformly measurable iff it is uniformly continuous (and that if $\mathbb{A}$ is weakly continuous on $(0,\infty)$, then it is strongly continuous on $(0,\infty)$ and hence strongly measurable on $\mathbf{R}_+$), and that for a C_0 semigroup $\mathbb{A}$, the function $\|\mathbb{A}(\cdot)\|$ is lower semicontinuous, hence measurable (the measurability of $\mathbb{A}f$ follows from Lemma F.1.3(a)).

We give below a sketch of the proofs for (h) and (i).

(h1) Now $t^{-1}(\mathbb{A}(t)x-x) \to 0 \Leftrightarrow t^{-1}(T\mathbb{A}(t)T^{-1}Tx-Tx) \to 0$, the claims on $\widetilde{A}$ follow from this. The claims on M are obvious; use them and (e1) for ω_A. The stability claims are again obvious.

(h2) If $L\mathbb{A}R$ is a semigroup and $t,s \geq 0$, then $L\mathbb{A}^t RL\mathbb{A}^s R = L\mathbb{A}^{t+s}R$, hence then $P\mathbb{A}^t P\mathbb{A}^s P = P\mathbb{A}^{t+s}P$. Conversely, assume that $P\mathbb{A}^t P\mathbb{A}^s P = P\mathbb{A}^{t+s}P$, hence $L\mathbb{A}^t RL\mathbb{A}^s R = L\mathbb{A}^{t+s}R$, for all $t,s>0$. Obviously, $\mathbb{A}^t_{LR}x = L\mathbb{A}^t Rx \to LRx = x$, as $t \to 0+$, hence $\mathbb{A}_{LR}$ has the C_0-property. (Note also that $P^2 = R(LR)L = P$, i.e., P is a projection.)

Obviously, $\omega_{A_{LR}} \leq \omega_A$. For $s \in \mathbf{C}^+_{\omega_A}$ and $x \in B$, we have $(s-A_{LR})^{-1}x = L\int_0^\infty e^{-st}\mathbb{A}^t Rx\,dt = L(s-A)^{-1}Rx$, by Lemma A.4.4(f).

If $y \in B_2$ and $Ry \in \mathrm{Dom}(A)$, then $Lt^{-1}[\mathbb{A}^t Ry - Ry] \to LARy$, hence then $y \in Dom(A)$ and $A_{LR}y = LARy$.

(i) By $A \subset \widetilde{A}$ we mean that $\mathrm{Dom}(A) \subset \mathrm{Dom}(\widetilde{A})$ and $A = \widetilde{A}$ on $\mathrm{Dom}(A)$. Let $\omega > \omega_A, \omega_{\widetilde{A}}$. Then $\omega I - \widetilde{A}$ maps $\mathrm{Dom}(A)$ one-to-one onto B and $\mathrm{Dom}(\widetilde{A})$ one-to-one onto B, hence $\mathrm{Dom}(\widetilde{A}) = \mathrm{Dom}(A)$, hence $A = \widetilde{A}$, hence $\mathbb{A} = \widetilde{\mathbb{A}}$, by (c3). □

The celebrated Hille–Yosida Theorem gives necessary and sufficient conditions for an operator to generate a C_0 semigroup:

Theorem A.4.3 (Hille–Yosida) *A linear operator $A : \mathrm{Dom}(A) \to B$ is the generator of a C_0-semigroup $\mathbb{A}$ s.t. $\|\mathbb{A}(t)\|_B \leq Me^{\omega t}$ $(t \geq 0)$ for some $M < \infty$ iff*

(i) A is closed and $\mathrm{Dom}(A)$ is dense in B;

(ii) $(\omega,\infty) \subset \sigma(A)^c$ and

$$\|(s-A)^{-n}\| \leq M(s-\omega)^{-n} \quad \textit{for all} \quad s>\omega, \; n = 1,2,3,\ldots \tag{A.39}$$

□

(See, e.g., Theorem I.5.3 of [Pazy] for the proof.)

The resolvents of the generators of semigroups has been studied extensively, here we list some important facts:

Lemma A.4.4 *Let $\mathbb{A}$ be a C_0-semigroup on a complex Banach space B, and let A be its generator. Define $H_1 := \mathrm{Dom}(A)$ and H_{-1} as in Lemma 6.1.16. Then, for all $x \in H$, the following holds:*

(a) $(s-A)^{-1}-(r-A)^{-1}=(r-s)(s-A)^{-1}(r-A)^{-1}\in \mathrm{H}(\mathbf{C}^+_{\omega_A};\mathcal{B}(H_{-1},H_1))$ *for a fixed* $r\in\sigma(A)^c$ *(this is called the* resolvent equation*).*

(b) $(s-A)^{-1}\in \mathrm{H}(\mathbf{C}^+_{\omega_A};\mathcal{B}(H,H_1))$.

(c1) $\|(s-A)^{-1}\|_{\mathcal{B}(H)}\le M/(\operatorname{Re} s-\omega_A)$ *(*$s\in\mathbf{C}^+_{\omega_A}$*) for some* $M<\infty$*.*

(c2) $\|s(s-A)^{-1}\|_{\mathcal{B}(H)}\le M_r$ *and* $\|(s-A)^{-1}\|_{\mathcal{B}(H,H_1)}\le M_r$ *(*$s\ge r$*), when* $r>\omega_A$*.*

(c3) $(s-A)^{-1}\in \mathrm{H}^\infty(\mathbf{C}^+_r;\mathcal{B}(H))\cap \mathrm{H}(\mathbf{C}^+_r;\mathcal{B}(H,H_1))\cap \mathrm{H}^2_{\text{strong}}(\mathbf{C}^+_r;\mathcal{B}(H))$ *for* $r>\omega_A$*.*

(d1) $\mathrm{Dom}(A)\ni s(s-A)^{-1}x\longrightarrow x$ *in* H*, as* $s\to+\infty$*.*

(d2) $H\ni A(s-A)^{-1}x\longrightarrow 0$ *in* H*, as* $s\to+\infty$*.*

(d3) $\mathrm{Dom}(A)\ni (s-A)^{-1}x\longrightarrow 0$ *in* $\mathrm{Dom}(A)$*, as* $s\to+\infty$*.*

(d4) $s(s-A)^{-1}(s-A)^{-1}x\to 0$ *in* $\mathrm{Dom}(A)$*, as* $s\to+\infty$*.*

(d5) $s(s-A)^{-1}s(s-A)^{-1}x\to x$ *in* H*, as* $s\to+\infty$*.*

(e1) The limits in (d1)–(d5) exist also as $\operatorname{Re} s\to+\infty$ *and* $s\in\{re^{i\theta}\,\big|\,r\ge 0,\ |\theta|<\pi/2-\varepsilon\}$ *for any* $\varepsilon>0$*.*

(e2) The limits in (d1)–(d5) exist also as $s\in\Sigma_{\theta,\omega}$*,* $|s|\to\infty$ *for some* $\theta>\pi/2$ *if the semigroup generated by* A *is analytic and* $\omega>\omega_A$*.*

(f) $(s-A)^{-1}x=\int_0^\infty e^{-st}\mathbb{A}(t)x\,dt$ *(*$s\in\mathbf{C}^+_{\omega_A}$*).*

(g) $A\mathbb{A}^t=\mathbb{A}'(t)=\mathbb{A}^tA$ *and* $A(s-A)^{-1}=s(s-A)^{-1}-I=(s-A)^{-1}A$ *on* $\mathrm{Dom}(A)$*, and* $(s-A)^{-1}\mathbb{A}^t=\mathbb{A}^t(s-A)^{-1}$ *and* $(s-A)^{-1}(r-A)^{-1}=(r-A)^{-1}(s-A)^{-1}$ *on* H*, for all* $s,r\in\sigma(A)^c$*,* $t\ge 0$*.*

We recall that $s\to+\infty$ refers to the limit at $+\infty$ along $\mathbf{R}$. Note that the results given for H can applied on H_1 and H_{-1} too if we restrict of extend A accordingly, because the three A's are isomorphic, as noted at Lemma 6.1.16. E.g., it follows from (a) that we also have $H\ni s(s-A)^{-1}x\longrightarrow x$, as $s\to+\infty$ for all $x\in H_{-1}$.

See Appendix D for holomorphic functions ($\mathrm{H}(\mathbf{C}^+_r)$).

Proof of Lemma A.4.4: (a) The resolvent equation (in $\mathcal{B}(H,H_1)$) is readily computed. The "$\in \mathrm{H}(\mathbf{C}^+_{\omega_A};\mathcal{B}(H_{-1},H_1))$" claim follows from (e).

(b) Equation (d1) implies that $\exists\ \frac{d}{ds}(s-A)^{-1}=-(s-A)^{-2}$, even in $\mathcal{B}(H,H_1)$ (use Lemma D.1.1).

(c1) This holds by the Hille–Yosida Theorem [Pazy, Theorem 1.5.3],

(c2) Because $s/(s-\omega_A)=1+\omega_A/(s-\omega_A)\le r/(r-\omega_A)$ ($s>r>\omega_A\ge 0$), and $s/(s-\omega_A)<1$ ($s>\omega_A<0$), we can set $M_r:=M\max\{1,r/(r-\omega_A)\}$ to obtain $\|s(s-A)^{-1}\|\le M_r$, by (c1). Because $A(s-A)^{-1}=s(s-A)^{-1}-I$, we can replace M_r by M_r+1 to obtain the second inequality.

(Obviously, we can allow s to belong to any sector $S_{r,T}:=\{s=r+z\,\big|\,\operatorname{Re} z>0,\ z/\operatorname{Re} z\le T\}$, or rectangular $R_{r,T}:=\{s\in\mathbf{C}^+_r\,\big|\,\operatorname{Im} s\le T\}$, where $r>\omega_A$, $T<\infty$.)

(c3) This follows from (a), (b) and Lemma A.4.5(i)&(v) (cf. Remark 6.1.9).

(d1)&(d2) Choose some $r>\omega_A$. Define $r_{x,s}:=\|x-s(s-A)^{-1}x\|_H=\|A(s-A)^{-1}x\|_H$. For $x\in H_1$ we have $r_{x,s}=\|(s-A)^{-1}Ax\|<M_r\|Ax\|/s\to 0$, by (c2), hence $r_{x,s}\to 0$ for all $x\in H$, by the uniform boundedness of $s(s-A)^{-1}$ and the density of $\mathrm{Dom}(A)$ in H. Thus (d1) and (d2) are true.

(d3) By (d2), $\|(s-A)^{-1}x\|_{\mathrm{Dom}(A)} := \|(s-A)^{-1}x\|_H + \|A(s-A)^{-1}x\|_H \to 0+0=0$.

(d4) This follows from (d3) and (c2).

(d5) Apply (d1) and Lemma A.3.1(j3) (see Lemma A.3.4(F1)).

(e1) This can be seen from the above proofs with slight modifications, because $s/\operatorname{Re} s = (\cos\theta)^{-1}$ is bounded.

(e2) The above proofs of (b1)–(b5) hold also when $s \in \Sigma_{\theta,\omega}$, $|s| \to \infty$ if the semigroup generated by A is analytic and θ is as in Lemma 9.4.2(a).

(f) See p. 20 of [Pazy].

(g) We obtain $(s-A)^{-1}\mathbb{A}^t = \mathbb{A}^t(s-A)^{-1}$ from (a) and Lemma A.4.2(c3). Use (a) and Lemma A.4.2(c1) for the others. □

We will also need the following "extended Datko's Theorem":

Lemma A.4.5 (Datko) *The following are equivalent for a C_0-semigroup $\mathbb{A}$ on H:*

(i) $\mathbb{A}$ is exponentially stable;

(ii) $\mathbb{A}(\cdot)x_0 \in \mathrm{L}^2(\mathbf{R}_+;H)$ for all $x_0 \in H$ (or equivalently, for all x_0 in a dense subset of H);

(iii) $\|\int_0^\infty \mathbb{A}(s)\phi(s)\,ds\|_H \le M\|\phi\|_2$ for all $\phi \in C_c^\infty((0,\infty);H)$;

(iv) $(s-A)^{-1} \in \mathrm{H}^\infty(\mathbf{C}^+;\mathcal{B}(H))$;

(v) $(s-A)^{-1} \in \mathrm{H}^2_{\mathrm{strong}}(\mathbf{C}^+;\mathcal{B}(H))$;

(vi) $\mathbb{A}^*$ satisfies some (hence all) of (i)–(vi).

Note also that $\mathbb{A}$ is exponentially stable iff $\mathbb{A}^*$ is. As the proof shows (see Theorem B.4.12), it is enough that (iii) holds for $\phi \in C_c^\infty((0,\infty);H_0)$, for some norming subspace $H_0 \subset H$. Also (ii) can be weakened: by Theorem 1.1 of [W88], it suffices that for some $p \in [1,\infty)$ we have $\int_0^\infty |\langle x_1, \mathbb{A}(t)x_0\rangle|^p\,dt < \infty$ for all $x_0, x_1 \in H$. See also [Sbook, Theorem 3.11.8] for a generalization.

Proof of Lemma A.4.5: 1° Obviously, $\mathbb{A}^*$ is exponentially stable iff $\mathbb{A}$ is; in particular, we only need to establish the equivalence between (i)–(v).

2° The equivalence (i)⇔(iv) is shown in [Prüss84] (and in Theorem 3.11.6 of [Sbook]); the equivalence (ii)⇔(v) follows from the Plancherel Theorem.

3° If (i) holds, then $\|\mathbb{A}\| \in \mathrm{L}^2$ (cf. Lemma A.4.2(a4)), hence then (ii) and (iii) hold. The implication (ii)⇒(i) is [Pazy, Theorem 4.4.1, p. 116], so we assume (iii) and prove (i) to complete the equivalence.

Let $x_0 \in H$ and set $f := \mathbb{A}(\cdot)^* x_0 : \mathbf{R}_+ \to H$. Then

$$|\int_0^\infty \langle f(s), \phi(s)\rangle_H\,ds| = |\int_0^\infty \langle x_0, \mathbb{A}(s)\phi(s)\rangle_H\,ds| = |\langle x_0, \int_0^\infty \mathbb{A}(s)\phi(s)\,ds\rangle_H| \le M\|x_0\|_H\|\phi\|_2 \tag{A.40}$$

for all $\phi \in C_c^\infty((0,\infty);H)$, hence $\|f\|_2 \le M\|x_0\|_H$, by Theorem B.4.12. Because $x_0 \in H$ was arbitrary, (ii) and hence (i) holds for $\mathbb{A}^*$; therefore, also $\mathbb{A}$ is exponentially stable.

4° *The "or equivalently" claim in (ii):* Assume that $\mathbb{A}x_0 \in \mathrm{L}^2$ for all $x_0 \in X$, where $X \subset H$ is dense. Then this map has a unique extension $\widetilde{\mathbb{A}} \in \mathcal{B}(H, \mathrm{L}^2(\mathbf{R};H))$, by Lemma A.3.10. Choose $\omega > \omega_A$, so that

$\mathbb{A} \in \mathcal{B}(H, \mathrm{L}^2_\omega(\mathbf{R}_+; H))$. Then $\widetilde{\mathbb{A}} = \mathbb{A}$ as elements of $\mathbb{A} \in \mathcal{B}(H, \mathrm{L}^2_\omega(\mathbf{R}_+; H))$, by density, hence (ii) holds ("for all $x_0 \in H$"). $\square$

We finish this section by presenting some standard conventions in control theory:

Lemma A.4.6 ($W \subset H \subset X^*$, $X \subset H \subset W^*$) *Let A generate a C_0-semigroup on a Hilbert space H. Fix $\alpha \in \sigma(A)^c$.*

Set $W := (\alpha - A)^{-1}H = \mathrm{Dom}(A)$ (with norm $\|(\alpha - A)\cdot\|_H$) and $X := (\bar{\alpha} - A^*)^{-1}H = \mathrm{Dom}(A^*)$ (with norm $\|(\bar{\alpha} - A^*)\cdot\|_H$). Then W^* can be identified with the completion of H w.r.t. $\|(\alpha - A)^{-1}\cdot\|_H$, and X^* can be identified with the completion of H w.r.t. $\|(\bar{\alpha} - A^*)^{-1}\cdot\|_H$; in particular, for any $w \in W$, $x \in W^*$ we have $\langle w, x\rangle_{W,W^*} = \langle(\alpha - A)w, (\bar{\alpha} - A^*)^{-1}x\rangle_H$ ($= \langle w, x\rangle_H$ when $x \in H$).

Moreover, (extended) $\alpha - A$ is an isometric isomorphism $W \to H$ and $H \to X^*$, and A (and its extension to H and restriction to $\mathrm{Dom}(A^2)$) generates isomorphic C_0-semigroups on W, H and X^*.

Furthermore, $\mathrm{Dom}(A) = \{x \in H \mid Ax \in H\}$, and $\beta - A \in \mathcal{GB}(W, H)$ for any $\beta \in \sigma(A^c)$. In particular all above spaces and their topologies are independent on $\alpha \in \sigma(A)^c$.

We shall set $H_1 := \mathrm{Dom}(A) := W$, $H^*_{-1} := W^*$ in Chapter 6.

Proof: This is well known (see Lemma A.4.6 or p. 532 of [Weiss-C] (or [S97b, Section 7] or [Sbook])), so only sketch part of the proof. By Lemma A.3.4(G3)&(G1), the norm on $W = \mathrm{Dom}(A)$ is equivalent to the graph norm on $\mathrm{Dom}(A)$ (in particular, W and its topology are independent on $\alpha \in \sigma(A)^c$), and W is a Banach space. In particular, $(\beta - A)^{-1} \in \mathcal{GB}(H, W)$ for any $\beta \in \sigma(A^c)$.

Let $\beta \in \sigma(A^c)$. Because $\beta - A \in \mathcal{GB}(W, H)$, we have $\beta - A^* \in \mathcal{GB}(H, W^*)$ (see Lemma A.3.24; thus $A^* \in \mathcal{B}(X, H) \cap \mathcal{B}(H, W^*)$). It follows that the norm of W^* becomes equivalent to $\|(\bar{\beta} - A^*)^{-1}\cdot\|_H$, hence W^* (as a TVS) is the completion of H w.r.t. this norm. The rest of the proof follows the same lines. $\square$

Notes

Most facts in this section are well known. The canonical references on semigroups are [Pazy] and [HP], but the list of suitable references for C_0-semigroup theory would be endless, including [Rud73], [CZ], [Prüss93] and [Sbook]. The notes for Chapter 3 of [Sbook] and those for Chapter 5 of [CZ] contain historical remarks on C_0-semigroups.