CHAPTER 9

Generalized Inverses of Linear Operators between Hilbert Spaces

1. Introduction

The observation that generalized inverses are like prose ("Good Heavens! For more than forty years I have been speaking prose without knowing it" – Molière, *Le Bourgois Gentilhomme*) is nowhere truer than in the literature of linear operators. In fact, generalized inverses of integral and differential operators were studied by Fredholm, Hilbert, Schmoidt, Bounitzky, Hurwitz, and others, before E. H. Moore introduced generalized inverses in an algebraic setting; see, e.g., the historic survey in Reid [**691**].

This chapter is a brief and biased introduction to generalized inverses of linear operators between Hilbert spaces, with special emphasis on the similarities to the finite–dimensional case. Thus the spectral theory of such operators is omitted.

Following the preliminaries in Section 2, generalized inverses are introduced in Section 3. Applications to integral and differential operators are sampled in Exs. 18–37. The minimization properties of generalized inverses are studied in Section 5. Integral and series representations of generalized inverses, and iterative methods for their computation are given in Section 6.

This chapter requires familiarity with the basic concepts of linear functional analysis, in particular, the theory of linear operators in Hilbert space.

2. Hilbert spaces and operators: Preliminaries and notation

In this section we have collected, for convenience, some preliminary results, which can be found, in the form stated here or in a more general form, in the standard texts on functional analysis; see, e.g., Taylor [**809**] and Yosida [**891**].

(A) Our Hilbert spaces will be denoted by $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$, etc. In each space, the *inner product* of two vectors $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$ and the *norm* is denoted by $\| \ \|$. The *closure* of a subset L of $\mathcal{H}$ will be denoted by $\overline{L}$ and its *orthogonal complement* by $L^\perp$. $L^\perp$ is a closed subspace of $\mathcal{H}$, and

$$L^\perp = \overline{L}^\perp \ .$$

The *sum*, $M + N$, of two subsets $M, N \subset \mathcal{H}$ is

$$M + N = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in M, \ \mathbf{y} \in N\} \ .$$

If M, N are subspaces of $\mathcal{H}$ and $M \cap N = \{\mathbf{0}\}$, then $M + N$ is called the *direct sum* of M and N, and denoted by $M \oplus N$. If in addition $M \subset N^\perp$ we denote their sum by $M \overset{\perp}{\oplus} N$ and call it the *orthogonal direct sum* of M and N. Even if the subspaces M, N are closed, their sum $M + N$ need not be closed; see, e.g., Ex. 1. An orthogonal direct sum of two closed subspaces is closed. Conversely, if L, M are closed subspaces of $\mathcal{H}$ and $M \subset L$, then

$$L = M \overset{\perp}{\oplus} (L \cap M^\perp) \ . \tag{1}$$

If (1) holds for two subspaces $M \subset L$, we say that L is *decomposable with respect to* M. See Exs. 5–6.

(B) The *(Cartesian) product* of $\mathcal{H}_1, \mathcal{H}_2$ will be denoted by

$$\mathcal{H}_{1,2} = \mathcal{H}_1 \times \mathcal{H}_2 = \{\{\mathbf{x}, \mathbf{y}\} : \mathbf{x} \in \mathcal{H}_1, \mathbf{y} \in \mathcal{H}_2\}$$

where $\{\mathbf{x}, \mathbf{y}\}$ is an ordered pair. $\mathcal{H}_{1,2}$ is a Hilbert space with inner product

$$\langle \{\mathbf{x}_1, \mathbf{y}_1\}, \{\mathbf{x}_2, \mathbf{y}_2\} \rangle = \langle \mathbf{x}_1, \mathbf{y}_1 \rangle + \langle \mathbf{x}_2, \mathbf{y}_2 \rangle \ .$$

Let $J_i : \mathcal{H}_i \to \mathcal{H}_{1,2}$, $i = 1, 2$ be defined by

$$J_1\mathbf{x} = \{\mathbf{x}, \mathbf{0}\} \quad \text{for all } \mathbf{x} \in \mathcal{H}_1$$

and

$$J_2\mathbf{y} = \{\mathbf{0}, \mathbf{y}\} \quad \text{for all } \mathbf{y} \in \mathcal{H}_2 \; .$$

The transformations J_1 and J_2 are isometric isomorphisms, mapping $\mathcal{H}_1$ and $\mathcal{H}_2$ onto

$$\mathcal{H}_{1,0} = J_1\mathcal{H}_1 = \mathcal{H}_1 \times \{\mathbf{0}\}$$

and

$$\mathcal{H}_{0,2} = J_2\mathcal{H}_2 = \{\mathbf{0}\} \times \mathcal{H}_2 \; ,$$

respectively. Here $\{\mathbf{0}\}$ is an appropriate zero space.

(C) Let $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ denote the *class of linear operators from* $\mathcal{H}_1$ *to* $\mathcal{H}_2$. In what follows we will use *operator* to mean a linear operator. For any $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ we denote the *domain* of T by $D(T)$, the *range* of T by $R(T)$, the *null space* of T by $N(T)$, and the *carrier* of T by $C(T)$, where

$$C(T) = D(T) \cap N(T)^{\perp} \; . \tag{2}$$

The *graph*, $G(T)$, of a $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is

$$G(T) = \{\{\mathbf{x}, T\mathbf{x}\} : \mathbf{x} \in D(T)\} \; .$$

Clearly, $G(T)$ is a subspace of $\mathcal{H}_{1,2}$, and $G(T) \cap \mathcal{H}_{0,2} = \{\mathbf{0}, \mathbf{0}\}$. Conversely, if G is a subspace of $\mathcal{H}_{1,2}$ and $G(T) \cap \mathcal{H}_{0,2} = \{\mathbf{0}, \mathbf{0}\}$, then G is the graph of a unique $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, defined for any point $\mathbf{x}$ in its domain

$$D(T) = J_1^{-1} P_{\mathcal{H}_{1,0}} G(T)$$

by

$$T\mathbf{x} = \mathbf{y} \; ,$$

where $\mathbf{y}$ is the unique vector in $\mathcal{H}_2$ such that $\{\mathbf{x}, \mathbf{y}\} \in G$, and $P_{\mathcal{H}_{1,0}}$ is the orthogonal projector: $\mathcal{H}_{1,2} \to \mathcal{A}_{1,0}$, see (L) below.

Similarly, for any $T \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ the *inverse graph* of T, $G^{-1}(T)$, is defined by

$$G^{-1}(T) = \{\{T\mathbf{y}, \mathbf{y}\} : \mathbf{y} \in D(T)\} \; .$$

A subspace G in $\mathcal{H}_{1,2}$ is an inverse graph of some $T \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ if and only if $G \cap \mathcal{H}_{1,0} = \{\mathbf{0}, \mathbf{0}\}$, in which case T is uniquely determined by G (von Neumann [**848**]).

(D) An operator $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is called *closed* if $G(T)$ is a closed subspace of $\mathcal{H}_{1,2}$. Equivalently, T is closed if

$$\mathbf{x}_n \in D(T),\ \mathbf{x}_n \to \mathbf{x}_0,\ T\mathbf{x}_n \to \mathbf{y}_0 \implies \mathbf{x}_0 \in D(T) \text{ and } T\mathbf{x}_0 = \mathbf{y}_0$$

where $\to$ denotes strong convergence. A closed operator has a closed null space. The subclass of *closed operators* in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is denoted by $\mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$.

(E) An operator $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is called *bounded* if its norm $\|T\|$ is finite, where

$$\|T\| = \sup_{0 \neq \mathbf{x} \in \mathcal{H}_1} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} \; .$$

The subclass of *bounded operators* in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is denoted by $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$. If $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, then it may be assumed, without loss of generality, that $D(T)$ is closed or even that $D(T) = \mathcal{H}_1$. A bounded $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ is closed if and only if $D(T)$ is closed. Thus we may write $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2) \subset \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$. Conversely, a closed $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$ is bounded if $D(T) = \mathcal{H}_1$. This statement is the *closed graph theorem*.

(F) Let $T_1, T_2 \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ with $D(T_1) \subset D(T_2)$. If $T_2\mathbf{x} = T_1\mathbf{x}$ for all $\mathbf{x} \in D(T_1)$, then T_2 is called an *extension* of T_1 and T_1 is called a *restriction* of T_2. These relations are denoted by

$$T_1 \subset T_2$$

or by

$$T_1 = (T_2)_{[D(T_1)]} \ .$$

Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and let the restriction of T to $C(T)$ be denoted by T_0

$$T_0 = T_{[C(T)]} \ .$$

Then

$$G(T_0) = \{\{\mathbf{x}, T\mathbf{x}\} : \mathbf{x} \in C(T)\}$$

satisfies

$$G(T_0) \cap \mathcal{H}_{1,0} = \{\mathbf{0}, \mathbf{0}\}$$

and hence is the inverse graph of an operator $S \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ with

$$D(S) = R(T_0) \ .$$

Clearly,

$$ST\mathbf{x} = \mathbf{x} \quad \text{for all } \mathbf{x} \in C(T) \ ,$$

and

$$TS\mathbf{y} = \mathbf{y} \quad \text{for all } \mathbf{y} \in R(T_0) \ .$$

Thus, if T_0 is considered as an operator in $\mathcal{L}(\overline{C(T)}, \overline{R(T_0)})$, then T_0 is invertible in its domain. The inverse T_0^{-1} is closed if and only if T_0 is closed. For $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, both $C(T)$ and T_0 nay be trivial; see, e.g., Exs. 2 and 4.

(G) An operator $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is called *dense* (or *densely defined*) if $\overline{D(T)} = \mathcal{H}_1$. Since any $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ can be viewed as an element of $T \in \mathcal{L}(\overline{D(T)}, \mathcal{H}_2)$, any operator can be assumed to be dense without loss of generality.

For any $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, the condition $\overline{D(T)} = \mathcal{H}_1$ is equivalent to

$$G(T)^{\perp} \cap \mathcal{H}_{1,0} = \{\mathbf{0}, \mathbf{0}\} \ ,$$

where

$$G(T)^{\perp} = \{\{\mathbf{y}, \mathbf{z}\} : \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{z}, T\mathbf{x} \rangle = 0 \text{ for all } \mathbf{x} \in D(T)\} \subset \mathcal{H}_{1,2} \ .$$

Thus for any dense $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, $G(T)^{\perp}$ is the inverse graph of a unique operator in $\mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$. This operator is $-T^*$, where T^*, the *adjoint* of T, satisfies

$$\langle T^*\mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{y}, T\mathbf{x} \rangle \quad \text{for all } \mathbf{x} \in D(T) \ .$$

(H) For any dense $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$,

$$\overline{N(T)} = R(T^*)^{\perp} \ , \quad N(T^*) = R(T)^{\perp} \ . \tag{3}$$

In particular, T $[T^*]$ has a dense range if and only if T^* $[T]$ is one–to–one.

(I) Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ be dense.

If both T and T^* have inverses, then $(T^{-1})^* = (T^*)^{-1}$.

T has a bounded inverse if and only if $R(T^*) = \mathcal{H}_1$.

T^* has a bounded inverse if $R(T) = \mathcal{H}_2$. The converse holds if T is closed.

T^* has a bounded inverse and $R(T^*) = \mathcal{H}_1$ if and only if T has a bounded inverse and $\overline{R(T)} = \mathcal{H}_1$ (Taylor [**809**], Goldberg [**294**]).

(J) An operator $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is called *closable* (or *preclosed*) if T has a closed extension. Equivalently, T is closable if

$$\overline{G(T)} \cap \mathcal{H}_{0,2} = \{\mathbf{0}, \mathbf{0}\} ,$$

in which case $\overline{G(T)}$ is the graph of an operator $\overline{T}$, called the *closure* of T. $\overline{T}$ is the minimal closed extension of T.

Since $G(T)^{\perp\perp} = \overline{G(T)}$ it follows that for a dense T, T^{**} is defined only if T is closable, in which case

$$T \subset T^{**} = \overline{T}$$

and

$$T = T^{**}$$

if and only if T is closed.

(K) A dense operator $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ is called *symmetric* if

$$T \subset T^*$$

and *self–adjoint* if

$$T = T^* ,$$

in which case it is called *non–negative*, and denoted by $T \succcurlyeq O$, if

$$\langle T\mathbf{x}, \mathbf{x} \rangle \geq 0 \quad \text{for all } \mathbf{x} \in D(T) .$$

If $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$ is dense, then T^*T and TT^* are non–negative, and $I + TT^*$ and $I + T^*T$ have bounded inverses (von Neumann [**846**]).

(L) An operator $T \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ is an *orthogonal projector* if

$$P = P^* = P^2 ,$$

in which case $R(P)$ is closed and

$$\mathcal{H} = R(P) \stackrel{\perp}{\oplus} N(P) .$$

Conversely, if L is a closed subspace of $\mathcal{H}$, then there is a unique orthogonal projector P_L such that

$$L = R(P_L) \quad \text{and} \quad L^\perp = N(P_L) .$$

(M) An operator $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$ is called *normally solvable* if $R(T)$ is closed, which, by (3), is equivalent to the following condition: The equation

$$T\mathbf{x} = \mathbf{y}$$

is consistent if and only if $\mathbf{y}$ is orthogonal to any solution $\mathbf{u}$ of

$$T^*\mathbf{u} = \mathbf{0} .$$

This condition accounts for the name "normally solvable".

For any $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$, the following statements are equivalent:

(a) T is normally solvable.

(b) The restriction $T_0 = T_{[C(T)]}$ has a bounded inverse.

(c) The non–negative number

$$\gamma(T) = \inf \left\{ \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} : \mathbf{0} \neq \mathbf{x} \in C(T) \right\} \tag{4}$$

is positive (Hestenes [**410**, Theorem 3.3]).

2.1. Exercises and examples.

EX. 1. *A nonclosed sum of closed subspaces.* Let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, and let

$$D = J_1 D(T) = \{\{\mathbf{x}, \mathbf{0}\} : \mathbf{x} \in D(T)\} .$$

Without loss of generality we assume that $D(T)$ is closed. Then D is closed. Also $G(T)$ is closed since T is bounded. But

$$G(T) + D$$

is nonclosed if $R(T)$ is nonclosed, since

$$\{\mathbf{x}, \mathbf{y}\} \in G(T) + D \iff \mathbf{y} \in R(T) \quad \text{(Halmos [\textbf{360}, p. 26])}.$$

EX. 2. *Unbounded linear functionals.* Let T be an unbounded linear functional on $\mathcal{H}$. Then $N(T)$ is dense in $\mathcal{H}$, and consequently $N(T)^\perp = \{\mathbf{0}\}$, $C(T) = \{\mathbf{0}\}$.

An example of such a functional on $L^2[0, \infty]$ is

$$T\mathbf{x} = \int_0^\infty t\mathbf{x}(t)dt .$$

To show that $N(T)$ is dense, let $\mathbf{x}_0 \in L^2[0, \infty]$ with $T\mathbf{x}_0 = \alpha$. Then a sequance $\{\mathbf{x}_n\} \subset N(T)$ converging to $\mathbf{x}_0$ is

$$\mathbf{x}_n(t) = \begin{cases} \mathbf{x}_0(t) & \text{if } t < 1 \text{ or } t > n+1 \\ \mathbf{x}_0(t) - \frac{\alpha}{nt} & \text{if } 1 \le t \le n+1 \end{cases}$$

Indeed,

$$\|\mathbf{x}_n - \mathbf{x}_0\|^2 = \int_1^{n+1} \frac{\alpha^2}{(nt)^2}\, dt = \frac{\alpha^2}{n(n+1)} \to 0 .$$

EX. 3. Let D be a dense subspace of $\mathcal{H}$, and let F be a closed subspace such that $F^\perp$ is finite dimensional. Then

$$\overline{D \cap F} = F \quad \text{(Erdelyi and Ben–Israel [\textbf{259}, Lemma 5.1])}.$$

EX. 4. *An operator with trivial carrier.* Let D be any proper dense subspace of $\mathcal{H}$ and choose $\mathbf{x} \notin D$. Let $F = [x]^\perp$, where $[\mathbf{x}]$ is the line generated by $\mathbf{x}$. Then $\overline{D \cap F} = F$, by Ex. 3. However, $D \not\subset F$, so we can choose a subspace $A \neq \{\mathbf{0}\}$ in D such that

$$D = A \oplus (D \cap F) .$$

Define $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ by

$$D(T) = D$$

and

$$T(\mathbf{y} + \mathbf{z}) = \mathbf{y} \quad \text{if } \mathbf{y} \in A,\ \mathbf{z} \in D \cap F .$$

Then

$$\begin{aligned} N(T) &= D \cap F , \\ \overline{N(T)} &= \overline{D \cap F} = F , \\ N(T)^\perp &= F^\perp = [\mathbf{x}] , \\ C(T) &= D(T) \cap N(T)^\perp = \{\mathbf{0}\} . \end{aligned}$$

EX. 5. Let L, M be subspaces of $\mathcal{H}$ and let $M \subset L$. Then

$$L = M \overset{\perp}{\oplus} (L \cap M^\perp) \tag{1}$$

if and only if

$$P_{\overline{M}}\mathbf{x} \in M \quad \text{for all } \mathbf{x} \in L .$$

In particular, a space is decomposable with respect to any closed subspace (Arghiriade [**26**]).

Ex. 6. Let L, M, N be subspaces of $\mathcal{H}$ such that

$$L = M \stackrel{\perp}{\oplus} N \,.$$

Then

$$M = L \cap N^{\perp} \,, \quad N = L \cap M^{\perp} \,.$$

Thus an orthogonal direct sum is decomposable with respect to each summand.

Ex. 7. *A bounded operator with nonclosed range.* Let ℓ^2 denote the Hilbert space of square summable sequences and let $T \in \mathcal{B}(\ell^2, \ell^2)$ be defined, for some $0 < k < 1$, by

$$T(\alpha_0,\, \alpha_1, \alpha_2,\, \dots\,,\, \alpha_n,\, \dots) = (\alpha_0,\, k\alpha_1,\, k^2\alpha_2,\, \dots\,,\, k^n\alpha_n,\, \dots) \,.$$

Consider the sequence

$$\mathbf{x}_n = \left(1,\, \frac{1}{2k},\, \frac{1}{3k^2},\, \dots\,,\, \frac{1}{nk^{n-1}},\, 0,\, 0,\, \dots\right) \,,$$

and the vector

$$\mathbf{y} = \lim_{n\to\infty} T\mathbf{x}_n = \left(1,\, \frac{1}{2},\, \frac{1}{3},\, \dots\,,\, \frac{1}{n},\, \dots\right) \,.$$

Then,

$$\mathbf{y} \in \overline{R(T)} \,,\; \mathbf{y} \notin R(T) \,.$$

Ex. 8. *Linear integral operators.* Let $L^2 = L^2[a,b]$, the Lebesgue square integrable functions on the finite interval $[a,b]$. Let $K(s,t)$ be an L^2–*kernel* on $a \le s, t, \le b$, meaning that the Lebesgue integral

$$\int_a^b \int_a^b |K(s,t)|^2 ds\, dt$$

exists and is finite; see, e.g., Smithies [**772**, Section 1.6].

Consider the two operators $T_1, T_2 \in \mathcal{B}(L^2, L^2)$ defined by

$$(T_1\mathbf{x})(s) = \int_a^b K(s,t)\mathbf{x}(t)dt \,, \quad a \le s \le b \,,$$

$$(T_2\mathbf{x})(s) = \mathbf{x}(s) - \int_a^b K(s,t)\mathbf{x}(t)dt \,, \quad a \le s \le b \,,$$

called *Fredholm integral operators of the first kind* and *the second kind*, respectively. Then

(a) $R(T_2)$ is closed.

(b) $R(T_1)$ is nonclosed unless it is finite dimensional.

More generally, if $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is completely continuous then $R(T)$ is nonclosed unless it is finite dimensional (Kammerer and Nashed [**458**, Proposition 2.5]).

Ex. 9. Let $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$. Then T is normally solvable if and only if T^* is. Also, T is normally solvable if and only if TT^* or T^*T is.

3. Generalized inverses of linear operators between Hilbert spaces

A natural definition of generalized inverses in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is the following one due to Tseng [**825**].

Definition 1. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then an operator $T^g \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ is a *Tseng generalized inverse* (abbreviated g.i.) of T if

$$R(T) \subset D(T^g) \tag{5}$$

$$R(T^g) \subset D(T) \tag{6}$$

$$T^gT\mathbf{x} = P_{\overline{R(T^g)}}\mathbf{x} \quad \text{for all } \mathbf{x} \in D(T) \tag{7}$$

$$TT^g\mathbf{y} = P_{\overline{R(T)}}\mathbf{y} \quad \text{for all } \mathbf{y} \in D(T^g) \,. \tag{8}$$

This definition is symmetric in T and T^g, thus T is a g.i. of T^g.

An operator $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ may have a unique g.i., or infinitely many g.i.'s or it may have none. We will show in Theorem 1 that T has a g.i. if and only if its domain is decomposable with respect to its null space,

$$\begin{aligned} D(T) &= N(T) \overset{\perp}{\oplus} (D(T) \cap N(T)^{\perp}) \\ &= N(T) \overset{\perp}{\oplus} C(T) \ . \end{aligned} \tag{9}$$

By Ex. 5, this condition is satisfied if $N(T)$ is closed. Thus it holds for all closed operators, and in particular for bounded operators. If T has g.i.'s, then it has a maximal g.i., some of whose properties are collected in Theorem 2. For bounded operators with closed range, the maximal g.i. coincides with the Moore–Penrose inverse, and will likewise be denoted by $T^{\dagger}$. See Theorem 3.

For operators T without g.i.'s, the maximal g.i. $T^{\dagger}$ can be "approximated" in several ways, with the objective of retaining as many of its useful properties as possible. One such approach, due to Erdélyi [**258**] is described in Definition 3 and Theorem 4.

Some properties of g.i.'s, when they exist, are given in the following three lemmas, due to Arghiriade [**26**], which are needed later.

LEMMA 1. If $T^g \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ is a g.i. of $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then $D(T)$ is decomposable with respect to $R(T^g)$.

PROOF. Follows from Ex. 5 since, for any $\mathbf{x} \in D(T)$

$$P_{\overline{R(T^g)}}\mathbf{x} = T^g T\mathbf{x} \ , \quad \text{by (7)} \ .$$

□

LEMMA 2. If $T^g \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ is a g.i. of $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then T is a one–to–one mapping of $R(T^g)$ onto $R(T)$.

PROOF. Let $\mathbf{y} \in R(T)$. Then

$$\mathbf{y} = P_{\overline{R(T)}}\mathbf{y} = TT^g\mathbf{y} \ , \quad \text{by (8)} \ ,$$

proving that $T(R(T^g)) = R(T)$.

Now we prove that T is one–to–one on $R(T^g)$. Let $\mathbf{x}_1, \mathbf{x}_2 \in R(T^g)$ satisfy

$$T\mathbf{x}_1 = T\mathbf{x}_2 \ .$$

Then

$$\mathbf{x}_1 = P_{\overline{R(T^g)}}\mathbf{x}_1 = T^g T\mathbf{x}_1 = T^g T\mathbf{x}_2 = P_{\overline{R(T^g)}}\mathbf{x}_2 = \mathbf{x}_2 \ .$$

□

LEMMA 3. If $T^g \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ is a g.i. of $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then:

$$N(T) = D(T) \cap R(T^g)^{\perp} \tag{10}$$

and

$$C(T) = R(T^g) \ . \tag{11}$$

PROOF. Let $\mathbf{x} \in D(T)$. Then, by Lemma 1,

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \ , \quad \mathbf{x}_1 \in R(T^g) \ , \quad \mathbf{x}_2 \in D(T) \cap R(T^g)^{\perp} \ , \quad \mathbf{x}_1 \perp \mathbf{x}_2 \ . \tag{12}$$

Now

$$\mathbf{x}_1 = P_{\overline{R(T^g)}}\mathbf{x} = T^g T(\mathbf{x}_1 + \mathbf{x}_2) = T^g T\mathbf{x}_1$$

and therefore

$$T^g T\mathbf{x}_2 = \mathbf{0}\ ,$$

which, by Lemma 2 with T and T^g interchanged, implies that

$$T\mathbf{x}_2 = \mathbf{0}\ , \tag{13}$$

hence

$$D(T) \cap R(T^g)^\perp \subset N(T)\ .$$

Conversely, let $\mathbf{x} \in N(T)$ be decomposed as in (12). Then

$$\begin{aligned}\mathbf{0} = T\mathbf{x} &= T(\mathbf{x}_1 + \mathbf{x}_2)\\ &= T\mathbf{x}_1\ , \quad \text{by (13)}\ ,\end{aligned}$$

which, by Lemma 2, implies that $\mathbf{x}_1 = \mathbf{0}$ and therefore

$$N(T) \subset D(T) \cap R(T^g)^\perp\ ,$$

completing the proof of (10).

Now

$$\begin{aligned}D(T) &= R(T^g) \stackrel{\perp}{\oplus} (D(T) \cap R(T^g)^\perp)\ , \quad \text{by Lemma 1}\ ,\\ &= R(T^g) \stackrel{\perp}{\oplus} N(T)\ ,\end{aligned}$$

which, by Ex. 6, implies that

$$R(T^g) = D(T) \cap N(T)^\perp\ ,$$

proving (11). □

The existence of g.i.'s is settled in the following theorem announced, without proof, by Tseng [**825**]. Our proof follows that of Arghiriade [**26**].

Theorem 1. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then T has a g.i. if and only if

$$D(T) = N(T) \stackrel{\perp}{\oplus} C(T)\ , \tag{9}$$

in which case, for any subspace $L \subset R(T)^\perp$, there is a g.i. T_L^g of T, with

$$D(T_L^g) = R(T) \stackrel{\perp}{\oplus} L \tag{14}$$

and

$$N(T_L^g) = L\ . \tag{15}$$

Proof. If T has a g.i., then (9) follows from Lemmas 1 and 3.

Conversely, suppose that (9) holds. Then

$$R(T) = T(D(T)) = T(C(T)) = R(T_0)\ , \tag{16}$$

where $T_0 = T_{[C(T)]}$ is the restriction of T to $C(T)$. The inverse T_0^{-1} exists, by Section 2(F), and satisfies

$$R(T_0^{-1}) = C(T)$$

and, by (16,)

$$D(T_0^{-1}) = R(T)\ .$$

For any subspace $L \subset R(T)^{\perp}$, consider the extension T_L^g of T_0^{-1} with domain

$$D(T_L^g) = R(T) \overset{\perp}{\oplus} L \tag{14}$$

and null space

$$N(T_L^g) = L \,. \tag{15}$$

From its definition, it follows that T_L^g satisfies

$$D(T_L^g) \supset R(T)$$

and

$$R(T_L^g) = R(T_0^{-1}) = C(T) \subset D(T) \,. \tag{17}$$

For any $\mathbf{x} \in D(T)$

$$\begin{aligned} T_L^g T\mathbf{x} &= T_L^g T P_{\overline{C(T)}} \mathbf{x} \,, \quad \text{by (9)} \\ &= T_0^{-1} T_0 P_{\overline{C(T)}} \mathbf{x} \,, \quad \text{by Ex. 5} \\ &= P_{\overline{R(T_L^g)}} \mathbf{x} \,, \quad \text{by (17)} \,. \end{aligned}$$

Finally, any $\mathbf{y} \in D(T_L^g)$ can be written, by (14), as

$$\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2 \,, \quad \mathbf{y}_1 \in R(T) \,, \quad \mathbf{y}_2 \in L \,, \quad \mathbf{y}_1 \perp \mathbf{y}_2 \,,$$

and therefore

$$\begin{aligned} TT_L^g \mathbf{y} &= TT_L^g \mathbf{y}_1 \,, \quad \text{by (15)} \\ &= T_0 T_0^{-1} \mathbf{y}_1 \\ &= \mathbf{y}_1 \\ &= P_{\overline{R(T)}} \mathbf{y} \,. \end{aligned}$$

Thus T_L^g is a g.i. of T. □

The g.i. T_L^g is uniquely determined by its domain (14) and null space (15); see Ex. 10.

The maximal choice of the subspace L in (14) and (15) is $L = R(T)^{\perp}$. For this choice we have the following

DEFINITION 2. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ satisfy (9). Then the *maximal g.i.* of T, denoted by $T^{\dagger}$, is the g.i. of T with domain

$$D(T^{\dagger}) = R(T) \overset{\perp}{\oplus} R(T)^{\perp} \tag{18}$$

and null space

$$N(T^{\dagger}) = R(T)^{\perp} \,. \tag{19}$$

By Ex. 10, the g.i. $T^{\dagger}$ so defined is unique. It is maximal in the sense that any other g.i. of T is a restriction of $T^{\dagger}$.

Moreover, $T^{\dagger}$ is dense, by (18), and has a closed null space, by (19). Choosing L as a dense subspace of $R(T)^{\perp}$ shows that an operator T may have infinitely many dense g.i.'s T_L^g. Also, T may have infinitely many g.i.'s T_L^g with closed null space, each obtained by choosing L as a closed subspace of $R(T)^{\perp}$. However, $T^{\dagger}$ is the unique dense g.i. with closed null space; see Ex. 11.

For closed operators, the maximal g.i. can be alternatively defined, by means of the following construction due to Hestenes [**410**], see also Landesman [**492**].

Let $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$ be dense. Since $N(T)$ is closed, it follows, from Ex. 5, that

$$D(T) = N(T) \overset{\perp}{\oplus} C(T) , \tag{9}$$

and therefore

$$G(T) = N \overset{\perp}{\oplus} C , \tag{20}$$

where, using the notation of Section 2(B), (C), and (F),

$$N = J_1 N(T) = G(T) \cap \mathcal{H}_{1,0} , \tag{21}$$

$$C = \{\{\mathbf{x}, T\mathbf{x}\} : \mathbf{x} \in C(T)\} . \tag{22}$$

Similarly, since T^* is closed, it follows from Section 2(G), that

$$G(T)^\perp = N^* \overset{\perp}{\oplus} C^* \tag{23}$$

with

$$N^* = J_2 N(T^*) = G(T)^\perp \cap \mathcal{H}_{0,2} , \tag{24}$$

$$C^* = \{\{-T^*\mathbf{y}, \mathbf{y}\} : \mathbf{y} \in C(T^*)\} . \tag{25}$$

Now

$$\begin{aligned} \mathcal{H}_{1,2} &= G(T) \overset{\perp}{\oplus} G(T)^\perp , \quad \text{since } T \text{ is closed} \\ &= (N \overset{\perp}{\oplus} C) \overset{\perp}{\oplus} (N^* \overset{\perp}{\oplus} C^*) , \quad \text{by (20) and (23)} \\ &= (C \overset{\perp}{\oplus} N^*) \overset{\perp}{\oplus} (C^* \overset{\perp}{\oplus} N) \\ &= G^\dagger \overset{\perp}{\oplus} G^{\dagger *} , \end{aligned} \tag{26}$$

where

$$G^\dagger = C \overset{\perp}{\oplus} N^* , \tag{27}$$

$$G^{\dagger *} = C^* \overset{\perp}{\oplus} N . \tag{28}$$

Since

$$G^\dagger \cap \mathcal{H}_{1,0} = \{\mathbf{0}, \mathbf{0}\} , \quad \text{by Section 2(F)} ,$$

it follows that $G^\dagger$ is the inverse graph of an operator $T^\dagger \in \mathcal{C}(\mathcal{H}_2, \mathcal{H}_1)$, with domain

$$\begin{aligned} J_2^{-1} P_{\mathcal{H}_{0,2}} G^\dagger &= T(C(T)) \overset{\perp}{\oplus} N(T^*) \\ &= R(T) \overset{\perp}{\oplus} R(T)^\perp , \quad \text{by (16) and (3)}, \end{aligned}$$

and null space

$$J_2^{-1} M^* = N(T^*) = R(T)^\perp$$

and such that

$$T^\dagger T \mathbf{x} = P_{\overline{C(T)}} \mathbf{x} , \quad \text{for any } \mathbf{x} \in N(T) \overset{\perp}{\oplus} C(T) ,$$

and

$$T T^\dagger \mathbf{y} = P_{\overline{R(T)}} \mathbf{y} , \quad \text{for any } \mathbf{y} \in R(T) \overset{\perp}{\oplus} R(T)^\perp .$$

Thus $T^\dagger$ is the maximal g.i. of Definition 2.

Similarly, $G^{\dagger *}$ is the graph of the operator $-T^{*\dagger} \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$, which is the maximal g.i. of $-T^*$.

This elegant construction makes obvious the properties of the maximal g.i., collected in the following:

Theorem 2. (Hestenes [**410**]). Let $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$ be dense. Then

(a) $T^\dagger \in \mathcal{C}(\mathcal{H}_2, \mathcal{H}_1)$,

(b) $D(T^\dagger) = R(T) \overset{\perp}{\oplus} N(T^*)$, $\quad N(T^\dagger) = N(T^*)$,

(c) $R(T^\dagger) = C(T)$,

(d) $T^\dagger T\mathbf{x} = P_{\overline{R(T^\dagger)}}\mathbf{x}$ for any $\mathbf{x} \in D(T)$,

(e) $TT^\dagger \mathbf{y} = P_{\overline{R(T)}}\mathbf{y}$ for any $\mathbf{y} \in D(T^\dagger)$,

(f) $T^{\dagger\dagger} = T$,

(g) $T^{*\dagger} = T^{\dagger *}$,

(h) $N(T^{*\dagger}) = N(T)$,

(i) T^*T and $T^\dagger T^{*\dagger}$ are non–negative and

$$(T^*T)^\dagger = T^\dagger T^{*\dagger} \ , \quad N(T^*T) = N(T) \ ,$$

(j) TT^* and $T^{*\dagger}T^\dagger$ are non–negative and

$$(TT^*)^\dagger = T^{*\dagger}T^\dagger \ , \quad N(TT^*) = N(T^*) \ .$$

□

For bounded operators with closed range, various characterizations of the maximal g.i. are collected in the following:

Theorem 3. (Petryshyn [**643**]). If $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ and $R(T)$ is closed, then $T^\dagger$ is characterized as the unique solution X of the following equivalent systems:

(a) $TXT = T \ , \ XTX = X \ , \ (TX)^* = TX \ , \ (XT)^* = XT$,

(b) $TX = P_{R(T)} \ , \ N(X^*) = N(T)$,

(c) $TX = P_{R(T)} \ , \ XT = P_{R(T^*)} \ , \ XTX = X$,

(d) $XTT^* = T^* \ , \ XX^*T^* = X$,

(e) $XT\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in R(T^*)$,
$X\mathbf{y} = \mathbf{0}$ for all $\mathbf{y} \in N(T^*)$,

(f) $XT = P_{R(T^*)} \ , \ N(X) = N(T^*)$,

(g) $TX = P_{R(T)} \ , \ XT = P_{R(X)}$. □

The notation $T^\dagger$ is justified by Theorem 3(a), which lists the four *Penrose equations* (1.1)–(1.4).

If $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ does not satisfy (9), then it has no g.i., by Theorem 1. In this case one can still approximate $T^\dagger$ by an operator that has some properties of $T^\dagger$, and reduces to it if $T^\dagger$ exists. Such an approach, due to Erdélyi [**258**], is described in the following

Definition 3. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and let T_r be the restriction of T defined by

$$D(T_r) = N(T) \overset{\perp}{\oplus} C(T) \ , \ N(T_r) = N(T) \ . \tag{29}$$

The (*Erdelyi*) *g.i.* of T is defined as $T_r^\dagger$, which exists since T_r satisfies (9).

The inverse graph of $T_r^\dagger$ is

$$G^{-1}(T_r) = \{\{\mathbf{x}, T\mathbf{x} + \mathbf{z}\} : \mathbf{x} \in C(T) \ , \ \mathbf{z} \in (T(C(T)))^\perp\} \ , \tag{30}$$

from which the following properties of $T_r^\dagger$ can be easily deduced.

Theorem 4. (Erdélyi [**258**]). Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and let its restriction T_r be defined by (29). Then

(a) $T_r^\dagger = T^\dagger$ if $T^\dagger$ exists,

(b) $D(T_r^\dagger) = T(C(T)) \overset{\perp}{\oplus} T(C(T))^\perp$,
and in general, $R(T) \not\subset D(T_r^\dagger)$,

(c) $R(T_r^\dagger) = C(T)$, $\overline{R(T_r^\dagger)} = N(T)^\perp$,
(d) $T_r^\dagger T\mathbf{x} = P_{\overline{R(T_r^\dagger)}}\mathbf{x}$ for all $\mathbf{x} \in D(T_r)$,
(e) $TT_r^\dagger\mathbf{y} = P_{\overline{R(T)}}\mathbf{y}$ for all $\mathbf{y} \in D(T_r^\dagger)$,
(f) $D((T_r^\dagger)_r^\dagger) = \overline{N(T)} \overset{\perp}{\oplus} C(T)$,
(g) $R((T_r^\dagger)_r^\dagger) = T(C(T))$,
(h) $N((T_r^\dagger)_r^\dagger) = \overline{N(T)}$,
(i) $T \subset (T_r^\dagger)_r^\dagger$ if (9) holds ,
(j) $T = (T_r^\dagger)_r^\dagger$ if and only if $N(T)$ is closed ,
(k) $T_r^{\dagger *} \subset (T^*)_r^\dagger$ if T is dense and closable. □

See also Ex. 15.

Exercises and examples.

Ex. 10. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ have g.i.'s and let L be a subspace of $R(T)^\perp$. Then the conditions

$$D(T_L^g) = R(T) \overset{\perp}{\oplus} L \tag{14}$$
$$N(T_L^g) = L \tag{15}$$

determine a unique g.i., which is thus equal to T_L^g as constructed in the proof of Theorem 1.

Proof. Let T^g be a g.i. of T satisfying (14) and (15), and let $\mathbf{y} \in D(T^g)$ be written as

$$\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2 , \quad \mathbf{y}_1 \in R(T) , \quad \mathbf{y}_2 \in L .$$

Then

$$\begin{aligned} T^g\mathbf{y} &= T^g\mathbf{y}_1 , \quad \text{by (15)} \\ &= T^gT\mathbf{x}_1 , \quad \text{for some } \mathbf{x}_1 \in D(T) \\ &= P_{\overline{R(T^g)}}\mathbf{x}_1 , \quad \text{by (7)} \\ &= P_{\overline{C(T)}}\mathbf{x}_1 , \quad \text{by (11)} . \end{aligned}$$

We claim that this determines T^g uniquely. For, suppose there is an $\mathbf{x}_2 \in D(T)$ with $\mathbf{y}_1 = T\mathbf{x}_2$. Then, as above,

$$T^g\mathbf{y} = P_{\overline{C(T)}}\mathbf{x}_2$$

and therefore

$$\begin{aligned} P_{\overline{C(T)}}\mathbf{x}_1 - P_{\overline{C(T)}}\mathbf{x}_2 &= P_{\overline{C(T)}}(\mathbf{x}_1 - \mathbf{x}_2) \\ &= \mathbf{0} \quad \text{since } \mathbf{x}_1 - \mathbf{x}_2 \in N(T) . \end{aligned}$$

□

Ex. 11. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ have g.i.'s. Then $T^\dagger$ is the unique dense g.i. with closed null space.

Proof. Let T^g be any dense g.i. with closed null space. Then

$$\begin{aligned} D(T^g) &= N(T^g) \overset{\perp}{\oplus} C(T^g) , \quad \text{by Theorem 1} \\ &= N(T^g \overset{\perp}{\oplus} R(T) , \quad \text{by (11)} , \end{aligned}$$

which, together with the assumptions $\overline{D(T^g)} = \mathcal{H}_2$ and $N(T^g) = \overline{N(T^g)}$, implies that

$$N(T^g) = R(T)^\perp .$$

Thus, T^g has the same domain and null space as $T^\dagger$, and therefore $T^g = T^\dagger$, by Ex. 10. □

Ex. 12. Let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ have a closed range $R(T)$ and let $T_1 \in \mathcal{B}(\mathcal{H}_1, R(T))$ be defined by

$$T_1\mathbf{x} = T\mathbf{x} \quad \text{for all } \mathbf{x} \in \mathcal{H}_1 .$$

Then

(a) T_1^* is the restriction of T^* to $R(T)$.

(b) The operator $T_1T_1^* \in \mathcal{B}(R(T), R(T))$ is invertible.

(c) $T^\dagger = P_{R(T^*)}T_1^*(T_1T_1^*)^{-1}P_{R(T)}$ (Kurepa [**485**]).

Ex. 13. Let $T \in \mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$. Then $R(T)$ is closed if and only if $T^\dagger$ is bounded (Landesman [**492**]).

PROOF. Follows from Section 2(M). □

Ex. 14. Let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ have closed range. Then

$$T^\dagger = (T^*T)^\dagger T^* = T^*(TT^*)^\dagger \quad \text{(Desoer and Whalen [\textbf{222}]).}$$

Ex. 15. For arbitrary $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ consider its extension $\widetilde{T}$ with

$$D(\widetilde{T}) = D(T) + \overline{N(T)} , \quad N(\widetilde{T}) = \overline{N(T)} , \quad \widetilde{T} = T \text{ on } D(T) , \tag{31}$$

which coincides with T if $N(T)$ is closed. Since $D(\widetilde{T})$ is decomposable with respect to $N(\widetilde{T})$, it might seem that $\widetilde{T}$ can be used to obtain $\widetilde{T}^\dagger$, a substitute for (possibly nonexisting) $T^\dagger$.

Show that $\widetilde{T}$ is not well defined by (31) if

$$D(T) \cap \overline{N(T)} \neq N(T) \quad \text{and} \quad N(\widetilde{T}) \neq D(\widetilde{T}) , \tag{32}$$

which is the only case of interest since otherwise $D(T)$ is decomposable with respect to $N(T)$ or $\widetilde{T}$ is identically O in its domain.

PROOF. By (32) there exist $\mathbf{x}_0$ and $\mathbf{y}$ such that

$$\mathbf{x}_0 \in D(T) \cap \overline{N(T)} , \quad \mathbf{x}_0 \notin N(T)$$

and

$$\mathbf{y} \in D(T) , \quad \mathbf{y} \notin \overline{N(T)} .$$

Then

$$\widetilde{T}(\mathbf{x}_0 + \mathbf{y}) = \widetilde{T}\mathbf{y} , \quad \text{since } \mathbf{x}_0 \in N(\widetilde{T})$$

and on the other hand

$$\begin{aligned} \widetilde{T}(\mathbf{x}_0 + \mathbf{y}) &= T(\mathbf{x}_0 + \mathbf{y}) , \quad \text{since } \mathbf{x}_0, \mathbf{y} \in D(T) \\ &\neq T\mathbf{y} , \quad \text{since } \mathbf{x}_0 \notin N(T) . \end{aligned}$$

□

Ex. 16. Let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ have closed range. Then

$$\|T^\dagger\| = \frac{1}{\gamma(T)} ,$$

where $\gamma(T)$ is defined in (4) (Petryshyn [**643**, Lemma 2]).

Ex. 17. Let $F \in \mathcal{B}(\mathcal{H}_3, \mathcal{H}_2)$ and $G \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_3)$ with $R(G) = \mathcal{H}_3 = R(F^*)$, and define $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ by $A = FG$, Then

$$\begin{aligned} A^\dagger &= G^*(GG^*)^{-1}(F^*F)^{-1}F^* \\ &= G^\dagger F^\dagger \quad \text{(Holmes [\textbf{422}, p. 223]).} \end{aligned}$$

Compare with Theorem 1.5 and Ex. 1.17.

Ex. 18. *Generalized inverses of linear integral operators.* In this exercise and in Exs. 19–25 below we consider the *Fredholm integral equation of the second kind*

$$x(s) - \lambda \int_a^b K(s,t)\, x(t)\, dt = y(s) , \quad a \le s \le b , \tag{33}$$

written for short as

$$(I - \lambda K)\, \mathbf{x} = \mathbf{y} ,$$

where all functions are complex, $[a, b]$ is a bounded interval, λ is a complex scalar and $K(s, t)$ is a L^2–*Kernel* on $[a, b] \times [a, b]$; see Ex. 8. Writing L^2 for $L^2[a, b]$, we need the following facts from the Fredholm thoery of integral equations; see, e.g., Smithies [**772**]. For any λ, K as above

(a) $(I - \lambda K) \in \mathcal{B}(L^2, l^2)$,

(b) $(I - \lambda K)^* = I - \overline{\lambda} K^*$, where $K^*(s,t) = \overline{K(t,s)}$.

(c) The null spaces $N(I - \lambda K)$ and $N(I - \overline{\lambda} K^*)$ have equal finite dimensions,

$$\dim N(I - \lambda K) = \dim N(I - \overline{\lambda} K^*) = n(\lambda) \text{ , say .} \tag{34}$$

(d) A scalar λ is called a *regular value* of K if $n(\lambda) = 0$, in which case the operator $I - \lambda K$ has an inverse $(I - \lambda K)^{-1} \in \mathcal{B}(L^2, L^2)$ written as

$$(I - \lambda K)^{-1} = I + \lambda R , \tag{35}$$

where $R = R(s, t; \lambda)$ is an L^2–kernel called the resolvent of K.

(e) A scalar λ is called an *eigenvalue* of K if $n(\lambda) > 0$, in which case any nonzero $\mathbf{x} \in N(I - \lambda K)$ is called an *eigenfunction* of K corresponding to λ. For any λ and, in particular, for any eigenvalue λ, both range spaces $R(I - \lambda K)$ and $R(I - \overline{\lambda} K^*)$ are closed and, by (3),

$$R(I - \lambda K) = N(I - \overline{\lambda} K^*)^{\perp} , \quad R(I - \overline{\lambda} K^*) = N(I - \lambda K)^{\perp} . \tag{36}$$

Thus, if λ is a regular value of K then (33) has, for any $\mathbf{y} \in L^2$, a unique solution given by

$$\mathbf{x} = (I + \lambda R)\, \mathbf{y} ,$$

that is

$$x(s) = y(s) + \lambda \int_a^b R(s, t, \lambda)\, y(t)\, dt , \ a \le s \le b . \tag{37}$$

If λ is an eigenvalue of K then (33) is consistent if and only if $\mathbf{y}$ is orthogonal to every $\mathbf{u} \in N(I - \overline{\lambda} K^*)$, in which case the general solution of (33) is

$$\mathbf{x} = \mathbf{x}_0 + \sum_{i=1}^{n(\lambda)} c_i \mathbf{x}_i , \quad c_i \text{ arbitrary scalars ,} \tag{38}$$

where $\mathbf{x}_0$ is a particular solution of (33) and $\{\mathbf{x}_1, \dots , \mathbf{x}_{n(\lambda)}\}$ is a basis of $N(I - \lambda K)$.

Ex. 19. *Pseudo resolvents.* Let λ be an eigenvalue of K. Following Hurwitz [**430**], an L^2–kernel $R = R(s, t, \lambda)$ is called a *pseudo resolvent* of K if for any $\mathbf{y} \in R(I - \lambda K)$, the function

$$x(s) = y(s) + \lambda \int_a^b R(s, t, \lambda)\, y(t)\, dt \tag{37}$$

is a solution of (33).

A pseudo resolvent was constructed by Hurwitz as follows.

Let λ_0 be an eigenvalue of K, and let $\{\mathbf{x}_1, \dots , \mathbf{x}_n\}$ and $\{\mathbf{u}_1, \dots , \mathbf{u}_n\}$ be orthonormal bases of $N(I - \lambda_0 K)$ and $N(I - \overline{\lambda_0} K^*)$ respectively. Then λ_0 is a regular value of the kernel

$$K_0(s,t) = K(s,t) - \frac{1}{\lambda_0} \sum_{i=1}^{n} u_i(s)\, \overline{x_i(t)} , \tag{39}$$

written for short as

$$K_0 = K - \frac{1}{\lambda_0} \sum_{i=1}^{n} \mathbf{u}_i \mathbf{x}_i^*$$

and the resolvent R_0 of K_0 is a pseudo resolvent of K, satisfying

$$\begin{aligned}(I+\lambda_0 R_0)(I-\lambda_0 K)\,\mathbf{x} &= \mathbf{x}\ , \quad \text{for all } \mathbf{x}\in R(I-\overline{\lambda_0}K^*)\\ (I-\lambda_0 K)(I+\lambda_0 R_0)\,\mathbf{y} &= \mathbf{y}\ , \quad \text{for all } \mathbf{y}\in R(I-\lambda_0 K)\\ (I+\lambda_0 R_0)\,\mathbf{u}_i &= \mathbf{x}_i\ , \quad i=1,\dots,n\ .\end{aligned} \tag{40}$$

PROOF. Follows from the matrix case, Ex. 2.53. □

EX. 20. A comparison with Theorem 2.2 shows that $I+\lambda R$ is a $\{1\}$–inverse of $I-\lambda K$, if R is a pseudo resolvent of K. As with $\{1\}$–inverses, the pseudo resolvent is nonunique. Indeed, for $R_0, \mathbf{u}_i, \mathbf{x}_i$ as above, the kernel

$$R_0 + \sum_{i,j,=1}^{n} c_{ij}\mathbf{x}_i\mathbf{u}_j^* \tag{41}$$

is a pseudo resolvent of K for any choice of scalars c_{ij}.

The pseudo resolvent constructed by Fredholm [**284**], who called the resulting operator $I+\lambda R$ a *pseudo inverse* of $I-\lambda K$, is the first explicit application, known to us, of a generalized inverse.

The class of all pseudo resolvents of a given kernel K is characterized as follows.

Let K be an L^2–kernel, let λ_0 be an eigenvalue of K and let $\{\mathbf{x}_1,\dots,\mathbf{x}_n\}$ and $\{\mathbf{u}_1,\dots,\mathbf{u}_n\}$ be orthonormal bases of $N(I-\lambda_0 K)$ and $N(I-\overline{\lambda_0}K^*)$ respectively. An L^2–kernel R is a pseudo resolvent of K if and only if

$$R = K + \lambda_0 KR - \frac{1}{\lambda_0}\sum_{i=1}^{n}\boldsymbol{\beta}_i\mathbf{u}_i^*\ , \tag{42a}$$

$$R = K + \lambda_0 RK - \frac{1}{\lambda_0}\sum_{i=1}^{n}\mathbf{x}_i\boldsymbol{\alpha}_i^*\ , \tag{42b}$$

where $\boldsymbol{\alpha}_i,\ \boldsymbol{\beta}_i\in L^2$ satisfy

$$\langle\boldsymbol{\alpha}_i,\mathbf{x}_j\rangle=\delta_{ij}\ ,\ \langle\boldsymbol{\beta}_i,\mathbf{u}_j\rangle=\delta_{ij}\ ,\quad i,j=1,\dots,n\ . \tag{43}$$

Here KR stands for the kernel $KR(s,t)=\int_a^b K(s,u)R(u,t)\,du$, etc.

If λ is a regular value of K then (42) reduces to

$$R=K+\lambda KR\ ,\ R=K+\lambda RK\ , \tag{44}$$

which uniquely determines the resolvent $R(s,t,\lambda)$ (Hurwitz [**430**]).

EX. 21. Let K, λ_0, $\mathbf{x}_i$, $\mathbf{u}_i$, and R_0 be as above. Then the maximal g.i. of $I-\lambda_0 K$ is

$$(I-\lambda_0 K)^\dagger = I+\lambda_0 R_0 - \sum_{i=1}^{n}\mathbf{x}_i\mathbf{u}_i^*\ , \tag{45}$$

corresponding to the pseudo resolvent

$$R = R_0 - \frac{1}{\lambda_0}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{u}_i^*\ . \tag{46}$$

EX. 22. Let $K(s,t)=u(s)\,\overline{v(t)}$, where

$$\int_a^b u(s)\,\overline{v(s)}\,ds = 0\ .$$

Then every scalar λ is a regular value of K.

EX. 23. Consider the equation

$$x(s)-\lambda\int_{-1}^{1}(1+3st)\,x(t)\,dt = y(s) \tag{47}$$

with $K(s,t)=1+3st$. The resolvent is

$$R(s,t;\lambda)=\frac{1+3st}{1-2\lambda}\ .$$

K has a single eigenvalue $\lambda = \frac{1}{2}$ and an orthonormal basis of $N(I - \frac{1}{2}K)$ is

$$\left\{ x_1(s) = \frac{1}{\sqrt{2}},\ x_2(s) = \frac{\sqrt{3}}{\sqrt{2}}\, s \right\}$$

which, by symmetry, is also an orthonormal basis of $N(I - \frac{1}{2}K^*)$. From (39) we get

$$\begin{aligned} K_0(s,t) &= K(s,t) - \frac{1}{\lambda_0} \sum u_i(s)\, \overline{x_i(t)} \\ &= (1 + 3st) - 2\left(\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} + \frac{\sqrt{3}}{\sqrt{2}} s \frac{\sqrt{3}}{\sqrt{2}} t \right) \\ &= 0\,, \end{aligned}$$

and the resolvent of $K_0(s,t)$ is therefore

$$R_0(s,t;\lambda) = 0\,.$$

If $\lambda \neq \frac{1}{2}$, then for each $y \in L^2[-1,1]$ equation (47) has a unique solution

$$x(s) = y(s) + \lambda \int_{-1}^{1} \frac{1 + 3st}{1 - 2\lambda} y(t)\, dt\,.$$

If $\lambda = \frac{1}{2}$, then (47) is consistent if and only if

$$\int_{-1}^{1} y(t)\, dt = 0\,, \qquad \int_{-1}^{1} t\, y(t)\, dt = 0\,,$$

in which case the general solution is

$$x(s) = y(s) + c_1 + c_2 s\,, \quad c_1, c_2 \text{ arbitrary}\,.$$

EX. 24. Let

$$K(s,t) = 1 + s + 3st\,, \quad -1 \leq s,t \leq 1\,.$$

Then $\lambda = \frac{1}{2}$ is the only eigenvalue and

$$\dim N(I - \tfrac{1}{2}K) = 1\,.$$

An orthonormal basis of $N(I - \frac{1}{2}K)$ is the single vector

$$x_1(s) = \frac{\sqrt{3}}{\sqrt{2}}\, s\,, \quad -1 \leq s \leq 1\,.$$

An orthonormal basis of $N(I - \frac{1}{2}K^*)$ is

$$u_1(s) = \frac{1}{\sqrt{2}}\,, \quad -1 \leq s \leq 1\,.$$

The Hurwitz kernel (39) is

$$\begin{aligned} K_0(s,t) &= (1 + s + 3st) - 2\left(\frac{1}{\sqrt{2}} \frac{\sqrt{3}}{\sqrt{2}} t \right) \\ &= 1 + s - \sqrt{3}t + 3st\,, \quad -1 \leq s,t \leq 1\,. \end{aligned}$$

Compute the resolvent R_0 of K_0, which is a pseudo resolvent of K. (*Hint*: Use the following exercise).

EX. 25. *Degenerate kernels.* A kernel $K(s,t)$ is called *degenerate* if it is a finite sum of products of L^2 functions, as follows:

$$K(s,t) = \sum_{i=1}^{m} f_i(s)\, \overline{g_i(t)}\,. \tag{48}$$

Degenerate kernels are convenient because they reduce the integral equation (33) to a finite system of linear equations. Also, any L^2–kernel can be approximated, arbitrarily close, by a degenerate kernel; see, e.g., Smithies [**772**, p. 40], and Halmos [**360**, Problem 137].

Let $K(s,t)$ be given by (48). Then

(a) The scalar λ is an eigenvalue of (48) if and only if $1/\lambda$ is an eigenvalue of the $m \times m$ matrix

$$B = [b_{ij}] \ , \quad \text{where } b_{ij} = \int_a^b f_j(s)\, \overline{g_i(s)}\, ds \ .$$

(b) Any eigenfunction of K $[K^*]$ corresponding to an eigenvalue λ $[\overline{\lambda}]$ is a linear combination of the m functions $f_1, \dots, f_m$ $[g_1, \dots, g_m]$.

(c) If λ is a regular value of (48), then the resolvent at λ is

$$R(s,t,;\lambda) = \frac{\det \begin{bmatrix} 0 & \vdots & f_1(s) & \cdots & f_m(s) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\overline{g_1(t)} & \vdots & & & \\ \vdots & \vdots & & I - \lambda B & \\ -\overline{g_m(t)} & \vdots & & & \end{bmatrix}}{\det(I - \lambda B)} \ .$$

See also Kantorovich and Krylov [**461**, Chapter II].

4. Generalized inverses of linear differential operators

This section deals with generalized inverses of closed dense operators $L \in \mathcal{C}(\mathcal{S}_1, \mathcal{S}_2)$ with $\overline{D(L)} = \mathcal{S}_1$, where:

(i) $\mathcal{S}_1, \mathcal{S}_2$ are spaces of (scalar or vector) functions which are either the Hilbert space $L^2[a,b]$ or the *space of continuous functions* $C[a,b]$, where $[a,b]$ is a given finite real interval. Since $C[a,b]$ is a dense subspace of $L^2[a,b]$, a closed dense linear operator mapping $C[a,b]$ into $\mathcal{S}_2$ may be considered as a dense operator in $\mathcal{C}(L^2[a,b], \mathcal{S}_2)$.

(ii) L is defined for all $\mathbf{x}$ in its domain $D(L)$ by

$$L\mathbf{x} = \ell\, \mathbf{x} \ , \tag{49}$$

where ℓ is a *differential expression*, for example, in the vector case,

$$\ell\, \mathbf{x} = A_1(t) \frac{d}{dt} \mathbf{x} + A_0(t) \mathbf{x} \ , \tag{50}$$

where $A_0(t), A_1(t)$ are $n \times n$ matrix coefficients, with suitable regularity conditions; see, e.g., Ex. 31 below.

(iii) The domain of L consists of those functions in $\mathcal{S}_1$ for which ℓ makes sense and $\ell\, \mathbf{x} \in \mathcal{S}_2$, and which satisfy certain conditions, such as *initial* or *boundary conditions*.

If a differential operator L is invertible and there is a kernel (function, or matrix in the vector case)

$$G(s,t) \ , \quad a \le s, t \le b \ ,$$

such that for all $\mathbf{y} \in R(L)$

$$(L^{-1}\mathbf{y})(s) = \int_a^b G(s,t)\mathbf{y}(t) dt \ , \quad a \le s \le b \ ,$$

then $G(s,t)$ is called the *Green's function* (or *matrix*) of L. In this case, for any $\mathbf{y} \in R(L)$, the unique solution of

$$L\mathbf{x} = \mathbf{y} \tag{51}$$

is given by

$$\mathbf{x}(s) = \int_a^b G(s,t)\mathbf{y}(t) dt \ , \quad a \le s \le b \ . \tag{52}$$

If L is not invertible, but there is a kernel $G(s,t)$ such that, for any $\mathbf{y} \in R(L)$, a particular solution of (51) is given by (52), then $G(s,t)$ is called a *generalized Green's function* (or *matrix*) of L. A generalized Green's function of L is therefore a kernel of an integral operator which is a generalized inverse of L.

Generalized Green's functions were introduced by Hilbert [**412**] in 1904, and consequently studied by Myller, Westfall and Bounitzky [**123**], Elliott ([**244**], [**245**]), and Reid [**688**]; see, e.g., the historical survey in [**691**].

Exercises and examples.

EX. 26. *Derivatives.* Let

$\mathcal{S}$ = the real space $L^2[0,\pi]$ of real valued functions,

$\mathcal{S}^1$ = the absolutely continuous functions $\mathbf{x}(t)$, $0 \le t \le \pi$, whose derivatives $\mathbf{x}'$ are in $\mathcal{S}$,

$\mathcal{S}^2 = \{\mathbf{x} \in \mathcal{S}^1 : \mathbf{x}' \in \mathcal{S}^1\}$,

and let L be the differential operator d/dt with

$$D(L) = \{\mathbf{x} \in \mathcal{S}^1 : \mathbf{x}(0) = \mathbf{x}(\pi) = 0\} .$$

Then

(a) $L \in \mathcal{C}(\mathcal{S},\mathcal{S})$, $\overline{D(L)} = \mathcal{S}$, $C(L) = D(L)$,

$$R(L) = \left\{\mathbf{y} \in \mathcal{S} : \int_0^\pi \mathbf{y}(t)\,dt = 0\right\} = \overline{R(L)} .$$

(b) The adjoint L^* is the operator $-d/dt$ with

$$D(L^*) = \mathcal{S}^1 , \quad C(L^*) = \mathcal{S}^1 \cap R(L) , \quad R(L^*) = \mathcal{S} .$$

(c) $L^*L = -\dfrac{d^2}{dt^2}$ with $D(L^*L) = \{\mathbf{x} \in \mathcal{S}^2 : \mathbf{x}(0) = \mathbf{x}(\pi) = 0\}$ and $R(L^*L) = \mathcal{S}$.

(d) $LL^* = -\dfrac{d^2}{dt^2}$ with $D(LL^*) = \{\mathbf{x} \in \mathcal{S}^2 : \mathbf{x}'(0) = \mathbf{x}'(\pi) = 0\}$ and $R(LL^*) = R(L)$.

(e) $L^\dagger$ is defined on $D(L^\dagger) = \mathcal{S}$ by

$$(L^\dagger\mathbf{y})(t) = \int_0^t \mathbf{y}(s)\,ds - \frac{t}{\pi}\int_0^\pi \mathbf{y}(s)\,ds , \quad 0 \le t \le \pi$$

(Hestenes [**410**, Example 1]).

EX. 27. For L of Ex. 26, determine which of the following equations hold and interpret your results:

(a) $L^{\dagger *} = L^{*\dagger}$,

(b) $L^\dagger = (L^*L)^\dagger L^* = L^*(LL^*)^\dagger$,

(c) $L^{\dagger\dagger} = L$.

EX. 28. *Gradients.* Let

$\mathcal{S}$ = the real space $L^2([0,\pi]\times[0,\pi])$ of real valued functions $\mathbf{x}(t_1,t_2)$, $0 \le t_1, t_2 \le \pi$.

$\mathcal{S}^1$ = the subclass of $\mathcal{S}$ with the properties

(i) $\mathbf{x}(t_1,t_2)$ is absolutely continuous in $t_1[t_2]$ for almost all $t_2[t_1]$, $0 \le t_1, t_2 \le \pi$;

(ii) the partial derivatives $\partial\mathbf{x}/\partial t_1$, $\partial\mathbf{x}/\partial t_2$ which exist almost everywhere are in $\mathcal{S}$,

and let L be the gradient operator

$$\ell\,\mathbf{x} = \begin{bmatrix} \dfrac{\partial\mathbf{x}}{\partial t_1} \\ \dfrac{\partial\mathbf{x}}{\partial t_2} \end{bmatrix}$$

with domain

$$D(L) = \left\{\mathbf{x} \in \mathcal{S}^1 : \begin{cases} \mathbf{x}(0,t_2) = \mathbf{x}(\pi,t_2) = 0 \text{ for almost all } t_2 , \\ \mathbf{x}(t_1,0) = \mathbf{x}(t_1,\pi) = 0 \text{ for almost all } t_1 , \end{cases} \quad 0 \le t_1, t_2 \le \pi\right\}$$

Then:

(a) $L \in \mathcal{C}(\mathcal{S}, \mathcal{S}\times\mathcal{S})$, $\overline{D(L)} = \mathcal{S}$.

(b) The adjoint L^* is the negative of the divergence operator

$$\ell^*\,\mathbf{y} = \ell^*\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = -\frac{\partial y_1}{\partial t_1} - \frac{\partial y_2}{\partial t_2}$$

with

$$D(L^*) = \{\mathbf{y} \in \mathcal{S} \times \mathcal{S} : \mathbf{y} \in C^1\} .$$

(c) L^*L is the negative of the Laplacian operator

$$L^*L = -\left[\frac{\partial^2}{\partial t_1^2} + \frac{\partial^2}{\partial t_2^2}\right]$$

(d) The Green's function of L^*L is

$$G(s_1, s_2, t_1, t_2) = \frac{4}{\pi^2} \sum_{m,n=1}^{\infty} \frac{1}{m^2+n^2} \sin(ms_1)\sin(ns_2)\sin(mt_1)\sin(nt_2) ,$$
$$0 \le s_i, t_j \le \pi .$$

(e) If

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathcal{S} \times \mathcal{S} ,$$

then

$$(L^\dagger \mathbf{y})(t_1, t_2) = \sum_{j=1}^{2} \int_0^\pi \int_0^\pi \frac{\partial}{\partial s_j} G(s_1, s_2, t_1, t_2)\, y_j(s_1, s_2)\, ds_1\, ds_2$$

(Landesman [**492**, Section 5]).

Ex. 29. *Ordinary linear differential equations with homogeneous boundary conditions.* Let
$\mathcal{S}$ = the real space $L^2[a, b]$ of real valued functions,
$C^k[a, b]$ = the real valued functions on $[a, b]$ with k derivatives and

$$\mathbf{x}^{(k)} = \frac{d^k \mathbf{x}}{dt^k} \in C[a, b] ,$$

$\mathcal{S}^k = \{\mathbf{x} \in C^{k-1}[a, b] : \mathbf{x}^{(k-1)}$ absolutely continuous , $\mathbf{x}^{(k)} \in \mathcal{S}\}$
and let L be the operator

$$\ell = \sum_{i=0}^{n} a_i(t) \left(\frac{d}{dt}\right)^i , \; a_i \in C^i[a, b] , \; i = 0, 1, \dots, n , \tag{53}$$
$$a_n(t) \ne 0 , \; a \le t \le b ,$$

with domain $D(L)$ consisting of all $\mathbf{x} \in \mathcal{S}^n$ which satisfy

$$M\widehat{\mathbf{x}} = 0 , \tag{54}$$

where $M \in \mathbb{R}_m^{m \times 2n}$ is a matrix with a specific null space $N(M)$, and $\widehat{\mathbf{x}} \in \mathbb{R}^{2n}$ is the boundary vector

$$\widehat{\mathbf{x}}^T = [\mathbf{x}(a), \mathbf{x}'(a), \cdots, \mathbf{x}^{(n-1)}(a); \mathbf{x}(b), \mathbf{x}'(b), \cdots, \mathbf{x}^{(n-1)}(b)] .$$

Finally let $\widetilde{L}$ be the operator ℓ of (53) with $D(\widetilde{L}) = \mathcal{S}^n$. Then
(a) $L \in \mathcal{C}(\mathcal{S}, \mathcal{S})$, $\overline{D(L)} = \mathcal{S}$.
(b) $\dim N(\widetilde{L}) = n = \dim N(\widetilde{L}^*)$.
(c) $N(L) \subset N(\widetilde{L})$, $N(L^*) \subset N(\widetilde{L}^*)$, hence $\dim N(L) \le n$ and $\dim N(L^*) \le n$.
(d) $R(L)$ is closed.
(e) The restriction $L_0 = L_{[C(L)]}$ of L to its carrier is a one–to–one mapping of $C(L)$ onto $R(L)$;

$$L_0 \in \mathcal{C}(C(L), R(L)) .$$

(f) $L_0^{-1} \in \mathcal{B}(R(L), C(L))$.
(g) $L^\dagger$, the extension of L_0^{-1} to all of $\mathcal{S}$ with $N(L^\dagger) = R(L)^\perp$ is bounded and satisfies

$$LL^\dagger \mathbf{y} = P_{R(L)} \mathbf{y} , \quad \text{for all } \mathbf{y} \in \mathcal{S}$$
$$L^\dagger L \mathbf{x} = P_{N(L)^\perp} \mathbf{x} , \quad \text{for all } \mathbf{x} \in D(L) .$$

For proofs of (a) and (d) see Halperin [**363**] and Schwartz [**739**]. The proof of (e) is contained in Section 2(F), and (f) follows from the closed graph theorem (Locker [**516**]).

Ex. 30. For L as in Ex. 29, find the generalized Green's function which corresponds to $L^\dagger$, i.e., find the kernel $L^\dagger(s,t)$ such that

$$(L^\dagger \mathbf{y})(s) = \int_a^b L^\dagger(s,t)\,\mathbf{y}(t)\,dt \quad \text{for all } \mathbf{y} \in D(L^\dagger) = \mathcal{S}\ .$$

SOLUTION. A generalized Green's function of $\widetilde{L}$ is (see Coddington and Levinson [**201**, Theorem 6.4])

$$\widetilde{G}(s,t) = \begin{cases} \sum\limits_{j=1}^{n} \dfrac{\mathbf{x}_j(s)\det(X_j(t))}{a_n(t)\det(X(t))}\ , & a \le t \le s \le b \\ 0\ , & a \le s \le t \le b \end{cases} \tag{55}$$

where

$\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ is an orthonormal basis of $N(L)$,

$\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n\}$ is an orthonormal basis of $N(\widetilde{L})$,

$$X(t) = \left[\mathbf{x}_j^{(i-1)}(t)\right]\ , \quad i,j = 1, \dots, n\ ,$$

$X_j(t)$ is the matrix obtained from $X(t)$ by replacing the j th column by $[0,0,\dots,0,1]^T$.

Since $R(L) \subset R(\widetilde{L})$ it follows, for any $\mathbf{y} \in R(L)$, that the general solution of

$$L\mathbf{x} = \mathbf{y}$$

is

$$\mathbf{x}(s) = \int_a^b \widetilde{G}(s,t)\,\mathbf{y}(t)\,dt + \sum_{i=1}^{n} c_i\,\mathbf{x}_i(s)\ , \quad c_i \text{ arbitrary}\ . \tag{56}$$

Writing the particular solution $L^\dagger\mathbf{y}$ in the form (56)

$$L^\dagger\mathbf{y} = \mathbf{x}_0 + \sum_{i=1}^{n} c_i\,\mathbf{x}_i\ , \tag{57}$$

$$\mathbf{x}_0(s) = \int_a^b \widetilde{G}(s,t)\,\mathbf{y}(t)\,dt\ ,$$

we determine its coefficients $\{c_1, \dots, c_n\}$ as follows:

(a) The coefficients $\{c_1, \dots, c_k\}$ are determined by $L^\dagger\mathbf{y} \in N(L)^\perp$, since, by (57),

$$\langle L^\dagger\mathbf{y}, \mathbf{x}_j\rangle = 0 \implies c_j = -\langle \mathbf{x}_0, \mathbf{x}_j\rangle\ , \quad j = 1, \dots, k\ .$$

(b) The remaining coefficients $\{c_{k+1}, \dots, c_n\}$ are determined by the boundary condition (54). Indeed, writing (57) as

$$L^\dagger\mathbf{y} = \mathbf{x}_0 + X\mathbf{c}\ , \quad \mathbf{c}^T = [c_1, \dots, c_n]\ ,$$

it follows from (54) that

$$M\widehat{\mathbf{x}}_0 + M\widehat{X}\mathbf{c} = \mathbf{0}\ , \quad \text{where } \widehat{X} = \begin{bmatrix} X(a) \\ X(b) \end{bmatrix}\ . \tag{58}$$

A solution of (58) is

$$\mathbf{c} = -(M\widehat{X})^{(1)} M\widehat{\mathbf{x}}_0\ , \tag{59}$$

where $(M\widehat{X})^{(1)} \in \mathbb{R}^{n\times m}$ is any $\{1\}$–inverse of $M\widehat{X} \in \mathbb{R}^{m\times n}$. Now $\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subset D(L)$, and therefore

$$M\widehat{X} = [O\ B]\ , \quad B \in \mathbb{R}^{m\times(n-k)}_{n-k}\ .$$

Thus, we may use in (59),

$$(M\widehat{X})^{(1)} = \begin{bmatrix} O \\ B^{(1)} \end{bmatrix}\ , \quad \text{for any } B^{(1)} \in B\{1\}\ ,$$

obtaining

$$\mathbf{c} = -\begin{bmatrix} O \\ B^{(1)} \end{bmatrix} M\widehat{\mathbf{x}}_0\ ,$$

which uniquely determines $\{c_{k+1}, \dots, c_n\}$.

Substituting these coefficients $\{c_1, \dots, c_n\}$ in (56) finally gives $L^\dagger(s,t)$ (Locker [**516**]). □

EX. 31. *The vector case.* Let $\mathcal{S}_n$ and $\mathcal{S}_n^k$ denote the spaces of n–dimensional vector functions whose components belong to $\mathcal{S}$ and $\mathcal{S}^k$, respectively, of Ex. 29. Let L be the differential operator

$$\ell\,\mathbf{x} = A_1(t)\frac{d\mathbf{x}}{dt} + A_0(t)\mathbf{x}\ , \quad a \le t \le b \tag{50}$$

where A_0, A_1 are $n \times n$ matrix functions satisfying[1]

(i) $A_0(t)$ is continuous on $[a,b]$.

(ii) $A_1(t)$ is continuously differentiable and nonsingular on $[a,b]$,

with domain $D(L)$ consisting of those vector functions $\mathbf{x} \in \mathcal{S}_n^1$ which satisfy

$$M\widehat{\mathbf{x}} = \mathbf{0}\ , \tag{54}$$

where $M \in \mathbb{R}_m^{m\times 2n}$ is a matrix with a specified null space $N(M)$ and $\widehat{\mathbf{x}} \in \mathbb{R}^{2n}$ is the boundary vector

$$\widehat{\mathbf{x}} = \begin{bmatrix}\mathbf{x}(a)\\ \mathbf{x}(b)\end{bmatrix}\ . \tag{60}$$

Let $\widetilde{L}$ be the differential operator (50) with domain $D(\widetilde{L}) = \mathcal{S}_n^1$. Then

(a) $L \in \mathcal{C}(\mathcal{S}_n, \mathcal{S}_n)\ , \quad \overline{D(L)} = \mathcal{S}_n$.

(b) The adjoint of L is the operator L^* defined by

$$\ell^*\mathbf{y} = -\frac{d}{dt}(A_1^*(t)\mathbf{y}) + A_0^*(t)\mathbf{y} \tag{61}$$

on its domain

$$\begin{aligned} D(L^*) &= \{\mathbf{y} \in \mathcal{S}_n^1 : \mathbf{y}^*(b)\mathbf{x}(b) - \mathbf{y}^*(a)\mathbf{x}(a) = 0 \text{ for all } \mathbf{x} \in D(L)\} \\ &= \left\{\mathbf{y} \in \mathcal{S}_n^1 : P^*\begin{bmatrix} I & O \\ O & -I\end{bmatrix}\widehat{\mathbf{y}} = \mathbf{0} \text{ for any } P \in \mathbb{R}_{2n-m}^{(2n-m)\times 2n} \text{ with} MP = O\right\} \end{aligned} \tag{62}$$

(c) $\dim N(\widetilde{L}) = n$.

(d) Let

$$k = \dim N(L) \quad \text{and} \quad k^* = \dim N(L^*)\ .$$

Then

$$\max\{0, n-m\} \le k \le \min\{n, 2n-m\}$$

and

$$k + m = k^* + n\ .$$

(e) $R(L) = N(L^*)^\perp\ , \quad R(L^*) = N(L)^\perp$,

hence both $R(L)$ and $R(L^*)$ are closed.

(f) Let

$$X(t) = [\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)]$$

be a *fundamental matrix* of $\widetilde{L}$, i.e., let the vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ form a basis of $N(\widetilde{L})$. Then

$$\widetilde{G}(s,t) = \tfrac{1}{2}\,\mathrm{sign}(s-t)X(s)X(t)^{-1} \tag{63}$$

is a generalized Green's matrix of $\widetilde{L}$.

(g) Let $(M\widehat{X})^{(1)}$ be any $\{1\}$–inverse of $M\widehat{X}$ where $\widehat{X} = \begin{bmatrix}X(a)\\ X(b)\end{bmatrix}$. Then

$$G(s,t) = \tfrac{1}{2}\,X(s)\left(\mathrm{sign}(s-t)I - (M\widehat{X})^{(1)}M\begin{bmatrix} I & O \\ O & -I\end{bmatrix}\widehat{X}\right)X(t)^{-1} \tag{64}$$

[1] Much weaker regularity conditions will do; see, e.g., Reid [**690**] and [**692**, Chapter III].

is a generalized Green's matrix of L (Reid [**691**] and [**692**, Chapter III]).

PROOF OF (g). For any $\mathbf{y} \in R(L)$, the general solution of

$$L\mathbf{x} = \mathbf{y} \tag{51}$$

is

$$\mathbf{x}(s) = \int_a^b \widetilde{G}(s,t)\,\mathbf{y}(t)\,dt + \sum_{i=1}^{n} c_i \mathbf{x}_i(s) \tag{56}$$

or

$$\mathbf{x} = \mathbf{x}_0 + X\mathbf{c}\,, \quad \mathbf{c}^T = [c_1, \dots, c_n]$$

and from (54) it follows that

$$\mathbf{c} = -(M\widehat{X})^{(1)} M \widehat{\mathbf{x}}_0 \tag{59}$$

and (64) follows by substituting (59) in (56). □

EX. 32. The differential expression

$$\ell\, x = \sum_{i=1}^{n} a_i(t)\, \frac{d^i x}{dt^i}\,, \quad x \text{ scalar function} \tag{53}$$

is a special case of

$$\ell \mathbf{x} = A_1(t)\frac{d\mathbf{x}}{dt} + A_0(t)\mathbf{x}\,, \quad \mathbf{x} \text{ vector function}\,. \tag{50}$$

EX. 33. *The class of all generalized Green's functions.* Let L be as in Ex. 31 and let $X_0(t)$ and $Y_0(t)$ be $n \times k$ and $n \times k^*$ matrix functions whose columns are bases of $N(L)$ and $N(L^*)$, respectively. Then a kernel $H(s,t)$ is a generalized Green's matrix of L if and only if

$$H(s,t) = G(s,t) + X_0(s)A^*(t) + B(s)Y_0^*(t)\,, \tag{65}$$

where $G(s,t)$ is any generalized Green's matrix of L (in particular (64)), and $A(t)$ and $B(s)$ are $n \times k$ and $n \times k^*$ matrix functions which are Lebesgue measurable and essentially bounded. (Reid [**690**]).

EX. 34. Let $X_0(t)$ and $Y_0(t)$ be as in Ex. 33. If $\Theta(t)$ and $\Psi(t)$ are Lebesgue measurable and essentially bounded matrix functions such that the matrices

$$\int_a^b \Theta^*(t)\, X_0(t)\, dt\,, \quad \int_a^b Y_0^*(t)\, \Psi(t)\, dt$$

are nonsingular, then L has a unique generalized Green's function $G_{\Theta,\Psi}$ such that

$$\begin{matrix} \int_a^b \Theta^*(s)\, G(s,t)\, ds = O \\ \int_a^b G(s,t)\, \Psi(t)\, dt = O \end{matrix} \qquad a \le s,t \le b\,. \tag{66}$$

Thus the generalized inverse determined by $G_{\Theta,\Psi}$ has null space spanned by the columns of Ψ and range which is the orthogonal complement of the columns of Θ. Compare with Section 2.6. (Reid [**690**]).

EX. 35. *Existence and properties of* $L^\dagger$. If in Ex. 34 we take

$$\Theta = X_0\,, \quad \Psi = Y_0\,,$$

then we get a generalized inverse of L which has the same range and null space as L^*. This generalized inverse is the analog of the *Moore–Penrose inverse* of L and will likewise be denoted by $L^\dagger$.

Show that $L^\dagger$ satisfies the four *Penrose equations* (1.1)–(1.4) as far as can be expected.

(a) $LL^\dagger L = L\,,$

(b) $L^\dagger L L^\dagger = L^\dagger\,,$

(c) $LL^\dagger = P_{R(L)}\,,$

$(LL^\dagger)^* = P_{R(L)}$ on $D(L^*)\,,$

(d) $L^\dagger L = P_{R(L^*)}$ on $D(L)\,,$

$(L^\dagger L)^* = P_{R(L^*)}$ (Loud [**518**], [**519**]).

Ex. 36. *Loud's construction of* $L^\dagger$. Just as in the matrix case (see Theorem 2.12(c) and Ex. 2.38) it follows here that

$$L^\dagger = P_{R(L^*)}\, G\, P_{R(L)} \ , \tag{67}$$

where G is any generalized Green's matrix.

In computing $P_{R(L^*)}$ and $P_{R(L)}$ we use Ex. 31(e) to obtain

$$P_{R(L^*)} = I - P_{N(L)} \ , \quad P_{R(L)} = I - P_{N(L^*)} \ . \tag{68}$$

Here $P_{N(L)}$ and $P_{N(L^*)}$ are integral operators of the first kind with kernels

$$K_{N(L)} = X_0(s)\left(\int_a^b X_0^*(u)X_0(u)du\right)^{-1} X_0^*(t) \tag{69}$$

and

$$K_{N(L^*)} = Y_0(s)\left(\int_a^b Y_0^*(u)Y_0(u)du\right)^{-1} Y_0^*(t) \ , \tag{70}$$

respectively, where X_0 and Y_0 are as in Ex. 33.

Thus, for any generalized Green's matrix $G(s,t)$, $L^\dagger$ has the kernel

$$L^\dagger(s,t) = G(s,t) - \int_a^b K_{N(L)}(s,u)\, G(u,t)\, du - \int_a^b G(s,u)\, K_{N(L^*)}(u,t)\, du \tag{71}$$
$$+ \int_a^b\int_a^b K_{N(L)}(s,u)\, G(u,v)\, K_{N(L^*)}(v,t)\, du\, dv \quad \text{(Loud [\textbf{519}])}.$$

Ex. 37. Let L be the differential operator given by

$$\ell\,\mathbf{x} = \mathbf{x}' - B(t)\mathbf{x} \ , \quad 0 \le t \le 1$$

with boundary conditions

$$\mathbf{x}(0) = \mathbf{x}(1) = \mathbf{0} \ .$$

Then the adjoint L^* is given by

$$\ell^*\,\mathbf{y} = -\mathbf{y}' - B(t)^*\mathbf{y}$$

with no boundary conditions.

Let $X(t)$ be a fundamental matrix for

$$\ell\,\mathbf{x} = \mathbf{0} \ .$$

Then $X(t)^{*-1}$ is a fundamental matrix for

$$\ell^*\,\mathbf{y} = \mathbf{0} \ .$$

Now $N(L) = \{\mathbf{0}\}$ and therefore $K_{N(L)} = O$. Also, $N(L^*)$ is spanned by the columns of $X(t)^{*-1}$, so by (70)

$$K_{N(L^*)}(s,t) = X(s)^{*-1}\left(\int_0^1 X(u)X(u)^{*-1}du\right) X(t)^{-1} \ . \tag{72}$$

A generalized Green's matrix for L is

$$G(s,t) = \begin{cases} X(s)X(t)^{-1} \ , & 0 \le s < t \le 1 \\ O \ , & 0 \le t < s \le 1 \end{cases} \tag{73}$$

Finally, by (71),

$$L^\dagger(s,t) = G(s,t) - \int_0^1 G(s,u)\, K_{N(L^*)}(u,t)\, du \ ,$$

with G and $K_{N(L^*)}$ given by (73) and (72), respectively (Loud [**519**, pp. 201–202]).

5. Minimal properties of generalized inverses

In this section, which is based on Erdélyi and Ben–Israel [**259**], we develop certain minimal properties of generalized inverses of operators between Hilbert spaces, analogous to the matrix case studied in Chapter 3.

DEFINITION 4. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and consider the linear equation

$$T\mathbf{x} = \mathbf{y} \; . \tag{74}$$

If the infimum

$$\|T\mathbf{x}' - \mathbf{y}\| = \inf_{\mathbf{x}\in D(T)} \|T\mathbf{x} - \mathbf{y}\| \tag{75}$$

is attained by a vector $\mathbf{x}' \in D(T)$, then $\mathbf{x}'$ is called an *extremal solution* of (74). Among the extremal solutions there may exist a unique vector $\mathbf{x}_0$ of least norm

$$\|\mathbf{x}_0\| < \|\mathbf{x}'\| \; ,$$

for all extremal solutions $\mathbf{x}' \neq \mathbf{x}_0$. Then $\mathbf{x}_0$ is called the *least extremal solution.*

Other names for extremal solutions are *virtual solutions* (Tseng [**828**]), and *approximate solutions.*

Example 38 shows that extremal solutions need not exist. Their existence is characterized in the following theorem.

THEOREM 5. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then

$$T\mathbf{x} = \mathbf{y} \tag{74}$$

has an extremal solution if and only if

$$P_{\overline{R(T)}}\mathbf{y} \in R(T) \; . \tag{76}$$

PROOF. For every $\mathbf{x} \in D(T)$

$$\begin{aligned}\|T\mathbf{x} - \mathbf{y}\|^2 &= \|P_{\overline{R(T)}}(T\mathbf{x} - \mathbf{y})\|^2 + \|P_{R(T)^\perp}(T\mathbf{x} - \mathbf{y})\|^2 \\ &= \|P_{\overline{R(T)}}(T\mathbf{x} - \mathbf{y})\|^2 + \|P_{R(T)^\perp}\mathbf{y}\|^2 \; .\end{aligned}$$

Thus

$$\|T\mathbf{x} - \mathbf{y}\| \geq \|P_{R(T)^\perp}\mathbf{y}\| \; , \quad \text{for all } \mathbf{x} \in D(T) \; ,$$

with equality if and only if

$$T\mathbf{x} = P_{\overline{R(T)}}\,\mathbf{y} \; . \tag{77}$$

Clearly,

$$\inf_{\mathbf{x}\in D(T)} \|T\mathbf{x} - \mathbf{y}\| = P_{\overline{R(T)}}\,\mathbf{y} \; , \tag{78}$$

which is attained if and only if (77) is satisfied for some $\mathbf{x} \in D(T)$. □

See also Ex. 45.

The existence of extremal solutions does not guarantee the existence of a least extremal solution; see, e.g., Ex. 40. Before settling this issue we require

LEMMA 4. Let $\mathbf{x}'$ and $\mathbf{x}''$ be extremal solutions of (74). Then

(a) $P_{N(T)^\perp}\mathbf{x}' = P_{N(T)^\perp}\mathbf{x}''$

(b) $P_{\overline{N(T)}}\,\mathbf{x}' \in N(T)$ if and only if $P_{\overline{N(T)}}\,\mathbf{x}'' \in N(T)$.

PROOF. (a) From (77),

$$T\mathbf{x}' = T\mathbf{x}'' = P_{\overline{R(T)}}\mathbf{y}$$

and hence

$$T(\mathbf{x}' - \mathbf{x}'') = \mathbf{0} \; , \tag{79}$$

proving (a).

(b) From (79),

$$\mathbf{x}' - \mathbf{x}'' = P_{\overline{N(T)}}(\mathbf{x}' - \mathbf{x}'')$$

and then

$$P_{\overline{N(T)}}\mathbf{x}' = P_{\overline{N(T)}}\mathbf{x}'' + (\mathbf{x}' - \mathbf{x}'') \; ,$$

proving (b). □

The existence of the least extremal solution is characterized in the following:

THEOREM 6. (Erdélyi and Ben–Israel [**259**]). Let $\mathbf{x}$ be an extremal solution of (74). There exists a least extremal solution if and only if

$$P_{\overline{N(T)}}\mathbf{x} \in N(T) \tag{80}$$

in which case, the least extremal solution is

$$\mathbf{x}_0 = P_{N(T)^\perp}\mathbf{x} \; . \tag{81}$$

PROOF. Let $\mathbf{x}'$ be an extremal solution of (74). Then

$$\begin{aligned} \|\mathbf{x}'\|^2 &= \|P_{\overline{N(T)}}\mathbf{x}'\|^2 + \|P_{N(T)^\perp}\mathbf{x}'\|^2 \\ &= \|P_{\overline{N(T)}}\mathbf{x}'\|^2 + \|P_{N(T)^\perp}\mathbf{x}\|^2 \; , \quad \text{by Lemma 4} \; , \end{aligned}$$

proving that

$$\|\mathbf{x}'\| \geq \|P_{N(T)^\perp}\mathbf{x}\|$$

with equality if and only if

$$P_{\overline{N(T)}}\mathbf{x}' = \mathbf{0} \; . \tag{82}$$

If. Let condition (80) be satisfied and define

$$\mathbf{x}_0 = \mathbf{x} - P_{\overline{N(T)}}\mathbf{x} \; .$$

Then $\mathbf{x}_0$ is an extremal solution since

$$T\mathbf{x}_0 = T\mathbf{x} \; .$$

Also

$$P_{\overline{N(T)}}\mathbf{x}_0 = \mathbf{0} \; ,$$

which, by (82), proves that $\mathbf{x}_0$ is the least extremal solution.

Only if. Let $\mathbf{x}_0$ be the least extremal solution of (74). Then, by (82),

$$\mathbf{x}_0 = P_{\overline{N(T)}}\mathbf{x}_0 + P_{N(T)^\perp}\mathbf{x}_0 = P_{N(T)^\perp}\mathbf{x} \; ,$$

and hence

$$\mathbf{x}_0 = \mathbf{x} - P_{\overline{N(T)}}\mathbf{x} \; .$$

But

$$T\mathbf{x}_0 = T\mathbf{x} ,$$

since both $\mathbf{x}_0$ and $\mathbf{x}$ are extremal solutions, and therefore

$$TP_{\overline{N(T)}} = \mathbf{0} ,$$

proving (80). □

As in the matrix case (see Corollary 3.3, p. 95), here too a unique generalized inverse is characterized by the property that it gives the least extremal solution whenever it exists. We define this inverse as follows:

DEFINITION 5. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, let

$$C(T) = D(T) \cap N(T)^{\perp} , \tag{2}$$

$$B(T) = D(T) \cap \overline{N(T)} , \tag{83}$$

and let $A(T)$ be a subspace satisfying

$$D(T) = A(T) \oplus \left(B(T) \overset{\perp}{\oplus} C(T) \right) . \tag{84}$$

(Examples 43 and 44 below show that, in the general case, this complicated decomposition cannot be avoided.) Let

$$G_0 = \{\{\mathbf{x}, T\mathbf{x}\} : \mathbf{x} \in C(T)\} , \quad G_1 = G(T)^{\perp} \cap \mathcal{H}_{0,2} = J_2 R(T)^{\perp} .$$

The *extremal g.i.* of T, denoted by $T_e^{\dagger}$, is defined by its inverse graph

$$G_0 + G_1 = \{\{\mathbf{x}, T\mathbf{x} + \mathbf{z}\} : \mathbf{x} \in C(T), \mathbf{z} \in R(T)^{\perp}\} .$$

The following properties of $T_e^{\dagger}$ are easy consequences of the above construction.

THEOREM 7. (Erdélyi and Ben–Israel [**259**]). Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then

(a) $D(T_e^{\dagger}) = T(C(T)) \overset{\perp}{\oplus} R(T)^{\perp}$, and in general, $R(T) \not\subset D(T_e^{\dagger})$.

(b) $R(T_e^{\dagger}) = C(T)$.

(c) $N(T_e^{\dagger}) = R(T)^{\perp}$.

(d) $TT_e^{\dagger}\mathbf{y} = P_{\overline{R(T)}}\mathbf{y}$, for all $\mathbf{y} \in D(T_e^{\dagger})$.

(e) $T_e^{\dagger}T\mathbf{x} = P_{\overline{R(T_e^{\dagger})}}\mathbf{x}$, for all $\mathbf{x} \in N(T) \overset{\perp}{\oplus} C(T)$. □

See also Exs. 41–42 below.

The extremal g.i. $T_e^{\dagger}$ is characterized in terms of the least extremal solution, as follows:

THEOREM 8. (Erdélyi and Ben–Israel [**259**]). The least extremal solution $\mathbf{x}_0$ of (74) exists if and only if

$$\mathbf{y} \in D(T_e^{\dagger}) , \tag{85}$$

in which case

$$\mathbf{x}_0 = T_e^{\dagger}\mathbf{y} . \tag{86}$$

PROOF. Assume (85). By Theorem 7(a)

$$P_{\overline{R(T)}}\mathbf{y} = \mathbf{y}_0 \in T(C(T)) \subset R(T) ,$$

and, by Theorem 5, extremal solutions do exist. Let $\mathbf{x}_0$ be the unique vector in $C(T)$ such that

$$P_{\overline{R(T)}}\mathbf{y} = \mathbf{y}_0 = T\mathbf{x}_0 .$$

Then, by Theorem 3(a), (c), and (e),

$$T_e^\dagger \mathbf{y} = T_e^\dagger \mathbf{y}_0 = T_e^\dagger T \mathbf{x}_0 = \mathbf{x}_0 \ ,$$

and by Theorem 3(d),

$$\|T\mathbf{x}_0 - \mathbf{y}\| = \|TT_e^\dagger \mathbf{y} - \mathbf{y}\| = \|P_{\overline{R(T)}}\mathbf{y} - \mathbf{y}\| = \|P_{R(T)^\perp}\mathbf{y}\| \ ,$$

which, by (78), shows that $\mathbf{x}_0$ is an extremal solution. Since

$$\mathbf{x}_0 \in R(T_e^\dagger) \subset N(T)^\perp \ ,$$

it follows, from Lemma 4, that

$$\mathbf{x}_0 = P_{N(T)^\perp}\mathbf{x}$$

for any extremal solution $\mathbf{x}$ of (74). By Theorem 6, $\mathbf{x}_0$ is the least extremal solution.

Conversely, let $\mathbf{x}_0$ be the least extremal solution whose existence we assume. By Theorem 2, $\mathbf{x}_0 \in C(T)$, and by Theorem 3(e),

$$T_e^\dagger T \mathbf{x}_0 = \mathbf{x}_0 \ .$$

Since $\mathbf{x}_0$ is an extremal solution, it follows from (77) that

$$T\mathbf{x}_0 = P_{\overline{R(T)}}\mathbf{y} \in T(C(T))$$

and therefore

$$\begin{aligned} \mathbf{x}_0 &= T_e^\dagger T \mathbf{x} = T_e^\dagger P_{\overline{R(T)}}\mathbf{y} \\ &= T_e^\dagger \mathbf{y} \ . \end{aligned}$$

□

If $N(T)$ is closed then $T_e^\dagger$ coincides with the maximal g.i. $T^\dagger$. Thus for closed operators, and in particular for bounded operators, $T_e^\dagger$ should be replaced by $T^\dagger$ in the statement of Theorem 8

5.1. Exercises and examples.

Ex. 38. *A linear equation without extremal solution.* Let T and $\mathbf{y}$ be as in Ex. 7. Then

$$T\mathbf{x} = \mathbf{y}$$

has no extremal solutions.

Ex. 39. It was noted in Ex. 8, that, in general, the Fredholm integral operator of the first kind has a nonclosed range. Consider the kernel

$$G(s,t) = \begin{cases} s(1-t) \ , & 0 \le s \le t \le 1 \\ t(1-s) \ , & 0 \le t \le s \le 1 \end{cases}$$

which is a generalized Green's function of the operator

$$-\frac{d^2}{dt^2} \ , \quad 0 \le t \le 1 \ .$$

Let $T \in \mathcal{B}(L^2[0,1], L^2[0,1])$ be defined by

$$(T\mathbf{x})(s) = \int_0^1 G(s,t)\,\mathbf{x}(t)\,dt \ .$$

Show that there exists a $\mathbf{y} \in L^2[0,1]$ for which

$$T\mathbf{x} = \mathbf{y}$$

has no extremal solution.

Ex. 40. *An equation without a least extremal solution.* Consider the unbounded functional on $L^2[0,\infty]$

$$T\mathbf{x} = \int_0^\infty t\mathbf{x}(t)\,dt$$

discussed in Ex. 2. Then the equation

$$T\mathbf{x} = 1$$

is consistent, and each of the functions

$$\mathbf{x}_n(t) = \begin{cases} \frac{1}{nt}\,, & 1 \le t \le n+1 \\ 0\,, & \text{otherwise} \end{cases}$$

is a solution, $n = 1, 2, \dots$. Since

$$\|\mathbf{x}_n\|^2 = \int_1^{n+1} \frac{1}{(nt)^2}\,dt = \frac{1}{n(n+1)} \to 0\,,$$

there is no extremal solution of least norm.

Ex. 41. *Properties of* $(T_e^\dagger)^\dagger$. By Theorem 7(a) and (c), it follows that $D(T_e^\dagger)$ is decomposable with respect to $N(T_e^\dagger)$. Thus $T_e^\dagger$ has a maximal (Tseng) g.i., denoted by $T_e^{\dagger\dagger}$. Some of its properties are listed below.

(a) $G(T_e^{\dagger\dagger}) = \{\{\mathbf{x}+\mathbf{z}, T\mathbf{x}\} : \mathbf{x} \in C(T),\, \mathbf{z} \in C(T)^\perp\}$.

(b) $D(T_e^{\dagger\dagger}) = C(T) \overset{\perp}{\oplus} C(T)^\perp$.

(c) $R(T_e^{\dagger\dagger}) = T(C(T))$.

(d) $N(T_e^{\dagger\dagger}) = C(T)^\perp$.

Ex. 42. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and let

$$D_0(T) = N(T) \overset{\perp}{\oplus} C(T)\,.$$

Then

(a) $D(T_e^{\dagger\dagger}) = C(T) \overset{\perp}{\oplus} \overline{N(T)} \overset{\perp}{\oplus} D_0(T)^\perp$, a refinement of Ex. 41(b).

(b) $D_0(T) \subset D(T) \cap D(T_e^{\dagger\dagger})$ and $T_{[D_0(T)]} = (T_e^{\dagger\dagger})_{[D_0(T)]}$.

(c) $T_e^{\dagger\dagger}$ is an extension of T if and only if $D(T)$ is decomposable with respect to $N(T)$, in which case $T_e^{\dagger\dagger}$ is an extension by zero to $\overline{N(T)} \overset{\perp}{\oplus} D(T)^\perp$.

Ex. 43. *An example of* $A(T) \neq \{\mathbf{0}\}$, $A(T) \subset D(T_e^{\dagger\dagger})$. Let T be the operator defined in Ex. 4. Then, by Ex. 4,

$$\begin{aligned} B(T) &= D(T) \cap \overline{N(T)} \\ &= D \cap (\overline{D \cap F}) \\ &= D \cap F \\ &= N(T)\,, \end{aligned}$$

and

$$C(T) = \{\mathbf{0}\}\,,$$

showing that

$$A(T) \neq \{\mathbf{0}\}\,, \quad \text{by (84)}\,.$$

Thus

$$A(T) = A \quad \text{of Ex. 4,}$$

and

$$D(T_e^\dagger) = A^\perp = N(T_e^\dagger)\,.$$

Finally, from $C(T)^\perp = \mathcal{H}$,

$$D(T_e^{\dagger\dagger}) = \mathcal{H} \supset A$$

with

$$N(T_e^{\dagger\dagger}) = \mathcal{H} .$$

Ex. 44. *An example of* $A(T) \neq \{\mathbf{0}\}$, $A(T) \cap D(T_e^{\dagger\dagger}) = \{\mathbf{0}\}$. Let $\mathcal{H}$ be a Hilbert space and let M, N be subspaces of $\mathcal{H}$ such that

$$M \neq \overline{M} , \quad N \neq \overline{N} \subset M^{\perp} .$$

Choose

$$\mathbf{y} \in \overline{M} \setminus M \quad \text{and} \quad \mathbf{z} \in M^{\perp} \setminus (N \overset{\perp}{\oplus} (N^{\perp} \cap M^{\perp})) ;$$

let

$$\mathbf{x} = \mathbf{y} + \mathbf{z}$$

and

$$D = M \oplus N \oplus [\mathbf{x}]$$

where $[\mathbf{x}]$ is the line spanned by $\mathbf{x}$. Define $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ on $D(T) = D$ by

$$T(\mathbf{u} + \mathbf{v} + \alpha\mathbf{x}) = \mathbf{v} + \alpha\mathbf{x} , \quad \mathbf{u} \in M , \quad \mathbf{v} \in N , \quad \alpha\mathbf{x} \in [\mathbf{x}] .$$

Then

$$C(T) = N , \quad N(T) = M , \quad A(T) = [\mathbf{x}]$$

and

$$\mathbf{x} \notin D(T_e^{\dagger\dagger}) .$$

Ex. 45. Let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$. Then

$$T\mathbf{x} = \mathbf{y} \tag{74}$$

has an extremal solution if and only if there is a positive scalar β such that

$$|\langle \mathbf{y}, \mathbf{z} \rangle|^2 \leq \beta \langle \mathbf{z}, AA^*\mathbf{z} \rangle , \quad \text{for every } \mathbf{z} \in N(AA^*)^{\perp}$$

(Tseng [**828**]; see also Holmes [**422**, Section 35]).

Ex. 46. Let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, $S \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_3)$ be normally solvable, and let

$$T_S = T_{[N(S)]}$$

denote the restriction of T to $N(S)$. If T_S is also normally solvable , then $T_S^{\dagger}$ is called the $N(S)$*–restricted pseudoinverse of* T. It is the unique solution X of the following five equations

$$\begin{aligned} SX &= O , \\ XTX &= X , \\ (TX)^* &= TX , \\ TXT &= T \quad \text{on } N(S) , \\ P_{N(S)}(XT)^* &= XT \quad \text{on } N(S) \end{aligned}$$

(Minamide and Nakamura [**553**]).

Ex. 47. Let T, S, and $T_S^{\dagger}$ be as in Ex. 46. Then for any $\mathbf{y}_0 \in \mathcal{H}_2$ and $\mathbf{z}_0 \in R(S)$, the least extremal solution of

$$T\mathbf{x} = \mathbf{y}_0$$

subject to

$$S\mathbf{x} = \mathbf{z}_0$$

is given by

$$\mathbf{x}_0 = T_S^{\dagger}(\mathbf{y}_0 - TS^{\dagger}\mathbf{z}_0) + S^{\dagger}\mathbf{z}_0$$

(Minamide and Nakamura [**553**]).

Ex. 48. Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ be Hilbert spaces, let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ with $R(T) = \mathcal{H}_2$ and let $S \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_3)$. For any $\mathbf{y} \in \mathcal{H}_2$, there is a unique $\mathbf{x}_0 \in \mathcal{H}_1$ satisfying

$$T\mathbf{x} = \mathbf{y} \tag{74}$$

and which minimizes the functional

$$\|S\mathbf{x}\|^2 + \|\mathbf{x}\|^2$$

over all solutions of (74). This $\mathbf{x}_0$ is given by

$$\mathbf{x}_0 = (I + S^*S)^{-1}T^\dagger \mathbf{y}_0$$

where $\mathbf{y}_0$ is the unique vector in $\mathcal{H}_2$ satisfying

$$\mathbf{y} = T(I + S^*S)^{-1}T^\dagger \mathbf{y}_0 \qquad \text{(Porter and Williams [\mathbf{649}])}.$$

Ex. 49. Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, T$, and S be as above. Then for any $\mathbf{y} \in \mathcal{H}_2, \mathbf{x}_1 \in \mathcal{H}_1$, and $\mathbf{y}_1 \in \mathcal{H}_2$ there is a unique $\mathbf{x}_0 \in \mathcal{H}_1$ which is a solution of

$$T\mathbf{x} = \mathbf{y} \tag{74}$$

and which minimizes

$$\|S\mathbf{x} - \mathbf{y}_1\|^2 + \|\mathbf{x} - \mathbf{x}_1\|^2$$

from among all solutions of (74). This $\mathbf{x}_0$ is given by

$$\mathbf{x}_0 = (I + S^*S)^{-1}(T^\dagger \mathbf{y}_0 + \mathbf{x}_0 + S^*\mathbf{y}_1)$$

where $\mathbf{y}_0$ is the unique vector in $\mathcal{H}_2$ satisfying

$$\mathbf{y} = T(I + S^*S)^{-1}(T^\dagger \mathbf{y}_0 + \mathbf{x}_1 + S^*\mathbf{y}_1) \qquad \text{(Porter and Williams [\mathbf{649}])}.$$

6. Series and integral representations and iterative computation of generalized inverses

Direct computational methods, in which the exact solution requires a finite number of steps (such as the elimination methods of Sections 7.2–7.4) cannot be used, in general, for the computation of generalized inverses of operators. The exceptions are operators with nice algebraic properties, such as the integral and differential operators of Exs. 18–37 with their finite–dimensional null spaces. In the general case, the only computable representations of generalized inverses involve infinite series, or integrals, approximated by suitable iterative methods. Such representations and methods are sampled in this section, based on Showalter and Ben–Israel [**761**], where the proofs, omitted here, can be found.

To motivate the idea behind our development consider the problem of minimizing

$$f(\mathbf{x}) = \langle A\mathbf{x} - \mathbf{y}, A\mathbf{x} - \mathbf{y}\rangle \ , \tag{87}$$

where $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ and $\mathcal{H}_1, \mathcal{H}_2$ are Hilbert spaces.

Treating $\mathbf{x}$ as a function $\mathbf{x}(t)$, $t \geq 0$, with $\mathbf{x}(0) = \mathbf{0}$, we differentiate (87):

$$\begin{aligned} \tfrac{d}{dt} f(\mathbf{x}) &= 2\Re\langle A\mathbf{x} - \mathbf{y}, A\dot{\mathbf{x}}\rangle \ , \quad \dot{\mathbf{x}} = \tfrac{d}{dt}\mathbf{x} \\ &= 2\Re\langle A^*(A\mathbf{x} - \mathbf{y}), \dot{\mathbf{x}}\rangle \end{aligned} \tag{88}$$

and setting

$$\dot{\mathbf{x}} = -A^*(A\mathbf{x} - \mathbf{y}) \ , \tag{89}$$

it follows from (88) that

$$\tfrac{d}{dt} f(\mathbf{x}) = -2\|A^*(A\mathbf{x} - \mathbf{y})\|^2 < 0 \ . \tag{90}$$

This version of the steepest descent method, given in Rosenbloom [**719**], results in $f(\mathbf{x}(t))$ being a monotone decreasing function of t, asymptotically approaching its infimum as $t \to \infty$. We expect $\mathbf{x}(t)$ to approach asymptotically $A^\dagger \mathbf{y}$, so by solving (89)

$$\mathbf{x}(t) = \int_0^t \exp\{-A^*A(t-s)\}A^*\mathbf{y}\,ds \tag{91}$$

and observing that $\mathbf{y}$ is arbitrary we get

$$A^\dagger = \lim_{t\to\infty} \exp\{-A^*A(t-s)\}A^*\,ds \tag{92}$$

which is the essence of Theorem 9.

Here as elsewhere in this section, the convergence is in the strong operator topology. Thus the limiting expression

$$A^\dagger = \lim_{t\to\infty} B(t) \quad \text{or} \quad B(t) \to A^\dagger \text{ or } t \to \infty \tag{93}$$

means that for all $\mathbf{y} \in D(A^\dagger)$

$$A^\dagger \mathbf{y} = \lim_{t\to\infty} B(t)\mathbf{y}$$

in the sense that

$$\lim_{t\to\infty} \|(A^\dagger - B(t))\mathbf{y}\| = 0 \; . \tag{94}$$

A numerical integration of (89) with suitably chosen step size similarly results in

$$A^\dagger = \sum_{k=0}^{\infty} (I - \alpha A^*A)^k \alpha A^* \; , \tag{95}$$

where

$$0 < \alpha < \frac{2}{\|A\|^2} \; , \tag{96}$$

which is the essence of Theorem 10.

In statements like (94) it is necessary do distinguish between points $\mathbf{y} \in \mathcal{H}_2$ relative to the given $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$. Indeed, the three cases

$$P_{\overline{R(A)}}\mathbf{y} \in R(AA^*) \; , \quad P_{\overline{R(A)}}\mathbf{y} \in (R(A) \setminus R(AA^*)) \; , \quad P_{\overline{R(A)}}\mathbf{y} \in (\overline{R(A)} \setminus R(A))$$

have different rates of convergence in (94). Here $\mathbf{x} \in (X \setminus Y)$ means $\mathbf{x} \in X$, $\mathbf{x} \notin Y$. We abbreviate these as follows:

$$\begin{aligned}
(\mathbf{y} \in \text{I}) \quad &\text{means} \quad P_{\overline{R(A)}}\mathbf{y} \in R(AA^*) \; , \\
(\mathbf{y} \in \text{II}) \quad &\text{means} \quad P_{\overline{R(A)}}\mathbf{y} \in (R(A) \setminus R(AA^*)) \; , \\
(\mathbf{y} \in \text{III}) \quad &\text{means} \quad P_{\overline{R(A)}}\mathbf{y} \in (\overline{R(A)} \setminus R(A)) \; .
\end{aligned} \tag{97}$$

We note that $A^\dagger \mathbf{y}$ is not defined for $(\mathbf{y} \in \text{III})$, a case which does not exist if $R(A)$ is closed.

Theorem 9. (Showalter and Ben–Israel [**761**]). Let $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ and define, for $t \geq 0$

$$\begin{aligned}
L_1(t) &= \int_0^t \exp\{-A^*A(t-s)\}\,ds \; , \\
L_2(t) &= \int_0^t \exp\{-AA^*(t-s)\}\,ds \; , \\
B(t) &= L_1(t)A^* = A^*L_2(t) \; .
\end{aligned} \tag{98}$$

Then:

(a) $\|(A^\dagger - B(t))\mathbf{y}\|^2 \leq \dfrac{\|A^\dagger \mathbf{y}\|^2 \|(AA^*)^\dagger \mathbf{y}\|^2}{\|(AA^*)^\dagger \mathbf{y}\|^2 + 2\|A^\dagger \mathbf{y}\|^2 t}$ if ($\mathbf{y} \in$ I) and $t \geq 0$.

(b) $\|(A^\dagger - B(t))\mathbf{y}\|^2$ is a decreasing function of $t \geq 0$, with limit zero as $t \to \infty$, if ($\mathbf{y} \in$ II).

(c) $\|(P_{\overline{R(A)}} - AB(t))\mathbf{y}\|^2 \leq \dfrac{\|\mathbf{y}\|^2 \|A^\dagger \mathbf{y}\|^2}{\|A^\dagger \mathbf{y}\|^2 + 2\|\mathbf{y}\|^2 t}$ if ($\mathbf{y} \in$ I) or ($\mathbf{y} \in$ II), and $t \geq 0$.

(d) $\|(P_{\overline{R(A)}} - AB(t))\mathbf{y}\|^2$ is a decreasing function of $t \geq 0$, with limit zero as $t \to \infty$, if ($\mathbf{y} \in$ III). □

Note that even though $A^\dagger \mathbf{y}$ is not defined for ($\mathbf{y} \in$ III), still

$$AB(t) \to P_{\overline{R(A)}} \quad \text{as } t \to \infty \ .$$

The discrete version of Theorem 9 is the following theorem.

THEOREM 10. (Showalter and Ben–Israel [**761**]). Let $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, let c be a real number, $0 < c < 2$, and let

$$\alpha = \frac{c}{\|A\|^2} \ .$$

For ant $\mathbf{y} \in \mathcal{H}_2$ define

$$\mathbf{x} = T^\dagger \mathbf{y} \quad \text{if } (\mathbf{y} \in \text{I}) \text{ or } (\mathbf{y} \in \text{II})$$

and define the sequence

$$\begin{aligned} \mathbf{y}_0 &= \mathbf{0} \ , \quad \mathbf{x}_0 = \mathbf{0} \ , \\ (\mathbf{y} - \mathbf{y}_{N+1}) &= (I - \alpha AA^*)\,(\mathbf{y} - \mathbf{y}_N) \quad \text{if } (\mathbf{y} \in \text{I}) \text{ or } (\mathbf{y} \in \text{II}) \text{ or } (\mathbf{y} \in \text{III}) \\ (\mathbf{x} - \mathbf{x}_{N+1}) &= (I - \alpha A^* A)\,(\mathbf{x} - \mathbf{x}_N) \quad \text{if } (\mathbf{y} \in \text{I}) \text{ or } (\mathbf{y} \in \text{II}) \\ N &= 1, 2, \dots \end{aligned}$$

Then the sequence

$$B_N = \sum_{k=0}^{N} (I - \alpha A^* A)^k \alpha A^* \ , \quad N = 0, 1, \dots \tag{99}$$

converges to $A^\dagger$ as follows:

(a) $\|(A^\dagger - B_N)\mathbf{y}\|^2 \leq \dfrac{\|A^\dagger \mathbf{y}\|^2 \|(AA^*)^\dagger \mathbf{y}\|^2}{\|(AA^*)^\dagger \mathbf{y}\|^2 + N[(2-c)c/\|A\|^2]\|A^\dagger \mathbf{y}\|^2}$ if ($\mathbf{y} \in$ I) and $N = 1, 2, \dots$

(b) $\|(A^\dagger - B_N)\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{x}_N\|^2$ converges monotonically to zero if ($\mathbf{y} \in$ II).

(c) $\|(P_{\overline{R(A)} - AB_N} \mathbf{y}\|^2 \leq \dfrac{\|\mathbf{y}\|^2 \|A^\dagger \mathbf{y}\|^2}{\|A^\dagger \mathbf{y}\|^2 + N[(2-c)c/\|A\|^2]\|\mathbf{y}\|^2}$ if ($\mathbf{y} \in$ I) or ($\mathbf{y} \in$ II) and $N = 1, 2, \dots$

(d) $\|(P_{\overline{R(A)} - AB_N} \mathbf{y}\|^2 = \|\mathbf{y} - \mathbf{y}_N\|^2$ converges monotonically to zero if ($\mathbf{y} \in$ III). □

The convergence $B_N \to A^\dagger$, in the uniform operator topology, was established by Petryshyn [**643**], restricting A to have closed range.

As in the matrix case, studied in Section 7.7, higher–order iterative methods are more efficient means of summing the series (95) than the first–order method (99). Two such methods, of order $p \geq 2$, are given in the following:

THEOREM 11. (Showalter and Ben–Israel [**761**]). Let A, α and $\{B_N : N = 0, 1, \dots\}$ be as in Theorem 10. Let p be an integer

$$p \geq 2$$

and define the sequence $\{C_{N,p} : N = 0, 1 \dots\}$ and $\{D_{N,p} : N = 0, 1 \dots\}$ as follows:

$$C_{0,p} = \alpha A^* \ , \quad C_{N+1,p} = C_{N,p} \sum_{k=0}^{p-1} (I - AC_{N,p})^k \ , \tag{100}$$

$$D_{0,p} = \alpha A^* \ , \quad D_{N+1,p} = D_{N,p} \sum_{k=0}^{p} \binom{p}{k} (-AD_{N,p})^{k-1} \ . \tag{101}$$

Then, for all $N = 0, 1, \dots$,

$$B_{(p^{N+1}-1)} = C_{N+1,p} = D_{N+1,p} \ . \tag{102}$$

□

Consequently $\{C_{N,p}\}$ and $\{D_{N,p}\}$ are pth–order iterative methods for computing $A^\dagger$, with the convergence rates established in Theorem 10; e.g.,

$$\|(A^\dagger - C_{N,p})\mathbf{y}\|^2 \le \frac{\|A^\dagger \mathbf{y}\|^2 \|(AA^*)^\dagger \mathbf{y}\|^2}{\|(AA^*)^\dagger \mathbf{y}\|^2 + (p^N - 1)[(2-c)c/\|A\|^2]\|A^\dagger \mathbf{y}\|^2} \quad \text{if } (\mathbf{y} \in \mathrm{I}) \text{ and } N = 1, 2, \dots$$

The series (100) is somewhat simpler to use if the term $(I - AC_{N,p})^k$ can be evaluated by only $k-1$ operator multiplications, e.g. for matrices. The form (101) is preferable otherwise, e.g. for integral operators,

6.1. Exercises and examples.

Ex. 50. (Zlobec [**904**]). Let $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ have closed range, let $\mathbf{b} \in \mathcal{H}_2$ and let[2] $B \in R(A^*, A^*)$. Then the sequence

$$\mathbf{x}_{k+1} = \mathbf{x}_k - B(A\mathbf{x}_k - \mathbf{b}) \ , \quad k = 0, 1, \dots \tag{103}$$

converges to $A^\dagger \mathbf{b}$ for all $\mathbf{x}_0 \in R(A^*)$ if

$$\rho(P_{R(A^*)} - BA) < 1$$

where $\rho(T)$ denotes the spectral radius of T; see, e.g. Taylor [**809**, p. 262].

The choice $B = \alpha A^*$ in (103) reduces it to the iterative method (99). Other choices of B are given in the following exercise.

Ex. 51. *Splitting methods.* Let A be as in Ex. 50, and write

$$A = M + N \ , \tag{104}$$

where $M \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ has closed range and $N(A) = N(M)$. Choosing

$$B = w\, M^\dagger \ , \quad w \neq 0$$

in (103) gives

$$\mathbf{x}_{k+1} = [(1-w)I - wM^\dagger N]\mathbf{x}_k + wM^\dagger \mathbf{b} \ , \quad \mathbf{x}_0 \in R(A^*) \ , \tag{105}$$

in particular, for $w = 1$,

$$\mathbf{x}_{k+1} = -M^\dagger N\, \mathbf{x}_k + M^\dagger \mathbf{b} \ , \quad \mathbf{x}_0 \in R(A^*) \ . \tag{106}$$

(Zlobec [**904**], Berman and Neumann [**88**], Berman and Plemmons [**89**]).

7. Frames

This section is based on Christensen's survey [**191**].

Let $\mathcal{H}$ be separable Hilbert space. A *basis* of $\mathcal{H}$ is a set $\{\mathbf{f}_n\} \subset \mathcal{H}$ such that every $\mathbf{f} \in \mathcal{H}$ is represented as

$$\mathbf{f} = \sum_n c_n \mathbf{f}_n \tag{107}$$

with unique scalar coefficients $\{c_n\}$. A basis is *unconditional* if (107) converges unconditionally for all $\mathbf{f} \in \mathcal{H}$. Unconditional bases are generalized as follows.

[2] For $S, T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ with closed ranges, $R(S,T) = \{Z : Z = SWR \text{ for some } W \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)\}$.

DEFINITION 6. (Duffin and Schaeffer [**237**, p. 358]). A sequence $\{\mathbf{f}_n\} \subset \mathcal{H}$ is:
(a) a *frame* for $\mathcal{H}$, if there exist constants $A, B > 0$ such that

$$A\|\mathbf{f}\|^2 \le \sum |\langle \mathbf{f}, \mathbf{f}_n \rangle|^2 \le B\|\mathbf{f}\|^2 \ , \ \forall \mathbf{f} \in \mathcal{H} \tag{108}$$

(b) a *Bessel sequence* if there is a $B > 0$ such that the upper bound holds in (108).

Consider the mapping $T : \ell^2 \to \mathcal{H}$ given by

$$T : \{c_n\} \to \sum_n c_n \mathbf{f}_n \tag{109}$$

$\{\mathbf{f}_n\}$ is a Bessel sequence if and only if T is a well defined operator from ℓ^2 into $\mathcal{H}$, in which case T is bounded, and its adjoint is

$$T^* : \mathcal{H} \to \ell^2 \ , \quad T^*\mathbf{f} = \{\langle \mathbf{f}, \mathbf{f}_n \rangle\} \ , \quad [\mathbf{191}, \text{Lemma } 2.2] \ . \tag{110}$$

If $\{\mathbf{f}_n\}$ is a frame, its *frame operator* $S : \mathcal{H} \to \mathcal{H}$ is $S = TT^*$, or

$$Sf = TT^*\mathbf{f} = \sum \langle \mathbf{f}, \mathbf{f}_n \rangle \mathbf{f}_n \ . \tag{111}$$

S is bounded and surjective, [**405**], allowing the representation

$$\mathbf{f} = SS^{-1}\mathbf{f} = \sum \langle \mathbf{f}, S^{-1}\mathbf{f}_n \rangle \mathbf{f}_n \ , \quad \forall \mathbf{f} \in \mathcal{H} \ , \quad [\mathbf{191}, \text{Theorem } 2.4] \ . \tag{112}$$

The coefficients $\langle \mathbf{f}, S^{-1}\mathbf{f}_n \rangle$ are not unique, however (112) converges unconditionally, showing frames to be generalizations of unconditional bases (uniqueness lost).

The following proof uses generalized inverses.

THEOREM 12. (Christensen [**191**, Theorem 2.5]). A sequence $\{\mathbf{f}_n\} \subset \mathcal{H}$ is a frame if and only if T is a well defined operator from ℓ^2 onto $\mathcal{H}$.

PROOF. *Only if*: Let $\{\mathbf{f}_n\}$ be a frame. Then T is a bounded operator from ℓ^2 into $\mathcal{H}$ (since $\{\mathbf{f}_n\}$ is a Bessel sequence), and T is surjective since S is.

If: Let T be a well defined operator from ℓ^2 onto $\mathcal{H}$. Therefore $\{\mathbf{f}_n\}$ is a Bessel sequence. Let $N(T)^\perp$ be the orthogonal complement of $N(T)$, the kernel of T, and let $\widetilde{T} = T_{[N(T)^\perp]} : N(T)^\perp \to \mathcal{H}$ be the restriction of T to $N(T)^\perp$. $\widetilde{T}$ is clearly bounded and bijective, and therefore has a bounded inverse $T^\dagger = \widetilde{T}^{-1} : \mathcal{H} \to N(T)^\perp$. Writing a decomposition of $T^\dagger\mathbf{f}, \mathbf{f} \in \mathcal{H}$, as $T^\dagger\mathbf{f} = \{(T^\dagger\mathbf{f})_n\}$ we have

$$\mathbf{f} = TT^\dagger\mathbf{f} = \sum (T^\dagger\mathbf{f})_n \mathbf{f}_n \ . \tag{113}$$

$$\begin{aligned} \therefore \ \|\mathbf{f}\|^4 &= \langle \mathbf{f}, \mathbf{f} \rangle^2 = |\langle \sum (T^\dagger\mathbf{f})_n \mathbf{f}_n, \mathbf{f} \rangle|^2 \\ &\le \sum |(T^\dagger\mathbf{f})_n|^2 \sum |\langle \mathbf{f}, \mathbf{f}_n \rangle|^2 \le \|T^\dagger\|^2 \|\mathbf{f}\|^2 \sum |\langle \mathbf{f}, \mathbf{f}_n \rangle|^2 \ . \end{aligned}$$

$$\therefore \ \sum |\langle \mathbf{f}, \mathbf{f}_n \rangle|^2 \ge \frac{\|\mathbf{f}\|^2}{\|T^\dagger\|^2} \ , \quad \forall \mathbf{f} \in \mathcal{H} \ ,$$

establishing the lower bound

$$A = \frac{1}{\|T^\dagger\|^2} = \frac{1}{\|S^{-1}\|} \ , \tag{114}$$

needed in (108) to make $\{\mathbf{f}_n\}$ a frame. ☐

The bound (114) was shown in [**190**] to be optimal. For further details see [**405**], [**190**]–[**191**], and their references.

Suggested further reading

Section 3. The annotated bibliography of Nashed and Rall [**594**] is an essential source.

For alternative or more general treatments of generalized inverses of operators see F. V. Atkinson ([**28**], [**29**]), Beutler ([**91**], [**92**]), Davis and Robinson [**211**], Hamburger [**364**], Hansen and Robinson [**367**], Hestenes [**411**], Holmes [**422**], Leach [**499**], Nashed ([**589**]–[**592**]), Nashed and Votruba ([**595**]–[**596**]), Pietsch [**644**], Porter and Williams ([**649**], [**650**]), Przeworska–Rolewicz and Rolewicz [**655**], Sheffield [**753**], Votruba [**851**], Wyler [**889**] and Zarantonello [**894**].

Sensitivity analysis of the Moore–Penrose inverse: Koliha [**471**], Moore and Nashed ([**573**], [**574**]), Nashed [**591**], Rakočević [**667**], Roch and Silbermann [**708**].

Integral equations: K. E. Atkinson [**30**], Courant and Hilbert [**206**], Kammerer and Nashed ([**457**]–[**458**]), Korganoff and Pavel-Parvu [**474**], Lonseth [**517**], and Rall [**670**].

Reproducing kernel Hilbert spaces: Alpay, Ball and Bolotnikov ([**16**, p. 286], Alpay, Bolotnikov and Loubaton [**17**, p. 84], Alpay, Bolotnikov and Rodman ([**18**, p. 258], [**19**, p. 33]), Aronszajn [**27**], Hilgers [**413**], Moore [**570**], Nashed and Wahba [**597**], Saitoh ([**723**], [**724**], [**725**]).

Generalized inverses of nonlinear mappings: Ben–Israel [**74**], Saitoh [**726**].

Drazin inverses of infinite matrices and linear operators: Campbell ([**144**], [**152**]), Castro González [**168**], Castro González and Koliha ([**169**], [**170**]), Koliha [**470**], Kuang [**480**], Nashed and Zhao [**598**].

Sensitivity analysis of the Drazin inverse: Castro and Koliha [**170**], Castro, Koliha and Straškraba [**171**], Castro, Koliha and Wei [**172**], Hartwig and Shoaf [**387**], Koliha [**472**], Koliha and Rakočević [**473**], Rakočević [**668**], Rakočević and Wei [**669**].

Section 4. For applications to differential operators and related areas see also Bradley ([**132**], [**133**]), Courant and Hilbert [**206**], Greub and Rheinboldt [**316**], Kallina [**450**], Kunkel and Mehrmann [**483**], Lent [**501**], Locker [**515**], Stakgold [**780**], Tucker [**830**], Van Hamme ([**836**], [**837**]), and Wyler [**890**].

Section 5. Groetsch [**325**], Ivanov and Kudrinskii [**434**], Tseng ([**827**], [**828**]), Wahba and Nashed [**852**].

Regularization plays an important role in the approximate solution of operator equations, see Engl, Hanke and Neubauer [**247**], Groetsch [**329**], Groetsch and Hanke [**331**], Groetsch and King [**333**], Groetsch and Neubauer [**334**], Gulliksson, Wedin and Wei [**354**], Hanke [**365**], Hansen [**368**], Hilgers ([**413**], [**414**], [**415**], [**416**]), Meleško [**537**], Morozov ([**577**], [**578**]), Nashed and Wahba [**597**], Neubauer ([**601**], [**602**]), Neumaier [**604**], Tihonov ([**816**], [**815**], [**817**], [**818**]), Tihonov and Arsenin [**819**].

Applications to approximation and spline approximation: Groetsch [**328**], Herring [**407**], Izumino [**435**], Jerome and Schumaker [**440**], Laurent ([**496**], [**497**]), Weinberger [**869**].

Section 6. Groetsch ([**326**], [**327**], [**330**]), Groetsch and Jacobs [**332**], Kammerer and Nashed ([**457**]–[**459**]), Lardy ([**494**], [**495**]), Showalter [**760**], Zlobec [**904**].