Inversion of analytically perturbed linear operators that are singular at the origin

Phil Howlett[†‖] Konstantin Avrachenkov[‡‖] Charles Pearce[§‖]
Vladimir Ejov[¶]

January 10, 2010

Abstract

Let H and K be Hilbert spaces and for each $z \in \mathbb{C}$ let $A(z) \in \mathcal{L}(H, K)$ be a bounded but not necessarily compact linear map with $A(z)$ analytic on a region $|z| < a$. If $A(0)$ is singular we find conditions under which $A(z)^{-1}$ is well defined on some region $0 < |z| < b$ by a convergent Laurent series with a finite order pole at the origin. We show that by changing to a standard Sobolev topology the method extends to closed unbounded linear operators and also that it can be used in Banach spaces where complementation of certain closed subspaces is possible. Our method is illustrated with several key examples[1].

1 Introduction

Let H and K be Hilbert spaces and consider bounded but not necessarily compact linear operators $A_0 \in \mathcal{L}(H, K)$ and $A_1 \in \mathcal{L}(H, K)$. Let $A(z) = A_0 + A_1 z$ be a linear perturbation of A_0 that depends on a single complex parameter $z \in \mathbb{C}$. When A_0 is non-singular the Neumann expansion can be used to calculate $(A_0 + A_1 z)^{-1}$. We refer to Courant and Hilbert [4] p. 18 and pp. 140–142 for further discussion.

Lemma 1.1 (Neumann) *Let $A_0 \in \mathcal{L}(H, K)$ and $A_1 \in \mathcal{L}(H, K)$ and suppose that ${A_0}^{-1}$ is well defined. Let $A(z) = A_0 + A_1 z$ where $z \in \mathbb{C}$. Then for some*

[†]Centre for Industrial and Applied Mathematics, University of South Australia, Mawson Lakes, Australia, (p.howlett@unisa.edu.au).

[‡]INRIA, Sophia Antipoles, France (k.avrachenkov@sophia.inria.fr).

[§]School of Mathematical Sciences, University of Adelaide, Adelaide, Australia, (charles.pearce@adelaide.edu.au).

[¶]Centre for Industrial and Applied Mathematics, University of South Australia, Mawson Lakes, Australia, (v.ejov@unisa.edu.au).

[‖]Supported by the Australian Research Council, Discovery Grant #DP0774504.

[1]This paper is based on preliminary work in [13, 14].

$b > 0$ *we have* $A(z)^{-1}$ *is well defined for* $|z| < b$ *with*

$$A(z)^{-1} = \sum_{j=0}^{\infty} (-1)^j (A_0{}^{-1} A_1)^j A_0{}^{-1} z^j.$$

When A_0 is singular we consider three different situations.

- A_0 is not 1–1.
- A_0 is 1–1 and $A_0(H)$ is closed but A_0 is not onto.
- A_0 is 1–1 but $A_0(H)$ is not closed.

We outline the procedure when A_0 is not 1–1. Let $M = A_0{}^{-1}(\{0\})$ and $N = A_1(M)$. If there is some $z_0 \neq 0$ for which $A(z_0)^{-1}$ is well defined then calculation of $(A_0 + A_1 z)^{-1} \in \mathcal{L}(K, H)$ can be reduced to a term in z^{-1} plus a similar projected calculation of $(A_{0,22} + A_{1,22} z)^{-1} \in \mathcal{L}(N^{\perp}, M^{\perp})$ where $A_{0,22} \in \mathcal{L}(M^{\perp}, N^{\perp})$ and $A_{1,22} \in \mathcal{L}(M^{\perp}, N^{\perp})$. If $A_{0,22}$ is non-singular the Neumann expansion can be applied to the projected problem and the original inverse can be represented on a region $0 < |z| < b$ by a convergent Laurent series with a pole of order 1 at the origin. If $A_{0,22}$ is not 1–1 then the reduction procedure can be applied again. Thus the procedure is essentially recursive. If the procedure terminates after a finite number of steps then the inverse operator $A(z)^{-1}$ is defined on some region $0 < |z| < b$ by a convergent Laurent series with a finite order pole. The other cases are manipulated so that a similar reduction procedure can be used. The method is not restricted to Fredholm operators.

We also consider unbounded operators. When $A_0 : \mathcal{D}(A_0) \subset H \mapsto K$ is a densely defined and closed unbounded linear operator we show that by changing to a standard Sobolev topology on H we can replace A_0 by a bounded operator and apply the previous results. Several key examples will be presented.

We will show that the procedure can be applied when $A(z) \in \mathcal{L}(X, Y)$ where X and Y are Banach spaces provided $X = M \oplus M'$ and $Y = N \oplus N'$ where $M = A_0{}^{-1}(\{0\})$ and $N = A_1(M)$ and M' and N' are complementary spaces. We consider some specific cases and an example of a perturbed Markov process.

We use augmented operators to extend the work on linear perturbations to polynomial perturbations and then to analytic perturbations.

2 Previous work

Much of the work on perturbed operators has been restricted to matrix operators [2, 8, 7, 20, 22], classes of differential operators [16, 22] or Fredholm operators [9] and has often been primarily concerned with analysis of the eigenspaces [15, 17]. The paper [7] by Gohberg *et al* on the local theory of regular analytic matrix functions uses a canonical system of root functions to compute a representation of the Laurent principal part of the inverse function near an isolated singular point. In this analysis the determinant of the matrix function plays a key diagnostic role. Although the earlier, beautifully written, paper by Vishik and

Lyusternik [22] is more general in scope the inversion formulae are developed for singularities on finite dimensional subspaces. The book [8] by Gohberg *et al* presents a systematic treatment of perturbation theory for Fredholm operators but once again relies on finite dimensional techniques. To extend the theory to more general classes some of the familiar algebraic techniques must be discarded or revised. In this paper we consider bounded but not necessarily compact linear operators and pay particular attention to cases where the null space is non trivial for the unperturbed operator but becomes trivial under perturbation. Our methodology follows early papers by Sain and Massey [19] and Howlett [12] on input retrieval in finite dimensional linear control systems, the PhD thesis by Avrachenkov [1] on analytic perturbations and their application and subsequent work by Howlett and Avrachenkov [13] and Howlett *et al* [14] on basic theoretical aspects of operator perturbation. Our approach was inspired by the work of Schweitzer and Stewart [20] on a corresponding matrix inversion problem but our technique depends on a geometric separation of the underlying spaces. The separation mimics the algebraic separation employed by Howlett [12] for matrix operators but does not depend directly on other established perturbation techniques. For this reason we defer to the PhD thesis by Avrachenkov [1], the paper by Avrachenkov *et al* [2], the book by Gohberg [9] and the fundamental work by Kato [15] for a more comprehensive review of the literature. Our work relies heavily on standard functional analysis for which we cite the classic texts by Courant and Hilbert [4], Diestel [5], Dunford and Schwartz [6], Hewitt and Stromberg [11], Luenberger [18], Singer [21] and Yosida [23].

3 Bounded operators: the basic inversion procedure

We use two key results established by Howlett and Avrachenkov [13]. These results are not widely available and so we repeat them here. Assume A_0 is not 1–1.

3.1 The key lemma

The following lemma establishes the basis for the inversion procedure.

Lemma 3.1 (Howlett & Avrachenkov) *Let H and K be Hilbert spaces and let $A_0, A_1 \in \mathcal{L}(H,K)$ be bounded linear maps. For each $z \in \mathbb{C}$ define $A(z) \in \mathcal{L}(H,K)$ by $A(z) = A_0 + A_1 z$. Suppose $M = {A_0}^{-1}(\{0\}) \neq \{0\}$ and let $N = A_1(M) \subset K$. If $A(z_0)^{-1}$ is well defined for some $z_0 \neq 0$ then A_1 is bounded below on M and N is a closed subspace of K.*

Proof. By the Banach Inverse Theorem (see Luenberger [18] p. 149) the map $(A_0 + A_1 z_0)$ is bounded below on H. Therefore we can find $\epsilon > 0$ such that

$$\|(A_0 + A_1 z_0)x\| \geq \epsilon \|x\|$$

for all $x \in H$. Since $A_0 m = 0$ it follows that

$$\|A_1 m\| \geq \frac{\epsilon}{|z_0|}\|m\|$$

for all $m \in M$. If $\{n_r\}$ is a Cauchy sequence in $N = A_1(M)$ then $n_r = A_1 m_r$ where $\{m_r\}$ is a corresponding sequence in M. Because A_1 is bounded below on M the sequence $\{m_r\}$ must also be a Cauchy sequence. If $m_r \to m$ and $n_r \to n$ then $A_1 m = n$. Thus $n \in A_1(M) = N$. $\square$

3.2 The key orthogonal decomposition

Since $M = {A_0}^{-1}(\{0\})$ is closed and since the orthogonal complement $M^\perp$ is also closed it follows that $H_1 = M$ and $H_2 = M^\perp$ are each Hilbert spaces. Let $P \in \mathcal{L}(H, H)$ denote the natural projection onto the subspace $M \subset H$ and define associated mappings $P_i \in \mathcal{L}(H, H_i)$ for $i = 1, 2$ by setting $P_1 = P$ and $P_2 = I - P$. Define $R \in \mathcal{L}(H, H_1 \times H_2)$ by the formula

$$Rx = \begin{pmatrix} P_1 x \\ P_2 x \end{pmatrix}$$

for each $x \in H$. Since $\langle Rx_1, Rx_2\rangle = \langle x_1, x_2\rangle$ for each $x_1, x_2 \in H$ the mapping R defines a unitary equivalence between H and $H_1 \times H_2$. In the same way note that $N = A_1(M)$ is closed and since $N^\perp$ is also closed it follows that $K_1 = N$ and $K_2 = N^\perp$ are each Hilbert spaces. Let $Q \in \mathcal{L}(K, K)$ denote the natural projection onto the subspace $N \subset K$ and define associated mappings $Q_j \in \mathcal{L}(K, K_j)$ for $j = 1, 2$ by setting $Q_1 = Q$ and $Q_2 = I - Q$. Define $S \in \mathcal{L}(K, K_1 \times K_2)$ by the formula

$$Sy = \begin{pmatrix} Q_1 y \\ Q_2 y \end{pmatrix}$$

for each $y \in K$. The mapping S defines a unitary equivalence between K and $K_1 \times K_2$. Now partition the operators A_0 and A_1 in the form

$$SA_0R^* = \begin{pmatrix} 0 & A_{0,12} \\ 0 & A_{0,22} \end{pmatrix} \quad \text{and} \quad SA_1R^* = \begin{pmatrix} A_{1,11} & A_{1,12} \\ 0 & A_{1,22} \end{pmatrix}$$

where $A_{0,ij}, A_{1,ij} \in \mathcal{L}(H_i, K_j)$ and where we note that $A_{0,11} = Q_1 A_0 P_1^* = 0$, $A_{0,12} = Q_1 A_0 P_2^*$, $A_{0,21} = Q_2 A_0 P_1^* = 0$, $A_{0,22} = Q_2 A_0 P_2^*$, $A_{1,11} = Q_1 A_1 P_1^*$, $A_{1,12} = Q_1 A_1 P_2^*$, $A_{1,21} = Q_2 A_1 P_1^* = 0$ and $A_{1,22} = Q_2 A_1 P_2^*$.

Remark 3.2 *Recall that if A_0 is not* 1–1 *and $(A_0 + A_1 z_0)^{-1}$ exists for some $z_0 \in \mathbb{C}$ with $z_0 \neq 0$ then A_1 is bounded below on H_1. Equivalently we can say that $A_{1,11} \in \mathcal{L}(H_1, K_1)$ is bounded below. It follows that $A_{1,11}$ is a* 1–1 *mapping of H_1 onto K_1.*

3.3 The key inversion formula

We use the notation introduced in the previous subsection.

Theorem 3.3 (Howlett & Avrachenkov) *Let $A_0 \in \mathcal{L}(H,K)$ with $H_1 = {A_0}^{-1}(\{0\}) \neq \{0\}$. Suppose $A_{1,11} \in \mathcal{L}(H_1,K_1)$ is a 1–1 mapping of H_1 onto $K_1 = A_1(H_1)$. The mapping $A(z) \in \mathcal{L}(H,K)$ is a 1–1 mapping of H onto K if and only if $z \neq 0$ and $(A_{0,22} + A_{1,22}z) \in \mathcal{L}(H_2,K_2)$ is a 1–1 mapping of $H_2 = H_1^{\perp}$ onto $K_2 = K_1^{\perp}$. In this case*

$$\begin{aligned} A(z)^{-1} &= P_1^* S A_{1,11}^{-1} Q_1/z \\ &\quad + \left[P_2^* - P_1^* A_{1,11}^{-1}(A_{0,12} + A_{1,12}z)/z\right](A_{0,22} + A_{1,22}z)^{-1} Q_2. \end{aligned} \tag{3.1}$$

Proof. Since

$$A(z) = S^* \begin{pmatrix} A_{1,11}z & A_{0,12} + A_{1,12}z \\ 0 & A_{0,22} + A_{1,22}z \end{pmatrix} R$$

where R and S are unitary operators it follows that $A(z)^{-1}$ exists if and only if

$$\begin{pmatrix} A_{1,11}z & A_{0,12} + A_{1,12}z \\ 0 & A_{0,22} + A_{1,22}z \end{pmatrix}^{-1}$$

exists. Let $x = R\xi$ and $y = S\eta$. The system of equations $A(z)x = y$ has a unique solution $x \in H$ for each $y \in K$ if and only if the system of equations

$$\begin{aligned} (A_{1,11}z)\xi_1 + (A_{0,12} + A_{1,12}z)\xi_2 &= \eta_1 \\ (A_{0,22} + A_{1,22}z)\xi_2 &= \eta_2 \end{aligned}$$

has a unique solution $\xi \in H_1 \times H_2$ for each $\eta \in K_1 \times K_2$. The latter system can be rewritten as

$$\begin{aligned} (A_{0,22} + A_{1,22}z)\xi_2 &= \eta_2 \\ (A_{1,11}z)\xi_1 &= \eta_1 - (A_{0,12} + A_{1,12}z)\xi_2 \end{aligned}$$

and so there is a unique solution if and only if $z \neq 0$ and both $A_{1,11}$ is a *1–1* mapping of H_1 onto K_1 and $(A_{0,22} + A_{1,22}z)$ is a 1–1 mapping of H_2 onto K_2. Therefore

$$\begin{aligned} \xi_2 &= (A_{0,22} + A_{1,22}z)^{-1}\eta_2 \\ \xi_1 &= A_{1,11}^{-1}\left[\eta_1 - (A_{0,12} + A_{1,12}z)\xi_2\right]/z \end{aligned}$$

and hence, by back substitution, $x = P_1^*\xi_1 + P_2^*\xi_2$ gives

$$x = \{P_1^* A_{1,11}^{-1} Q_1/z + \left[P_2^* - P_1^* A_{1,11}^{-1}(A_{0,12} + A_{1,12}z)/z\right](A_{0,22} + A_{1,22}z)^{-1} Q_2\}y.$$

Thus we obtain the desired formula for $A(z)^{-1}$. $\square$

Remark 3.4 *If $A_{0,22} \in \mathcal{L}(H_2, K_2)$ is a 1–1 mapping of H_2 onto K_2 then ${A_{0,22}}^{-1}$ is well defined and for some real number $b > 0$ the operator $(A_{0,22} + A_{1,22}z) \in \mathcal{L}(H_2, K_2)$ is defined by a convergent Neumann series in the region $|z| < b$. Thus the operator $A(z)^{-1}$ is defined in the region $0 < |z| < b$ by a convergent Laurent series with a pole of order 1 at $z = 0$.*

Example 1 (Continuous spectrum) *Each element in the space $L^2(\mathbb{R})$ can be represented by a Fourier integral and defined by a continuously distributed spectral density. A bounded linear operator on $L^2(\mathbb{R})$ can be regarded as a linear transformation on a continuous spectrum. Let*

$$w(t) = \frac{2\sin(u_0 t)}{t}$$

where $u_0 \in \mathbb{R}$ and $u_0 > 0$. Define $A_0 : L^2(\mathbb{R}) \mapsto L^2(\mathbb{R})$ by the formula

$$A_0 x(t) = x(t) - [x * w](t) = x(t) - \frac{1}{\pi}\int_{\mathbb{R}} x(\tau) w(t-\tau) d\tau$$

for all $t \in \mathbb{R}$. The Fourier cosine and sine transforms are defined by

$$\mathcal{F}_c[p](u) = \frac{1}{\pi}\int_{\mathbb{R}} p(t)\cos(ut)dt \quad \text{and} \quad \mathcal{F}_s[p](u) = \frac{1}{\pi}\int_{\mathbb{R}} p(t)\sin(ut)dt$$

for each $p \in L^2(\mathbb{R})$. It is well known that p can be reconstructed by the formula

$$p(t) = \int_{\mathbb{R}} \left[\mathcal{F}_c[p](u)\cos(ut) + \mathcal{F}_s[p](u)\sin(ut)\right] dt$$

and that the correspondence $p \in L^2(\mathbb{R}) \Leftrightarrow (\mathcal{F}_c[p], \mathcal{F}_s[p]) \in L^2(\mathbb{R}) \times L^2(\mathbb{R})$ is unique. If $p, q \in L^2(\mathbb{R})$ then

$$\mathcal{F}_c[p*q](u) = \mathcal{F}_c[p](u)\mathcal{F}_c[q](u) - \mathcal{F}_s[p](u)\mathcal{F}_s[q](u) \quad \text{and}$$
$$\mathcal{F}_s[p*q](u) = \mathcal{F}_c[p](u)\mathcal{F}_s[q](u) + \mathcal{F}_s[p](u)\mathcal{F}_c[q](u).$$

Since $\mathcal{F}_c[w](u) = \chi_{(-u_0,u_0)}(u)$ and $\mathcal{F}_s[w](u) = 0$ it follows that

$$\mathcal{F}_c[A_0 x](u) = \mathcal{F}_c[x](u) - \mathcal{F}_c[x*w](u) = \mathcal{F}_c[x](u)\left[1 - \chi_{(-u_0,u_0)}(u)\right] \quad \text{and}$$
$$\mathcal{F}_s[A_0 x](u) = \mathcal{F}_s[x](u) - \mathcal{F}_s[x*w](u) = \mathcal{F}_s[x](u)\left[1 - \chi_{(-u_0,u_0)}(u)\right]$$

for each $x \in L^2(\mathbb{R})$. Define $A_1 : L^2(\mathbb{R}) \mapsto L^2(\mathbb{R})$ by $A_1 x = x$ for all $x \in L^2(\mathbb{R})$ and consider the equation $(A_0 + A_1 z)x = y$. The solution is given by $x = (A_0 + A_1 z)^{-1} y$ provided the inverse exists. Taking a Fourier cosine transform of the original equation gives

$$\mathcal{F}_c[x](u)\left[(1+z) - \chi_{(-u_0,u_0)}(u)\right] = \mathcal{F}_c[y](u)$$

and hence

$$\begin{aligned}
\mathcal{F}_c[x](u) &= \mathcal{F}_c[y](u)\chi_{(-u_0,u_0)}(u) \cdot \frac{1}{z} + \mathcal{F}_c[y](u)\left[1 - \chi_{(-u_0,u_0)}(u)\right] \cdot \frac{1}{1+z} \\
&= \mathcal{F}_c[y*w](u) \cdot \frac{1}{z} + \left[\mathcal{F}_c[y](u) - \mathcal{F}_c[y*w](u)\right] \cdot \left[1 - z + z^2 - \cdots\right]
\end{aligned}$$

for $|z| < 1$. *In similar fashion a Fourier sine transform of the original equation gives*

$$\mathcal{F}_s[x](u)\left[(1+z) - \chi_{(-u_0,u_0)}(u)\right] = \mathcal{F}_s[y](u)$$

from which it follows that

$$\mathcal{F}_s[x](u) \quad = \quad \mathcal{F}_s[y*w](u)\cdot\frac{1}{z} + [\mathcal{F}_s[y](u) - \mathcal{F}_s[y*w](u)]\cdot[1 - z + z^2 - \cdots]$$

for $|z| < 1$. *Therefore the solution is*

$$x(t) = (y*w)(t)\cdot\frac{1}{z} + [y(t) - (y*w)(t)]\cdot[1 - z + z^2 - \cdots]$$

for $|z| < 1$. *Note that the Laurent series has a pole of order* 1 *provided* $(y*w) \neq 0$. *By considering the Fourier transforms it can be seen that* $(y*w) = 0$ *if and only if* $\mathcal{F}_c[y](u) = 0$ *and* $\mathcal{F}_s[y](u) = 0$ *for almost all* $u \in (-u_0, u_0)$.

Remark 3.5 *If* $A(z_0) \in \mathcal{L}(H,K)$ *is non-singular then* $(A_{0,22} + A_{1,22}z_0) \in \mathcal{L}(H_2, K_2)$ *is also non-singular. If* $A_{0,22} \in \mathcal{L}(H_2,K_2)$ *is onto but not* $1-1$ *then Theorem* 3.3 *can be applied to the operator* $(A_{0,22} + A_{1,22}z)$. *Thus the procedure is essentially recursive.*

Example 2 *Let* $u : [-\pi,\pi] \mapsto \mathbb{R}$ *be defined by* $u(t) = (-1)\cdot\mathrm{sgn}(t)$ *for all* $t \in [-\pi,\pi]$. *That is*

$$u(t) = \begin{cases} 1 & \text{for } t \in (-\pi, 0) \\ 0 & \text{for } t = -\pi, 0, \pi \\ -1 & \text{for } t \in (0,\pi). \end{cases}$$

Let $H = K = L^2([-\pi,\pi])$. *Define* $A_0 : H \mapsto K$ *by setting*

$$\begin{aligned} A_0x(t) &= \frac{1}{16}\left[(x*u)(t) + (x*u)(-t)\right] \\ &= \frac{1}{16}\int_{-\pi}^{\pi} x(s)\left[u(t-s) + u(-t-s)\right]ds \\ &= -\frac{1}{8}\left[X(t) + X(-t)\right] + \frac{1}{8}\left[X(t-\pi) + X(-t+\pi)\right] \end{aligned}$$

where $X(t) = \int_{[0,t]} x(s)ds$ *and where we have used the periodic extension of* $u(t)$ *as required in the convolution integral. The functions*

$$\boldsymbol{e}_0 = 1,\ \boldsymbol{e}_1(t) = \cos t,\ \boldsymbol{f}_1(t) = \sin t,\ \boldsymbol{e}_2(t) = \cos 2t,\ \boldsymbol{f}_2(t) = \sin 2t,\ \ldots$$

form an orthogonal basis for L *and hence we can represent each element* $f \in L$ *as an infinite sequence*

$$f = (a_0, a_1, b_1, a_2, b_2, \ldots) \Leftrightarrow a_0 + \sum_{n=1}^{\infty}[a_n\boldsymbol{e}_n + b_n\boldsymbol{f}_n]$$

of Fourier coefficients. Note that $A_0\boldsymbol{e}_n = 0$*,* $A_0\boldsymbol{f}_{2m} = 0$ *and* $A_0\boldsymbol{f}_{2m-1} = \boldsymbol{e}_{2m-1}/(2m-1)$ *for all* $m, n \in \mathbb{N}$*. The null space* $M = A_0^{-1}(\{0\})$ *is defined by*

$$M = \{x \mid x \in H \text{ and } x = (a_0, a_1, 0, a_2, b_2, a_3, 0, a_4, b_4, \ldots)\}$$

and the orthogonal complement $M^\perp = A_0^{-1}(\{0\})^\perp$ *is defined by*

$$M^\perp = \{x \mid x \in H \text{ and } x = (0, 0, b_1, 0, 0, 0, b_3, 0, 0, \ldots)\}.$$

Both M *and* $M^\perp$ *are infinite dimensional spaces. In terms of the Fourier coefficients the mapping* $A_0 \in \mathcal{L}(H, K)$ *can be described by the relationship*

$$A_0(a_0, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, \ldots) = (0, b_1, 0, 0, 0, b_3/3, 0, 0, 0, \ldots).$$

Let $A_1 = I$*. The perturbed operator* $(A_0 + A_1 z) : H \mapsto K$ *can be defined by an equivalent transformation* $(A_0 + A_1 z) : \ell^2 \mapsto \ell^2$ *using the formula*

$$\begin{aligned}&(A_0 + A_1 z)(a_0, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5, \ldots) \\ &\quad = \quad (a_0 z, b_1 + a_1 z, b_1 z, a_2 z, b_2 z, b_3/3 + a_3 z, b_3 z, a_4 z, b_4 z, \ldots)\end{aligned}$$

where a_0*,* a_n *and* b_n *are the usual Fourier coefficients. Solving a simple set of equations shows that the equivalent inverse transformation* $(A_0 + A_1 z)^{-1} : \ell^2 \mapsto \ell^2$ *is defined by*

$$\begin{aligned}&(A_0 + A_1 z)^{-1}(c_0, c_1, d_1, c_2, d_2, c_3, d_3, c_4, d_4, c_5, d_5, \ldots) \\ &\quad = \quad \left(\frac{c_0}{z}, \frac{c_1}{z} - \frac{d_1}{z^2}, \frac{d_1}{z}, \frac{c_2}{z}, \frac{d_2}{z}, \frac{c_3}{z} - \frac{d_3}{3z^2}, \frac{d_3}{z}, \frac{c_4}{z}, \frac{d_4}{z}, \ldots\right)\end{aligned}$$

where c_0*,* c_n *and* b_n *are the usual Fourier coefficients. The inverse operator has a pole of order* 2 *at the origin. Write* $H = M \times M^\perp$ *and* $K = N \times N^\perp$ *where* $N = A_1(M) = M$ *and* $N^\perp = M^\perp$*. Now using an infinite dimensional matrix notation*

$$(A_0 + A_1 z) = \left[\begin{array}{cccccc|ccc} z & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & z & 0 & 0 & 0 & \cdots & 1 & 0 & \cdots \\ 0 & 0 & z & 0 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & 0 & 0 & z & 0 & \cdots & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & z & \cdots & 0 & \frac{1}{3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \hline 0 & 0 & 0 & 0 & 0 & \cdots & z & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & z & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] = \left[\begin{array}{c|c} Iz & A_{0,12} \\ \hline 0 & Iz \end{array}\right]$$

and hence

$$(A_0 + A_1 z)^{-1} = \left[\begin{array}{c|c} I \cdot \dfrac{1}{z} & -A_{0,12} \cdot \dfrac{1}{z^2} \\ \hline 0 & I \cdot \dfrac{1}{z} \end{array}\right].$$

Remark 3.6 *If the procedure is applied recursively to generate a sequence*

$$M_1^\perp \supset M_2^\perp \supset \cdots$$

of complementary spaces and if $M_n^\perp$ is finite dimensional for some $n \in \mathbb{N}$ then the recursive procedure terminates after a finite number of steps and the Laurent series has a finite order pole and converges on some region $0 < |z| < b$.

Remark 3.7 *If the action of the operators is restricted to a finite dimensional subspace for the purpose of numerical calculation then the Laurent series for the inverse of the perturbed restricted operator has at most a finite order pole.*

The recursive procedure may continue indefinitely as the following example shows.

Example 3 *Consider the mappings on ℓ^2 defined by the infinite matrices*

$$A_0 = \begin{bmatrix} 0 & A_{0,12} \\ 0 & A_{0,22} \end{bmatrix} = \left[\begin{array}{c|cccc} 0 & 1 & 0 & 0 & \cdots \\ \hline 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right]$$

and

$$A_1 = \begin{bmatrix} A_{1,11} & A_{1,12} \\ 0 & A_{1,22} \end{bmatrix} = \left[\begin{array}{c|cccc} 1 & 0 & 0 & 0 & \cdots \\ \hline 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] = I$$

and the linearly perturbed infinite matrix

$$A(z) = \begin{bmatrix} A_{1,11}z & A_{0,12} + A_{1,12}z \\ 0 & A_{0,22} + A_{1,22}z \end{bmatrix} = \left[\begin{array}{c|cccc} z & 1 & 0 & 0 & \cdots \\ \hline 0 & z & 1 & 0 & \cdots \\ 0 & 0 & z & 1 & \cdots \\ 0 & 0 & 0 & z & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] = (A_0 + Iz).$$

The reduced problem to calculate $(A_{0,22} + A_{1,22}z)^{-1}$ is the same as the original problem to calculate $A(z)^{-1}$. By an elementary calculation

$$\begin{aligned} (A_0 + Iz)^{-1} &= \begin{bmatrix} z^{-1} & -z^{-2} & z^{-3} & -z^{-4} & \cdots \\ 0 & z^{-1} & -z^{-2} & z^{-3} & \cdots \\ 0 & 0 & z^{-1} & -z^{-2} & \cdots \\ 0 & 0 & 0 & z^{-1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \\ &= I \cdot \frac{1}{z} + (-1)A_0 \cdot \frac{1}{z^2} + (-1)^2 A_0^2 \cdot \frac{1}{z^3} + \cdots . \end{aligned}$$

In general this series does not converge near $z = 0$ but if $y = \sum_{j=1}^{n} y_j \boldsymbol{e}_j$ for some natural number $n \in \mathbb{N}$ then only the first n terms are non-zero and the series converges for all $z \neq 0$ with a pole of order at most n at the origin.

4 Bounded operators: the adjoint inversion formula

These results were originally proposed in [14]. Assume A_0 is 1–1 and $A_0(H)$ is closed but A_0 is not onto.

4.1 The adjoint operator

Let $A_0 \in \mathcal{L}(H, K)$. The Hilbert space adjoint $A_0{}^* \in \mathcal{L}(K, H)$ is defined by the relationship $\langle x, A_0{}^* y \rangle = \langle A_0 x, y \rangle$ for all $x \in H$ and $y \in K$. The following standard result is used.

Lemma 4.1 *Let $A_0 \in \mathcal{L}(H, K)$ and let $A_0{}^* \in \mathcal{L}(K, H)$ denote the Hilbert space adjoint. If $A_0{}^{-1}(\{0\}) = \{0\}$ and $A_0(H)$ is closed but $A_0(H) \neq K$ then $[A_0{}^*]^{-1}(\{0\}) \neq \{0\}$ and $A_0{}^*(K) = H$. Thus the adjoint operator $A_0{}^*$ is onto but not 1–1.*

Remark 4.2 *If $A^{-1} \in \mathcal{L}(K, H)$ is well defined then $[A^*]^{-1} = [A^{-1}]^* \in \mathcal{L}(H, K)$ is also well defined.*

Lemma 4.1 and Remark 4.2 provide a basis for the inversion procedure when $A_0{}^{-1}(\{0\}) = \{0\}$ and $A_0(H)$ is closed but $A_0(H) \neq K$.

Theorem 4.3 *Let $A_0 \in \mathcal{L}(H, K)$ and suppose $A_0{}^{-1}(\{0\}) = \{0\}$ with $A_0(H)$ closed but $A_0(H) \neq K$. If the inverse operator $A(z_0)^{-1} = (A_0 + A_1 z_0)^{-1}$ is well defined for some $z_0 \neq 0$ then $[A(z_0)^*]^{-1} = (A_0{}^* + A_1{}^* \overline{z_0})^{-1} = [A(z_0)^{-1}]^*$ is also well defined. If Theorem 3.3 can be applied to show that for some $b > 0$ the inverse operator $[A(z)^*]^{-1}$ is well defined for $0 < |z| < b$ then $A(z)^{-1} = [\{A(z)^*\}^{-1}]^*$ is also well defined for $0 < |z| < b$.*

Proof. Apply the original inversion formula to the adjoint operator $A(z)^*$ and recover the desired series from the formula $A(z)^{-1} = [\{A(z)^*\}^{-1}]^*$. $\square$

5 Bounded operators: non-closed range

We begin with an important and well-known observation.

Lemma 5.1 *If $A_0 \in \mathcal{L}(H, K)$ and $A_0(H)$ is not closed then A_0 is not bounded below.*

Outline of Proof. If $y \in \overline{A_0(H)} \setminus A_0(H)$ then we can find $\{x_n\} \in H$ such that $\|A_0 x_n - y\|_K \to 0$ as $n \to \infty$. If $\|x_n\|_H$ is bounded then by the Eberlein–Shmulyan theorem (see Yosida [23] pp. 141-145) there is a subsequence $\{x_{n(m)}\}$ and some $x \in H$ such that $x_{n(m)}$ converges weakly to x. It follows that $A_0 x_{n(m)}$ converges weakly to $A_0 x$ and hence that $A_0 x = y \in A_0(H)$. This is a contradiction. Thus $\{x_n\}$ is unbounded and since $\{A_0 x_n\}$ is bounded it follows that A_0 is not bounded below. □

When $A_0(H)$ is not closed in K the essence of the difficulty is that K is an inappropriate image space with an inappropriate topology. We restrict the image space and define a new topology.

Definition 5.2 *Let $M = A_0(\{0\})^{-1}$ be the null space of A_0. Let $\langle \cdot, \cdot \rangle_E : A_0(H) \times A_0(H) \mapsto \mathbb{C}$ be defined by the formula*

$$\langle y, v \rangle_E = \langle y, v \rangle_K + \langle x_M^\perp, u_M^\perp \rangle_H$$

for each $y, v \in A_0(H)$ where $x_M^\perp, u_M^\perp \in M^\perp$ are the uniquely defined elements with $A_0 x_M^\perp = y$ and $A_0 u_M^\perp = v$.

Lemma 5.3 *The space $K_E = \{A_0(H), \langle \cdot, \cdot \rangle_E\}$ is a Hilbert space.*

The mapping $A_{0,E} \in \mathcal{L}(H, K_E)$ defined by $A_{0,E} x = A_0 x$ for all $x \in H$ is onto but not necessarily 1–1. Of course it may well be true that K_E can be regarded as a closed subspace of some larger Hilbert space K' in which case the mapping $A_{0,E} \in \mathcal{L}(H, K')$ is no longer onto. In any case the original inversion formulae can now be applied to the perturbed operator $(A_{0,E} + A_{1,E} z) \in \mathcal{L}(H, K')$ where we assume the perturbation operator $A_{1,E} \in \mathcal{L}(H, K')$.

Example 4 (A modified integral operator) *Let $H = K = L^2([0,1])$. Note that the space $L^2([0,1])$ can be generated by the limits of all Cauchy sequences of continuous functions $\{x_n\} \in \mathcal{C}_0([0,1])$ in $L^2([0,1])$ satisfying $x_n(0) = x_n(1) = 0$. Define $A_0 \in \mathcal{L}(H, K)$ by setting $A_0 x(t) = \mathcal{X}(1) - X(t)$ where*

$$X(t) = \int_0^t x(s) ds \quad \textit{and} \quad \mathcal{X}(u) = \int_0^u X(t) dt.$$

If we define $x_n \in H$ by

$$x_n(s) = \sin n\pi s$$

then $\|x_n\| = 1/\sqrt{2}$ for all $n \in \mathbb{N}$ but we have

$$A_0 x_n(t) = \frac{\cos n\pi t}{n\pi}$$

and hence $\|A_0 x_n\| \to 0$ as $n \to \infty$. Therefore A_0 is not bounded below and $A_0(H)$ is not closed in K. For instance if we define $y_0 \in K$ by the formula

$$\begin{aligned} y_0(t) &= \begin{cases} \frac{1}{2} & \textit{for} \quad 0 < t < \frac{1}{2} \\ -\frac{1}{2} & \textit{for} \quad \frac{1}{2} < t < 1 \end{cases} \\ &= \frac{2}{\pi} \left[\cos \pi t - \frac{\cos 3\pi t}{3} + \frac{\cos 5\pi t}{5} - \cdots \right] \end{aligned}$$

then $y_0 \in \overline{A_0(H)} \backslash A_0(H)$. In general there are many non-differentiable functions on the boundary of the set $A_0(H)$. Define a new scalar product on $A_0(H)$ by setting

$$\langle y, v\rangle_E = \langle y, v\rangle_K + \langle x, u\rangle_H$$

where x, u are the unique solutions to $y = A_0 x$ and $v = A_0 u$. If we use the new topology induced by the new scalar product then non-differentiable functions, such as y_0, are removed from the image space. The image space now consists of those functions $y \in L^2([0,1])$ with generalized derivative $y' \in L^2([0,1])$ such that $\int_0^1 y(t)dt = 0$, and with $\|y\|_E^2 = \|y\|_2^2 + \|y'\|_2^2$.

6 Unbounded operators that are both densely defined and closed

We refer the reader to the book by Yosida [23] pp. 193-201 for general results about densely defined closed linear operators and their adjoints. Suppose $A_0 : \mathcal{D}(A_0) \subset H \mapsto K$ is a densely defined and closed unbounded linear operator. For each $\varphi, \psi \in \mathcal{D}(A_0)$ define a new inner product

$$\langle \varphi, \psi\rangle_E = \langle \varphi, \psi\rangle_H + \langle A_0\varphi, A_0\psi\rangle_K$$

and corresponding norm $\|\varphi\|_E = [\langle \varphi, \varphi\rangle_E]^{1/2}$. The space $H_E = (\mathcal{D}(A_0), \langle \cdot, \cdot\rangle_E)$ is a Hilbert space. We denote the new mapping by $A_{0,E} : H_E \mapsto K$. In practice the operator A_0 may be defined on a dense subset $\mathcal{C} \subset H$ but may not be closed. In such cases the set $H_E \subset H$ is defined as the completion of $\mathcal{C}$ in the new norm. The point $x \in H$ will belong to H_E if there exists a sequence $\{\varphi_n\} \in \mathcal{C}$ with $\|\varphi_n - x\|_E \to 0$ as $n \to \infty$. Thus we must also have $y \in K$ with $\|A_0\varphi_n - y\|_K \to 0$. The completion is guaranteed if we allow the limit process to define an appropriate equivalence class.

Lemma 6.1 *The mapping $A_{0,E} : H_E \mapsto K$ is a bounded linear mapping. That is $A_{0,E} \in \mathcal{L}(H_E, K)$.*

Remark 6.2 *$A_0(H)$ is closed if and only if A_0 is bounded below on $\mathcal{D}(A_0)$.*

Theorem 6.3 (J. von Neumann) *If A is both densely defined and closed then A^*A and AA^* are self-adjoint with $(I + A^*A)^{-1} \in \mathcal{L}(H)$ and $(I + AA^*)^{-1} \in \mathcal{L}(K)$.*

Lemma 6.4 *The new adjoint mapping ${A_{0,E}}^* \in \mathcal{L}(K, H_E)$ is defined in terms of the original adjoint mapping ${A_0}^* : \mathcal{D}({A_0}^*) \subset K \mapsto H$ by the formulae*

$${A_{0,E}}^* = {A_0}^*(I + A_0{A_0}^*)^{-1} = (I + {A_0}^*A_0)^{-1}{A_0}^*.$$

Since the operator $A_{0,E} : H_E \mapsto K$ is a bounded linear mapping the original inversion formula can now be applied to a linearly perturbed operator $(A_{0,E} + A_{1,E}z) \in \mathcal{L}(H_E, K)$ where we assume the perturbation operator $A_{1,E} \in \mathcal{L}(H_E, K)$.

Example 5 (The differentiation operator) *Let $H = L^2([0,1])$ and define $A_0\varphi(t) = \varphi'(t)$ for all $\varphi \in \mathcal{C}_0^1([0,1])$ and all $t \in [0,1]$. For each $\{\varphi_n\} \in \mathcal{C}_0^1([0,1])$ with*

$$\int_{[0,1]} \left\{|\varphi_m(t) - \varphi_n(t)|^2 + |\varphi_m{}'(t) - \varphi_n{}'(t)|^2\right\} dt \to 0$$

as $m, n \to \infty$ there exist functions x and y such that

$$\int_{[0,1]} |\varphi_n(t) - x(t)|^2 dt \to 0 \quad \text{and} \quad \int_{[0,1]} |\varphi_n{}'(t) - y(t)|^2 dt \to 0$$

as $n \to \infty$. We say $y = x'$ is the generalized derivative of x. Note that

$$\|x\|^2 = \int_0^1 \left|\int_0^t x'(s)ds\right|^2 dt \le \frac{1}{2}\| x'\|^2.$$

The Hilbert space H_E is the completion of the space $\mathcal{C}_0^1([0,1])$ with the inner product

$$\langle x, u\rangle_E = \int_0^1 [x(t)\overline{u}(t) + x\,'(t)\overline{u}\,'(t)]dt$$

and the norm

$$\|x\|_E = \left[\int_0^1 \{|x(t)|^2 + |x\,'(t)|^2\}dt\right]^{1/2}.$$

It can be shown that

$$H_E = \{x \mid \ x \in \mathcal{C}_0^0([0,1]) \text{ and } x\,' \in L^2([0,1])\}.$$

The space $H_E = H_0^1([0,1])$ is an elementary example of a Sobolev space. Define the generalized differentiation operator $A_{0,E} : H_E \mapsto K$ by the formula $A_{0,E}x = \lim_{n\to\infty} A_0\varphi_n$ where $\varphi_n \in \mathcal{C}_0^1([0,1])$ and $\varphi_n \to x$ in H_E as $n \to \infty$. Thus $A_{0,E}x = x'$ is simply the generalized derivative. It follows from the inequality above that $A_{0,E}$ is bounded below and hence $A_{0,E}(H_E)$ is closed. It is also obvious that $\|A_{0,E}x\| \le \|x\|_E$ and so $A_{0,E} \in \mathcal{L}(H_E, K)$. For the original mapping $A_0 : \mathcal{C}_0^1([0,1]) \subset L^2([0,1]) \mapsto L^2([0,1])$ consider the adjoint mapping ${A_0}^*$. If ${A_0}^*\eta = \xi$ then

$$\int_0^1 \varphi\,'(t)\overline{\eta}(t)dt = \int_0^1 \varphi(t)\overline{\xi}(t)dt \quad \Rightarrow \quad \int_0^1 \varphi\,'(t)\left[\overline{\eta}(t) + \int_0^t \overline{\xi}(s)ds\right] dt = 0$$

for all $\varphi \in \mathcal{C}_0^1([0,1])$. Hence η is differentiable and $\xi = -\eta\,' = {A_0}^*\eta$. Now consider the adjoint of the generalized mapping. If ${A_{0,E}}^*\eta = \zeta$ then

$$\int_0^1 \varphi\,'(t)\overline{\eta}(t)dt = \int_0^1 [\varphi(t)\overline{\zeta}(t) + \varphi\,'(t)\overline{\zeta}\,'(t)]dt$$

and therefore

$$\int_0^1 \varphi\,'(t)\left[\overline{\eta}(t) - \overline{\zeta}\,'(t) + \int_0^t \overline{\zeta}(s)ds\right] dt = 0$$

for all $\varphi \in \mathcal{C}_0^1([0,1])$. *Hence* $\zeta\,'$ *is differentiable and* $\zeta - \zeta\,'' = -\eta\,'$. *It follows that*

$$(I + {A_0}^* A_0) {A_{0,E}}^* = {A_0}^* \quad \Leftrightarrow \quad {A_{0,E}}^* = (I + {A_0}^* A_0)^{-1} {A_0}^*.$$

Example 6 (Discrete spectrum) *Each element in the space* $L^2([0,1])$ *can be represented by a Fourier series and defined by a countably infinite discrete spectrum. A bounded linear operator on any subspace of* $L^2([0,1])$ *can be regarded as a linear transformation on a discrete spectrum. Let* $H = H^2([0,1]) \cap H_0^1([0,1])$ *be the Hilbert space of measurable functions* $x : [0,1] \mapsto \mathbb{C}$ *with*

$$\int_{[0,1]} \left[|x(t)|^2 + |x\,'(t)|^2 + |x\,''(t)|^2 \right] dt < \infty,$$

and $x(0) = x(1) = 0$ *and with inner product given by*

$$\langle x_1, x_2 \rangle_H = \int_{[0,1]} [x_1(t)\overline{x}_2(t) + {x_1}'(t){\overline{x}_2}'(t) + {x_1}''(t){\overline{x}_2}''(t)]\, dt.$$

Let $K = L^2([0,1])$ *be the Hilbert space of measurable functions* $y : [0,1] \mapsto \mathbb{C}$. *Define* $A_0, A_1 \in \mathcal{L}(H,K)$ *by setting*

$$A_0 x = x\,'' + \pi^2 x \quad \text{and} \quad A_1 x = x$$

for all $x \in H$. *Note that* $\|x\,''\|_K^2 \le \|x\|_H^2$. *For each* $y \in K$ *and* $z \in \mathbb{C}$ *we wish to find* $x \in H$ *to solve the differential equation*

$$[x\,''(t) + \pi^2 x(t)] + z x(t) = y(t).$$

This equation can be written in the form $(A_0 + A_1 z)x = y$ *and hence the solution is given by* $x = (A_0 + A_1 z)^{-1} y$ *provided the inverse exists. If* $\boldsymbol{e}_k : [0,1] \mapsto \mathbb{C}$ *is defined by* $\boldsymbol{e}_k(t) = \sqrt{2}\sin k\pi t$ *for each* $k = 1, 2, \ldots$ *and all* $t \in [0,1]$ *then each* $x \in H$ *can be written as* $x = \sum_{k=1}^{\infty} x_k \boldsymbol{e}_k$ *where* $x_k \in \mathbb{C}$ *and* $\sum_{k=1}^{\infty} k^4 |x_k|^2 < \infty$ *and each* $y \in K$ *can be written as* $y = \sum_{k=1}^{\infty} y_k \boldsymbol{e}_k$ *where* $y_k \in \mathbb{C}$ *and* $\sum_{k=1}^{\infty} |y_k|^2 < \infty$. *The operator* A_0 *is singular because* $A_0 \boldsymbol{e}_1 = 0$. *Nevertheless* $(A_0 + A_1 z)$ *is non-singular for* $0 < |z| < 3\pi^2$ *and equating coefficients in the respective Fourier series gives the solution*

$$x_1 = y_1/z \quad \text{and} \quad x_k = (-1)y_k/[\pi^2(k^2-1) - z] \text{ for } k \ge 2.$$

By writing the solution in the form

$$\begin{aligned} x &= \frac{y_1 \boldsymbol{e}_1}{z} - \sum_{k=2}^{\infty} \frac{y_k \boldsymbol{e}_k}{\pi^2(k^2-1)} \left[1 + \frac{z}{\pi^2(k^2-1)} + \cdots \right] \\ &= y_1 \boldsymbol{e}_1 \cdot \frac{1}{z} - \sum_{k=2}^{\infty} \frac{y_k \boldsymbol{e}_k}{\pi^2(k^2-1)} \cdot 1 - \sum_{k=2}^{\infty} \frac{y_k \boldsymbol{e}_k}{[\pi^2(k^2-1)]^2} \cdot z - \cdots \end{aligned}$$

for $0 < |z| < 3\pi^2$ *we can see that the expansion is a Laurent series with a pole of order* 1 *at* $z = 0$.

7 Inversion of perturbed linear operators on Banach space

Our results can also be applied in some Banach spaces. The material in this section is based on the general theory of Banach spaces described in [5], [6], [11], [15], [21] and [23]. In particular we use the terminology of Kato [15] in the following matter. Let X be a Banach space over the field $\mathbb{C}$ of complex numbers. The space X^* is the space of all bounded conjugate linear functionals on X. Thus $X^* = \overline{\mathcal{L}(X, \mathbb{C})}$ and for each $f \in X^*$ and each $z_1, z_2 \in \mathbb{C}$ and $x_1, x_2 \in X$ we have

$$\langle f, z_1 x_1 + z_2 x_2 \rangle = \overline{z_1} \langle f, x_1 \rangle + \overline{z_2} \langle f, x_2 \rangle.$$

Remark 7.1 *In this section are guided by the following observations. Let X, Y be Banach spaces with $A_0, A_1 \in \mathcal{L}(X, Y)$. Let $M = {A_0}^{-1}(\{0\})$ and $N = A_1(M)$. If $X = M \oplus M'$ and $Y = N \oplus N'$ where M' and N' are complementary spaces then the inversion procedure used for Hilbert spaces can be applied in exactly the same way. The problem is that not all closed subspaces can be complemented in Banach space. Our investigation is therefore directed towards the case when the subspaces M and N can be defined by bounded linear projections.*

7.1 General projection methods

We note the following well known result which is a compilation of results given in [5] p. 37, [11] p. 232 and [21] p. 111.

Theorem 7.2 *If X is a uniformly convex Banach space then X is reflexive and strictly convex. If M is a closed linear subspace of X then for each $x \in X$ there is a uniquely defined element $x_M \in M$ such that $||x - x_M|| \leq ||x - m||$ for all $m \in M$.*

Unfortunately the projection $P_M : X \mapsto X$ defined by $P_M(x) = x_M$ is generally non linear and so it is not possible to proceed as before. In certain special cases the above projection is linear.

7.2 Linear projections

We begin with an example of a linear projection in a uniformly convex space.

Example 7 *Let $p \in \mathbb{R}$ with $1 < p < \infty$ and let $X = l^p$. Let $f \in X^* = l^q$ where $1/p + 1/q = 1$ and $\langle f, x \rangle = \sum_{k=1}^{\infty} f_k \overline{x_k}$ for each $x \in l^p$. The space l^p is uniformly convex. If*

$$M = \{m \mid m \in l^p \text{ and } \langle f, m \rangle = 0\}$$

then the natural projection $\xi = P_M x$ is defined by

$$\xi_k = x_k - \frac{\hat{f}_k |f_k|^{q-1} \overline{\langle f, x \rangle}}{||f||_q^q}$$

where $\hat{f}_k = \overline{f_k}/|f_k|$. *The projection is clearly linear in* x.

In some cases linear projections can be constructed. It is well known for instance, that bounded linear projections can always be constructed if the subspace is closed and finite dimensional or finite co-dimensional. Let X, Y be Banach spaces over the field $\mathbb{C}$ and let X^*, Y^* be the corresponding adjoint spaces. We need the following elementary result.

Lemma 7.3 *Let* $T \in \mathcal{L}(X, Y)$. *If* T *is a* 1–1 *mapping of* X *onto* Y *then the adjoint mapping* $T^* \in \mathcal{L}(Y^*, X^*)$ *is a* 1–1 *mapping of* Y^* *onto* X^*. *Note that if* $f \in X^*$ *and* $g = f \circ T^{-1} \in Y^*$ *then* $f = T^* g$.

Example 8 *Let* $A_0, A_1 \in \mathcal{L}(X, Y)$ *and suppose that for some finite linearly independent set* $\{f_j\}_{j=1,2,\ldots,r} \subset X^*$ *the subspace* $M = {A_0}^{-1}(\{0\}) \subset X$ *is defined by*

$$M = \{m \mid \langle f_j, m\rangle = 0 \text{ for each } j = 1, 2, \ldots, r\}.$$

Choose $\{x_k\}_{k=1,2,\ldots,r} \subset X$ *such that* $\langle f_j, x_k\rangle = \delta_{jk}$ *where* δ_{jk} *is the Kronecker delta and define* $P_M \in \mathcal{L}(X)$ *by the formula*

$$P_M(x) = x - \sum_{k=1}^{r} \overline{\langle f_k, x\rangle} x_k.$$

Since P_M *and* $I - P_M$ *are linear projections we can write* $X = P_M(X) \oplus [I - P_M](X)$. *Suppose there exists* $z_0 \in \mathbb{C}$ *with* $z_0 \neq 0$ *so that* ${T_0}^{-1} = z_0(A_0 + A_1 z_0)^{-1} \in \mathcal{L}(Y, X)$ *is well defined and let* $\{g_j\}_{j=1,2,\ldots,r} \subset Y^*$ *be given by the formula*

$$g_j = [{T_0}^*]^{-1} f_j = \overline{z_0}(A_0^* + A_1^* \overline{z_0})^{-1} f_j.$$

For each $m \in M$ *we have* $\langle g_j, A_1 m\rangle = \langle [{T_0}^{-1}]^* f_j, T_0 m\rangle = \langle f_j, m\rangle$ *and so the subspace* $N = A_1(M) \subset Y$ *has the form*

$$N = \{n \mid \langle g_j, n\rangle = 0 \text{ for each } j = 1, 2, \ldots, r\}.$$

Since $\{g_j\}_{j=1,2,\ldots,r} \subset Y^*$ *are linearly independent there exist* $\{y_k\}_{k=1,2,\ldots,r} \subset Y$ *such that* $\langle g_j, y_k\rangle = \delta_{jk}$ *and we can define* $Q_N \in \mathcal{L}(Y)$ *by the formula*

$$Q_N(y) = y - \sum_{k=1}^{r} \overline{\langle g_k, y\rangle} y_k$$

for each $y \in Y$. *Because* Q_N *and* $I - Q_N$ *are linear projections we can write* $Y = Q_N(Y) \oplus [I - Q_N](Y)$.

7.3 Mean transition times for a perturbed Markov process

The intrinsic structure of a Markov process can be substantially changed by a small perturbation. For instance the perturbation may introduce state transitions that are not possible in the original unperturbed process. In such cases the mean passage times between states can be calculated by finding the inverse of a perturbed linear operator. We introduce our example with a brief motivating discussion but refer to [3] for technical information about the terminology. Consider a Markov chain defined on the discrete state space $S = \{0, \frac{1}{r}, \frac{2}{r}, \ldots, 1\}$ with transition probabilities defined by the matrix

$$T = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{r} & \frac{1}{r} & \frac{1}{r} & \cdots & \frac{1}{r} & 0 \\ \frac{1}{r+1} & \frac{1}{r+1} & \frac{1}{r+1} & \cdots & \frac{1}{r+1} & \frac{1}{r+1} \end{bmatrix}.$$

The transition matrix acts on a discrete probability measure $\boldsymbol{\pi} \in \mathbb{R}^{1\times(r+1)}$ to produce a transformed discrete probability measure $T(\boldsymbol{\pi}) \in \mathbb{R}^{1\times(r+1)}$ defined by the formula $T(\boldsymbol{\pi}) = \boldsymbol{\pi} \cdot T$. Thus for each $j = 0, 1, \ldots, r$ we have

$$T(\pi_j) = \sum_{k=j}^{r} \frac{\pi_k}{k+1}.$$

If we write $\pi_j = \Delta\xi_j$ and define the cumulative probability $\xi_j = \Delta\xi_0 + \Delta\xi_1 + \cdots + \Delta\xi_j$ then by an appropriate sum of the above equations we obtain

$$T(\xi_j) = \xi_j + (j+1) \sum_{k=j+1}^{r} \frac{\Delta\xi_k}{k+1}.$$

An analogous Markov process with a continuous state space $[0, 1]$ is defined by the formula

$$T\xi([0,t]) = \xi([0,t]) + t \int_{(t,1]} \frac{d\xi([0,s])}{s}$$

for all $t \in [0, 1]$. The transformation T now acts on the space of regular countably additive measures on $[0, 1]$. The following example shows how our proposed inversion procedure can be applied to the calculation of mean passage times for the perturbed continuous state Markov process defined by the operator $T_\epsilon = (1-\epsilon)I + \epsilon T$. The mean passage times are determined by the deviation operator $[I - T_\epsilon + T_\epsilon^\infty]^{-1} - T_\epsilon^\infty$. The properties of the deviation operator are described in [3]. Markov processes with continuous state space are discussed in [23].

Example 9 *Let* $X = \mathcal{C}([0, 1])$ *be the Banach space of continuous complex valued functions on* $[0, 1]$ *and* $X^* = \mathrm{rca}([0, 1])$ *the corresponding adjoint space of regular* *countably additive complex valued measures on* $[0,1]$. *Define a continuous state Markov process* $T : X^* \mapsto X^*$ *by the formula*

$$T\xi([0,t]) = \xi([0,t]) + t\int_{(t,1]} \frac{d\xi([0,s])}{s}$$

for $t \in [0,1)$ *with* $T\xi([0,1]) = \xi([0,1])$. *Consider the transformation* $T_\epsilon : X^* \mapsto X^*$ *defined by*

$$T_\epsilon = (1-\epsilon)I + \epsilon T$$

where $I : X^* \mapsto X^*$ *is the identity transformation. The transformation* T_ϵ *is a perturbation of the identity that allows a small probability of transition between states. We will investigate the key operator*

$$[I - T_\epsilon + T_\epsilon^\infty]^{-1}$$

where $T_\epsilon^\infty = lim_{n\to\infty} T_\epsilon^n$. *We can see that*

$$dT\xi([0,t]) = \left(\int_{(t,1]} \frac{d\xi([0,s])}{s}\right) dt$$

and if we define $E : X \mapsto X$ *by setting*

$$E\varphi(s) = \frac{1}{s}\int_{[0,s]} \varphi(t)dt$$

for each $\varphi \in X$ *then*

$$\begin{aligned}\langle T\xi, \varphi\rangle &= \int_{[0,1]} \overline{\varphi(t)}\left(\int_{(t,1]} \frac{d\xi([0,s]}{s}\right) dt \\ &= \int_{[0,1]}\left(\frac{1}{s}\int_{[0,s]}\overline{\varphi(t)}dt\right) d\xi([0,s]) \;=\; \langle \xi, E\varphi\rangle.\end{aligned}$$

Thus $T = E^*$. *For each* $n = 0,1,\ldots$ *it is not difficult to show that*

$$E^{n+1}\varphi(s) = \int_{[0,s]} w_n(s,t)\varphi(t)dt$$

where $w_n(s,t) = [\ln(s/t)]^n/[n!s]$. *Note that* $w_n(s,t) \geq 0$ *for* $t \in (0,s]$ *with*

$$\int_{[0,s]} w_n(s,t)dt = 1$$

and $w_n(s,t) \downarrow 0$ *uniformly in* t *for* $t \in [\sigma, s]$ *for each* $\sigma > 0$ *as* $n \to \infty$. *It follows that* $E^{n+1}\varphi(s) \to \varphi(0)\chi_{[0,1]}(s)$ *for each* $s \in [0,1]$. *Hence we deduce that*

$$\langle T^{n+1}\xi, \varphi\rangle = \langle \xi, E^{n+1}\varphi\rangle \to \xi([0,1])\overline{\varphi(0)}$$

for each $\varphi \in X$. If we define the Dirac measure $\delta \in X^*$ by the formula $\langle \delta, \varphi \rangle = \overline{\varphi(0)}$ then we can say that $T^{n+1}\xi \to T^\infty \xi = \xi([0,1])\delta$ in the weak sense. Let $\varphi \in X$ be any fixed test function and let τ be a positive real number. We can find $N \in \mathbf{N}$ such that*

$$|\langle T^k \xi, \varphi \rangle - \xi([0,1])\overline{\varphi(0)}| < \tau$$

for all $k \geq N+1$. It follows that

$$\begin{aligned}
&|\langle T_\epsilon^{n+1}\xi, \varphi\rangle - \xi([0,1])\overline{\varphi(0)}| \\
&\quad = \left| \sum_{k=0}^{n+1} \binom{n+1}{k} (1-\epsilon)^{n+1-k}\epsilon^k \left[\langle T^k\xi, \varphi\rangle - \xi([0,1])\overline{\varphi(0)} \right] \right| \\
&\quad \leq \sum_{k=0}^{N} \binom{n+1}{k} (1-\epsilon)^{n+1-k}\epsilon^k \left| \langle T^k\xi, \varphi\rangle - \xi([0,1])\overline{\varphi(0)} \right| \\
&\qquad\qquad + \left[\sum_{k=N+1}^{n} \binom{n+1}{k} (1-\epsilon)^{n+1-k}\epsilon^k \right] \tau \\
&\quad \leq \left[\sum_{k=0}^{N} \binom{n+1}{k} (1-\epsilon)^{n+1-k}\epsilon^k \left| \langle T^k\xi, \varphi\rangle - \xi([0,1])\overline{\varphi(0)} \right| \right] + \tau
\end{aligned}$$

and hence

$$\limsup_{n\to\infty} |\langle T_\epsilon^{n+1}\xi, \varphi)\rangle - \xi([0,1])\overline{\varphi(0)}| \leq \tau.$$

Since τ is arbitrary it follows that $\langle T_\epsilon^{n+1}\xi, \varphi\rangle \to \xi([0,1])\overline{\varphi(0)}$ for each $\varphi \in X$. Thus we say that $T_\epsilon^{n+1}\xi \to T_\epsilon^\infty \xi = \xi([0,1])\delta$ in the weak sense. Hence $T_\epsilon^\infty = T^\infty$. If we define $A_0 = T^\infty$ and $A_1 = I - T$ then the equation*

$$[I - T_\epsilon + T_\epsilon^\infty]\xi = \eta \quad \Leftrightarrow \quad [T^\infty + \epsilon(I-T)]\xi = \eta$$

can be rewritten as $(A_0 + A_1\epsilon)\xi = \eta$ where A_0 is singular. The null space of A_0 is given by

$$M = {A_0}^{-1}(\{0\}) = \{\mu \mid \mu([0,1]) = 0\}$$

and the projection $P_M : X^* \mapsto X^*$ onto M is defined by

$$\mu = P_M \xi = \xi - \xi([0,1])\delta$$

for each $\xi \in X^*$. We wish to find a simple description for the space $N = A_1(M)$. On the one hand if $\nu = (I-T)\mu$ then $\langle \nu, \varphi\rangle = \langle \mu, \varphi - E\varphi\rangle$ for $\varphi \in X$. If we write $\chi_{[0,1]}$ to denote the characteristic function of the interval $[0,1]$ then since $E\chi_{[0,1]} = \chi_{[0,1]}$ it follows that

$$\nu([0,1]) = \langle \nu, \chi_{[0,1]}\rangle = \langle \mu, \chi_{[0,1]} - E\chi_{[0,1]}\rangle = 0.$$

On the other hand suppose $\nu([0,1]) = 0$. If we set $\psi = \varphi - E\varphi$ then $\psi \in X$ and $\psi(0) = 0$. By solving an elementary differential equation it can be seen that

$\varphi - E\varphi(1)\chi_{[0,1]} = \psi - F\psi$ where

$$F\psi(s) = \int_{(s,1]} \frac{\psi(t)}{t} dt.$$

Note that $F\psi(0) = E\varphi(1) - \varphi(0)$ is well defined. Define $\langle \mu, \psi \rangle = \langle \nu, \psi - F\psi \rangle$ for each $\psi \in X$ with $\psi(0) = 0$. Since $\langle \nu, \chi_{[0,1]} \rangle = 0$ we deduce that

$$\langle \nu, \varphi \rangle = \langle \nu, \varphi - E\varphi(1)\chi_{[0,1]} \rangle = \langle \nu, \psi - F\psi \rangle = \langle \mu, \psi \rangle = \langle \mu, \varphi - E\varphi \rangle$$

for each $\varphi \in X$. Therefore $\nu = (I - T)\mu$ and hence

$$N = A_1(M) = \{\nu \mid \nu([0,1]) = 0\}$$

and the projection $Q_N : X^* \mapsto X^*$ is defined by

$$\nu = Q_N \eta = \eta - \eta([0,1])\delta$$

for each $\eta \in X^*$. By applying an appropriate decomposition to the given equation with $\mu = P_M \xi \in M$ and $\nu = Q_N \eta \in N$ and by noting that $A_0 \delta = \delta$ and $A_1 \mu = \mu(I - E)$ we obtain

$$\epsilon \mu(I - E) + \xi([0,1])\delta = \nu + \eta([0,1])\delta.$$

By equating corresponding terms we have $\epsilon\mu(I - E) = \nu$ and $\xi([0,1])\delta = \eta([0,1])\delta$. The former equation means that $\epsilon\langle \mu, \varphi - E\varphi \rangle = \langle \nu, \varphi \rangle$ for each $\varphi \in X$ and could be rewritten as $\epsilon \langle \mu, \psi \rangle = \langle \nu, \psi - F\psi \rangle$ for each $\psi \in X$ with $\psi(0) = 0$. Thus $\epsilon\mu = \nu(I - F)$. Since $\xi = \mu + \xi([0,1])\delta$ the solution is given by

$$\xi = \frac{1}{\epsilon}\nu(I - F) + \eta([0,1])\delta = \frac{1}{\epsilon} Q_N \eta (I - F) + (I - Q_N)\eta.$$

As expected there is a pole of order one at $\epsilon = 0$.

8 Polynomial perturbations

We extend the above results for linear perturbations on Hilbert space to polynomial perturbations. In view of the previous remarks we assume all operators are bounded.

8.1 The augmented operator notation

Let H and K be Hilbert spaces and let $\{A_i\} \subset \mathcal{L}(H, K)$. For $k \in \mathbb{N}$ with $k > 1$ define $\mathcal{A}_0^{(k)} \in \mathcal{L}(H^k, K^k)$ by setting

$$\mathcal{A}_0^{(k)} = \begin{bmatrix} A_0 & 0 & \cdots & 0 \\ A_1 & A_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{k-1} & A_{k-2} & \cdots & A_0 \end{bmatrix}$$

and $\mathcal{A}_r^{(k)} \in \mathcal{L}(H^k, K^k)$ for $r \geq 1$ by setting

$$\mathcal{A}_r^{(k)} = \begin{bmatrix} A_{rk} & A_{rk+1} & \cdots & A_{rk-k+1} \\ A_{rk+1} & A_{rk+2} & \cdots & A_{rk-k+2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{rk+k-1} & A_{rk+k-2} & \cdots & A_{rk} \end{bmatrix}.$$

For $r = 0, 1, \ldots$ and each $X \in H^k$ the value $\mathcal{A}_r^{(k)}(X)$ is defined by formal matrix multiplication. Define $\mathcal{D} : \mathcal{L}(H, K) \mapsto \mathcal{L}(H^k, K^k)$ by

$$\mathcal{D}(A) = \begin{bmatrix} A & 0 & \cdots & 0 & 0 \\ 0 & A & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & A & 0 \\ 0 & 0 & \cdots & 0 & A \end{bmatrix}$$

for each $A \in \mathcal{L}(H, K)$ and define $\mathcal{Z}(z) \in \mathcal{L}(H^k, H^k)$ by

$$\mathcal{Z}(z) = \begin{bmatrix} 0 & 0 & \cdots & 0 & zI \\ I & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}$$

for all $z \in \mathbb{C}$. We will normally write

$$\mathcal{Z} = [E_2, E_3, \ldots, E_k, zE_1].$$

We note that

$$\begin{array}{rcl} \mathcal{Z}^2 & = & [E_3, E_4, \ldots, E_k, , zE_1, zE_2] \\ \mathcal{Z}^3 & = & [E_4, E_5, \ldots, zE_1, zE_2, zE_3] \\ \cdots & & \cdots \\ \mathcal{Z}^k & = & z[E_1, E_2, \ldots, E_{k-1}, E_k] = z\mathcal{I} \end{array}$$

and in general for $r = 0, 1, \ldots$ and $s = 0, 1, \ldots, k-1$ we have

$$\mathcal{Z}^{rk+s} = z^r[E_{s+1}, E_{s+2}, \ldots, E_k, zE_1, \ldots, zE_s].$$

Since $\mathcal{Z}^k = zI$ it follows that $\|\mathcal{Z}\| = |z|^{1/k}$.

8.2 Some equivalent series

Let $\{A_i\}_{i=0,1,\ldots} \subset \mathcal{L}(H, K)$.

Lemma 8.1 $\sum_{i=0}^{\infty} \mathcal{D}(A_i)\mathcal{Z}^i = \sum_{r=0}^{\infty} \mathcal{A}_r^{(k)} z^r$ *for all* $z \in \mathbb{C}$.

Proof. Expand the LHS and collect terms according to the powers of z. $\square$

Lemma 8.2 *$\sum_{i=0}^{\infty} \mathcal{D}(A_i)\mathcal{Z}^i$ converges for $\|\mathcal{Z}\| < b^{1/k}$ if and only if $\sum_{i=0}^{\infty} A_i z^i$ converges for $|z| < b$.*

Proof. Suppose the LHS converges when $\|\mathcal{Z}\| < b^{1/k} \Leftrightarrow |z| < b$. By expanding the first column of the LHS we see that the series

$$A_s + A_{s+k}z^k + A_{s+2k}z^{2k} + \cdots$$

converges when $|z| < b$ for each $s = 0, 1, \ldots, k-1$. Hence the series

$$\sum_{s=0}^{k-1} \left[A_s + A_{s+k}z^k + A_{s+2k}z^{2k} + \cdots \right] z^s \;=\; \sum_{i=0}^{\infty} A_i z^i$$

also converges for $|z| < b$. Thus the RHS converges when $|z| < b$. The reverse implication is established by a similarly elementary argument. $\square$

Lemma 8.3

$$\left(\sum_{i=0}^{\infty} \mathcal{D}(A_i)\mathcal{Z}^i \right) \left(\sum_{i=0}^{\infty} \mathcal{D}(X_i)\mathcal{Z}^i \right) = \mathcal{Z}^m$$

is valid for some non-negative integer m if and only if

$$\left(\sum_{i=0}^{\infty} A_i z^i \right) \left(\sum_{i=0}^{\infty} X_i z^i \right) = z^m I$$

is also valid.

Proof. Both identities are true if and only if

$$\sum_{j=0}^{i} A_j X_{i-j} = \delta_{im} I$$

for each $i = 0, 1, \ldots$ where δ_{im} is the Kronecker delta. If $\|A_j\| \leq h^{j+1}$ for some $h \in \mathbb{R}$ with $h > 0$ then $\|X_j\| \leq k^{j+1}$ for some $k \in \mathbb{R}$ with $k > 0$. An inductive argument justifying geometric bounds of this type can be found in [12]. $\square$

8.3 More notation

For $s < k$ define $\mathcal{X}_i^{(k,s)} \in \mathcal{L}(K^k, H^k)$ by

$$\mathcal{X}_0^{(k,s)} = \begin{bmatrix} 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ X_0 & 0 & \cdots & 0 & \cdots & 0 \\ X_1 & X_0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ X_{s-1} & X_{s-2} & \cdots & X_0 & \cdots & 0 \end{bmatrix},$$

$$\mathcal{X}_1^{(k,s)} = \begin{bmatrix} X_s & \cdots & X_0 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ X_k & \cdots & X_{k-s} & \cdots & X_0 \\ \vdots & & \vdots & & \vdots \\ X_{s+k-1} & \cdots & X_{k-1} & \cdots & X_{s+1} \end{bmatrix}$$

and

$$\mathcal{X}_r^{(k,s)} = \begin{bmatrix} X_{rk+s} & \cdots & X_{rk-k+s+1} \\ \vdots & & \vdots \\ X_{(r+1)k+s-1} & \cdots & X_{rk+s} \end{bmatrix}$$

for $r > 1$.

8.4 The polynomial inversion formulae

The polynomial inversion formula is equivalent to a corresponding linear inversion using augmented operators.

Theorem 8.4 *The inverse operator*

$$(A_0 + \cdots + A_k z^k)^{-1} \in \mathcal{L}(K, H)$$

is given by the formula

$$(A_0 + \cdots + A_k z^k)^{-1} = \frac{1}{z^{rk+s}}(X_0 + X_1 z + \cdots)$$

where $r \in \{0, 1, \ldots\}$ *and* $s \in \{0, 1, \ldots, k-1\}$ *if and only if the inverse operator*

$$(\mathcal{A}_0^{(k)} + \mathcal{A}_1^{(k)} z)^{-1} \in \mathcal{L}(K^k, H^k)$$

is given by the formula

$$(\mathcal{A}_0^{(k)} + \mathcal{A}_1^{(k)} z)^{-1} = \frac{1}{z^r}(\mathcal{X}_0^{(k)} + \mathcal{X}_1^{(k)} z + \cdots)$$

when $s = 0$ *and*

$$(\mathcal{A}_0^{(k)} + \mathcal{A}_1^{(k)} z)^{-1} = \frac{1}{z^{r+1}}(\mathcal{X}_0^{(k,s)} + \mathcal{X}_1^{(k,s)} z + \cdots)$$

when $s \in \{1, 2, \ldots, k-1\}$.

Proof. The proof follows by expanding the various expressions in the statement of the theorem and equating corresponding elements. □

Example 10 *Let*

$$A_0 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \textit{and} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

and define $A(z) = A_0 + A_1 z + A_2 z^2$*. An elementary calculation shows that*

$$X(z) = A(z)^{-1} = \frac{1}{z^3} \begin{bmatrix} z^2 & -1 \\ -z^2 & 1+z \end{bmatrix}$$

for $z \neq 0$ *and hence we have*

$$X_0 = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}, \quad X_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \textit{and} \quad X_2 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$$

with $X_j = 0$ *for* $j > 2$*. We note also that*

$$\mathcal{A}_0^{(2)} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \textit{and} \quad \mathcal{A}_1^{(2)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

If we define $\mathcal{A}^{(2)}(z) = \mathcal{A}_0^{(2)} + \mathcal{A}_1^{(2)} z$ *then another elementary calculation shows that*

$$\mathcal{X}(z) = \mathcal{A}^{(2)}(z)^{-1} = \frac{1}{z^2} \begin{bmatrix} 0 & 0 & z^2 & -z \\ 0 & z & -z^2 & z \\ z & -1 & 0 & 0 \\ -z & 1 & 0 & z \end{bmatrix}$$

for $z \neq 0$*. By comparing the various expressions we can see that* $k = 2$*,* $r = 1$ *and* $s = 1$*. Thus we write* $\mathcal{X}(z) = \mathcal{X}^{(2,1)}(z)$ *and observe that*

$$\mathcal{X}_0^{(2,1)} = \begin{bmatrix} 0 & 0 \\ X_0 & 0 \end{bmatrix}, \quad \mathcal{X}_1^{(2,1)} = \begin{bmatrix} X_1 & X_0 \\ X_2 & X_1 \end{bmatrix}, \quad \mathcal{X}_2^{(2,1)} = \begin{bmatrix} 0 & X_2 \\ 0 & 0 \end{bmatrix}$$

with $\mathcal{X}_j^{(2,1)} = 0$ *for* $j \geq 3$.

9 Analytic perturbations

Suppose $A(z)$ converges in the region $|z| < a$. If $[A(z_0)]^{-1}$ is well defined for some $z_0 \neq 0$ with $|z_0| < a$ then by the Banach Inverse Theorem we can find $\delta > 0$ such that

$$\|A(z_0)x\| \geq \delta \|x\|$$

for all $x \in H$. Because the series converges in norm there exists m such that

$$\left\| \left[\sum_{i=0}^{m} A_i z_0^i \right] x \right\| \geq \frac{\delta}{2} \|x\|$$

for all $x \in H$. Hence $[A_m(z_0)]^{-1}$ is well defined. Since $A_m(z)$ is a polynomial perturbation we can calculate $[A_m(z)]^{-1}$ and

$$[A(z)]^{-1} = [A_m(z) + R_m(z)]^{-1} = [I + A_m(z)^{-1} R_m(z)]^{-1} A_m(z)^{-1}.$$

10 Conclusions and future work

We have shown that when $A(0)$ is singular it is nevertheless possible that the operator $A(z)^{-1}$ is well defined on some region $0 < |z| < b$ by a Laurent series expansion with a finite order pole at $z = 0$. In future work we will analyse the perturbation of certain generalized inverse operators defined by a well posed minimum norm problem. For instance we wish to find a bounded linear operator $X(z) = X_0(z)$ of Hilbert-Schmidt type which solves the problem

$$\min_{X(z)} \|A(z)X(z)B(z) - C(z)\|_2$$

where $A(z)$, $B(z)$ are bounded linear operators, $C(z)$ is a bounded linear operator of Hilbert-Schmidt type and $\|\cdot\|_2$ is the Hilbert-Schmidt norm. If $A(0)$ and $B(0)$ are singular but $A(z)^{-1}$ and $B(z)^{-1}$ are well defined for $z \neq 0$ we believe the fundamental solution $X_0(z)$ is a Tichonov regularisation that can be represented near $z = 0$ by a Maclaurin series. We note Theorem 2.3 on pp. 144–145 in the book by Gohberg *et al* [9] which guarantees a well posed problem and recent work by Golub *et al* [10] on numerical solution of a corresponding matrix problem which suggests the Tichonov regularisation.

References

[1] Konstantin Avrachenkov, *Analytic perturbation theory and its applications*, PhD thesis, University of South Australia, 1999.

[2] Kostantin E. Avrachenkov, M. Haviv and P.G. Howlett, 2001, *Inversion of analytic matrix functions that are singular at the origin*, SIAM J. Matrix Anal. Appl., **22**, No. 4, pp. 1175–1189.

[3] K.E. Avrachenkov, Jerzy Filar and Moshe Haviv, 2002, Singular perturbations of Markov Chains and Decision Processes, *Handbook of Markov Decision Processes, Methods and Applications*, Eds. Eugene A. Feinberg and Adam Shwartz, International Series in Operations Research & Management Science, **40**, Kluwer, pp. 113–150.

[4] R. Courant and D. Hilbert, 1953, *Methods of Mathematical Physics, Vol. 1*, Interscience Publishers.

[5] Joseph Diestel, 1975, *Geometry of Banach Spaces - Selected Topics*, Lecture Notes in Mathematics, Eds. A. Dold and B. Eckerman, Vol. 485, Springer-Verlag.

[6] Nelson Dunford and Jacob Schwartz, 1988, *Linear Operators, Part I: General Theory*, Wiley Classics, John Wiley and Sons.

[7] I. Gohberg, M.A. Kaashoek and F. van Schagen, 1993, *On the local theory of regular analytic matrix functions*, Lin. Alg. Appl. **182** pp. 9–25.

[8] Israel Gohberg, Peter Lancaster and Leiba Rodman, 2006, *Invariant Subspaces of Matrices with Applications*, Classics in Applied Mathematics, Vol. 51, SIAM.

[9] Israel Gohberg, Seymour Goldberg and Marinus A. Kaashoek, *Classes of Linear Operators, Vol. 1*, Operator Theory: Advances and Applications, **49**, Birkhauser, 1990.

[10] G.H. Golub, P.C. Hansen and D.P. O'Leary, 1999, *Tichonov regularization and total least squares*, SIAM J. Matrix Anal. and Appl., **21**, 1, pp. 185–194.

[11] Edwin Hewitt and Karl Stromberg, 1975, *Real and Abstract Analysis*, Graduate Texts in Mathematics, **25**, Springer-Verlag.

[12] P.G.Howlett, 1982, *Input retrieval in finite dimensional linear systems*, J. Austral. Math. Soc., Series B (now the ANZIAM Journal), **23**, pp. 357–382.

[13] P.G. Howlett and K.E. Avrachenkov, 2001, *Laurent series for the inversion of perturbed lionear operators on Hilbert space*, Progress in Optimisation III, Contributions from Australasia, A. Rubinov Ed., Kluwer, pp. 325–342, 2001.

[14] P.G. Howlett, V. Ejov and K.E. Avrachenkov, 2003, *Inversion of perturbed linear operators that are singular at the origin*, Proceedings of 42nd IEEE Conference on Decision and Control, Maui, Hawai, pp. 5628–5631 (on Compact Disc), December 2003.

[15] Tosio Kato, 1995, *Perturbation Theory for Linear Operators*, Classics in Mathematics, Reprint of the 1980 Edition, Springer.

[16] M.V. Keldysh, 1951, *On eigenvalues and eigenfunctions of certain classes of non self-adjoint operators*, Dokl. Akad. Nauk., USSR **77**, pp. 11–14.

[17] Peter Lancaster, *Lambda-matrices and Vibrating Systems*, Dover Publications, 2002.

[18] D.G. Luenberger, 1969, *Optimization by vector space methods*, John Wiley & Sons.

[19] M.K.Sain and J.L.Massey, 1969, *Invertibility of linear time invariant dynamical systems*, IEEE Trans. Auto. Control, **AC-14**, pp. 141–149.

[20] P. Schweitzer and G.W. Stewart, 1993, *The Laurent expansion of pencils that are singular at the origin*, Linear Algebra and its Applications, **183**, pp. 237–254.

[21] Ivan Singer, 1970, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Springer-Verlag.

[22] M.I. Vishik; L.A. Lyusternik, 1960, *The solution of some perturbation problems in the case of matrices and self-adjoint and non-self-adjoint differential equations*, Russian Math. Surveys **15**, pp. 1–73.

[23] Kosaku Yosida, *Functional Analysis*, Fifth Edition, Springer-Verlag, 1978.