480c

ANGLES BETWEEN SUBSPACES

AND

APPLICATION TO PERTURBATION THEORY

ANGLES BETWEEN SUBSPACES AND APPLICATION TO PERTURBATION THEORY

BY

Nagwa A. Sherif, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

August, 1977

MASTER OF SCIENCE (1977) MCMASTER UNIVERSITY

(Mathematics) Hamilton, Ontario

TITLE: Angles Between Subspaces and Application to Perturbation Theory

AUTHOR: Nagwa Abd El.Monem Hassan Sherif, B.Sc.(Hons.) (Cairo University)

SUPERVISOR: Dr. Z.V. Kovarik

NUMBER OF PAGES: (vii) (149)

ABSTRACT

It is known that when two subspaces of a Hilbert space are in some sense close to each other, then there exists a unitary operator which is called the direct rotation. This operator maps one of the subspaces onto the other while being as close to identity as possible. In this thesis, we study such a pair of subspaces, and the application of the angles between them to the invariant subspace perturbation theory. We also develop an efficient algorithm for computing the direct rotation for pairs of subspaces of relatively small dimension.

To my husband

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENT

I express my thanks to Dr. Z.V. Kovarik for his assistance and encouragement during the preparation of this thesis. Without his help the completion of this work would have been impossible.

Appreciation is due also to the Mathematics Department of McMaster University.

Finally, I thank my family for their support throughout.

INTRODUCTION

Pairs of linear subspaces of a real n-dimensional inner product space of equal dimensions have been studied since 1875 [21]. Since then, it is known that this pair of subspaces has a number of angles equal to the dimension of each of them as unitary invariants. A treatment of the subject in somewhat modern style is in [16]. The subject was developed by S.N. Afriat [1] and others. The extension to the case of Hilbert space was completely analysed by C. Davis [8].

In chapter one, we study such a pair of subspaces of a Hilbert space. We define the direct rotation which maps one of the subspaces onto the other. This direct rotation was introduced by C. Davis [8] and T. Kato[24],§§1.4.6, 1.6.8].

The study of the direct rotation is greatly simplified using the idea of a source space, and the operator angle Θ. Following [11], we present a detailed study of the direct rotation and a complete set of unitary invariants of a pair of subspaces. We conclude this chapter by studying the extremal properties of the direct rotation.

In chapter 2, we study the operator equation $BX - XA = Q$ in different settings. We show that under certain conditions, the equation has a unique solution. Also we give an explicit formula for the solution in special cases. This equation will be of later use in chapters 3 and 4.

Chapter 3 is devoted to the case when a pair of subspaces consists of reducing subspaces of A and A+H where A and H are Hermitian operators and H is small in a sense specified in the text. Through this, we can shed some light on the behaviour of eigenvectors under perturbation. In the finite dimensional case, we give bounds on the difference between eigenvectors of a Hermitian matrix and those of a Hermitian perturbed matrix. In the infinite dimensional case, a Hermitian operator may not have eigenvalues, but still has reducing subspaces; in this case, we give bounds on the difference between corresponding reducing subspaces of A and A+H in terms of the operator angle Θ.

Chapter 4 is mainly concerned with the generalization of chapter 3 to the case where A is a closed (possibly nonselfadjoint) linear operator and the generalization is done from a different point of view.

Chapter 5 is devoted to algorithms for computing the direct rotation and the angles between subspaces. We define the angle bisector and prove some of its properties. We discuss and compare different methods for computing the direct rotation and introduce an algorithm, which is efficient for subspaces of low dimensions.

For the convenience of the reader, we include two appendices which contain the background necessary throughout the thesis. In appendix A, the polar representation of a bounded linear operator is presented.

In appendix B, we give some known results about the singular values of a completely continuous operator and the relation between unitary invariant norms and the singular values.

CHAPTER 1

The Separation of Two Subspaces

§1.1 The Aperture of Two Linear Manifolds.

The concept of the aperture of two linear manifolds was introduced by B. Nagy [38], and independently of him, by M.G. Krein and M.A. Krasnoselskii [27].

Let $\mathcal{H}$ be a Hilbert space, and let M and N be two linear manifolds in $\mathcal{H}$.

Definition 1.1.1

The aperture of two linear manifolds in $\mathcal{H}$ is defined as the norm of the difference of the operators which project $\mathcal{H}$ on the closures of these two linear manifolds. This aperture is denoted by $\delta(M,N)$:

(1.1.1) $\delta(M,N) = \|P-Q\| = \|Q-P\| = \|(I-P) - (I-Q)\|$,

where P and Q are the operators of projection onto $\bar{M}$ and $\bar{N}$. i.e. $P^2 = P$ and $P^* = P$, range $P = \bar{M}$, similarly for Q.

From this definition, it follows that

(1) $\delta(M,N) = \delta(\bar{M},\bar{N}) = \delta(\mathcal{H}\ominus M, \mathcal{H}\ominus N)$

(2) $\delta(M,N) \leq 1$, and equality holds if there exists a nonzero element of one of these manifolds, which is orthogonal to the other. This property follows from $\|(P-Q)h\|^2 = \|P(I-Q)h - (I-P)Qh\|^2 = \|P(I-Q)h\|^2 + \|(I-P)Qh\|^2 \leq \|(I-Q)h\|^2 + \|Qh\|^2 = \|h\|^2$.

Now, given any two subspaces of a Hilbert space, or equivalently two projectors P and Q, we have the following

Theorem 1.1.2 ([24], p. 56)

Two orthogonal projections P and Q such that $\|P-Q\|<1$ *are unitarily equivalent, that is, there is a unitary operator* U *with the property* $Q = UPU^{-1}$.

Proof

Let $R = (P-Q)^2$, then R commutes with P and Q. Similarly $(I-P-Q)^2$ commutes with P and Q, since I-P is a projector. We define $U = [QP + (I-Q)(I-P)]\,(I-R)^{-1/2} =$

$$= (I-R)^{-1/2}\,[QP + (I-Q)(I-P)].$$

U is well defined since $\|P-Q\| < 1$ so that $(I-R)^{-1/2}$ is obtainable, say, by Maclaurin series. It is easy to show that $U^*U = UU^* = I$ and $UP = QU$, since R commutes with P and Q. From $(I-Q)UP = 0$ it follows that $UP\mathcal{H} \subset Q\mathcal{H}$. Similarly $(I-P)U^*Q = 0$ implies that $U^*Q\mathcal{H} \subset P\mathcal{H}$, so that $UP\mathcal{H} = Q\mathcal{H}$, i.e. U is a unitary operator, taking $P\mathcal{H}$ onto $Q\mathcal{H}$ and $(I-P)\mathcal{H}$ onto $(I-Q)\mathcal{H}$.

Remark 1

A sufficient but not necessary condition for the existence of such an operator U is $\|P-Q\| < 1$. A necessary condition is $\dim P\mathcal{H} = \dim Q\mathcal{H}$. This condition is sufficient in the finite dimensional case, but it is far from being sufficient in infinite dimensional $\mathcal{H}$. See (1.3.2) below.

An equivalent definition of the aperture of two linear manifolds is given in [2] as follows:

Definition 1.1.3

$$\delta(M,N) = \max \{\sup_{\substack{f\in\bar{N} \\ \|f\|=1}} \|(I-P)f\|, \sup_{\substack{g\in\bar{M} \\ \|g\|=1}} \|(I-Q)g\|\}$$

where $\|(I-P)f\| = d(f,\bar{M})$, the distance between the point f and $\bar{M}$. The importance of this formula is that it can be used to define the aperture of two linear manifolds in a Banach space.

Remark 2

Other measures of the difference between the subspaces $P\mathcal{H}$ and $Q\mathcal{H}$ are:

(1) For a unit vector $x = Px$, to find how large $Qx-x$ is, Davis [9] estimates the following:

$\sup\{\|Qx-x\|;\ \|x\| = 1,\ x = Px\}$, and

(2) $\sup\{\inf[\|y-x\|,\ \|y\| = 1;\ y = Qy],\ \|x\| = 1,\ x = Px\}$

A much stronger result than theorem 1.1.2 was given by Kato ([24], p. 57):

Theorem 1.1.3

Let P and Q be two orthogonal projections, with $M = R(P)$, and $N = R(Q)$, such that $\|(I-Q)P\| = \delta<1$.

Then there are the following alternatives: Either (i) Q maps M onto N one-to-one and bicontinuously and $\|P-Q\| = \|(I-P)Q\| = \|(I-Q)P\| = \delta$; or (ii) Q maps M onto a proper subspace $N_0 \subset N$ one-to-one, and bicontinuously, if Q_0 is the orthogonal projection on N_0. Thus

$$\|P-Q_0\| = \|(I-P)Q_0\| =$$
$$= \|(I-Q_0)P\| = \|(I-Q)P\| = \delta$$

$$\|P-Q\| = \|(I-P)Q\| = 1.$$

§1.2 <u>The idea of a source space</u>

Throughout, $\mathcal{H}$ will denote a separable Hilbert space. It is known that bounded operators on $\mathcal{H}$ admit matrix representations, completely analogous to the well known matrix representations of operators on finite dimensional spaces. We will specify subspaces of $\mathcal{H}$ by their projectors. Having a fixed subspace $P\mathcal{H}$ of $\mathcal{H}$, where P denotes the operator of projection on $P\mathcal{H}$, we will study operators on $\mathcal{H}$ in terms of the orthogonal decomposition of $\mathcal{H}$ into $P\mathcal{H}$ and $(I-P)\mathcal{H}$. To facilitate this idea, we define $E_0 : K(E_0) \to \mathcal{H}$ and $E_1 : K(E_1) \to \mathcal{H}$, where E_0 and E_1 are isometric mappings of some new Hilbert spaces into $\mathcal{H}$, having ranges $R(E_0) = P\mathcal{H}$ and $R(E_1) = (I-P)\mathcal{H}$. Here K( ) stands for the source space of an isometry, R( ) for the range and N( ) for the null space.

Now $E^*_0E_0 = I$, $E_0E_0^* = P$, $R(E_0^*) = K(E_0)$. Since $N(E_0^*) = R(E_0)^\perp = (I-P)\mathcal{H}$, one has $E_0^*E_1 = 0$. Similarly $E_1^*E_1 = I$, $E_1E_1^* = I-P$, $R(E_1^*) = K(E_1)$ and $E_1^*E_0 = 0$. Now every $x \in \mathcal{H}$ can be written as $x = Px + (I-P)x$. If we can write $x_0 = E_0^*x$ and $x_1 = E_1^*x$, then

$$(1.2.1) \qquad x = (E_0 \quad E_1)\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = E_0\, x_0 + E_1\, x_1 .$$

But $\quad (E_0 \quad E_1)\begin{pmatrix} E_0{}^* \\ E_1{}^* \end{pmatrix} x = x \quad$ for any $x \in \mathcal{N}$

and $\quad \begin{pmatrix} E_0{}^* \\ E_1{}^* \end{pmatrix} (E_0 \quad E_1) \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$

Thus $\quad (E_0 \quad E_1)^{-1} = \begin{pmatrix} E_0{}^* \\ E_1{}^* \end{pmatrix}$

The corresponding notation for operators is

$$(1.2.2) \qquad A = (E_0 \quad E_1)\begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix}\begin{pmatrix} E_0{}^* \\ E_1{}^* \end{pmatrix}$$

This equation defines the new operators appearing in it, i.e. $A_{00} = E_0{}^* A E_0$ is an operator from $K(E_0)$ to $K(E_0)$ and $A_{11} = E_1{}^* A E_1$ is an operator from $K(E_1)$ to $K(E_1)$; similarly $A_{01} = E_0{}^* A E_1$ from $K(E_1)$ to $K(E_0)$ and $A_{10} = E_1{}^* A E_0$ from $K(E_0)$ to $K(E_1)$.

If we agree that the sign $\simeq$ is to be read as "is represented by", we can rewrite equations (1.2.1) and (1.2.2) as follows.

$$x \simeq \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \quad A \simeq \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix}$$

$$P \simeq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad I-P = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

The usual rules of matrix multiplication are applicable here. However, the notion of representing operators on $\mathfrak{H}$ by 2×2 block matrices becomes treacherous, because there is more than one way to represent them.

§1.3 Unitary Application of One Subspace Onto Another.

To say that two subspaces are close, we must see how one can be changed to the other by a unitary transformation. The unitaries V in question, will then be those such that

(1.3.1) $VP = QV$,

consequently $V(I-P) = (I-Q)V$.

Thus the dimensions agree:

$$(1.3.2)\begin{cases} \dim P\mathfrak{H} = \dim Q\mathfrak{H}, \\ \dim (I-P)\mathfrak{H} = \dim (I-Q)\mathfrak{H}. \end{cases}$$

In §1.2 we gave a representation of operators in terms of the decomposition by E_0 and E_1. Similarly, when decomposing $\mathfrak{H}$ according to $Q\mathfrak{H}$ and $(I-Q)\mathfrak{H}$, we can define F_0: $K(F_0) \to \mathfrak{H}$, F_1: $K(F_1) \to \mathfrak{H}$. These are isometric mappings of the new Hilbert spaces into $\mathfrak{H}$, with ranges $R(F_0) = Q\mathfrak{H}$ and $R(F_1) =$

$(I-Q)\mathcal{H}$. Here $F_0F_0^* = Q$ and $F_1F_1^* = I-Q$. Henceforth, all operators will be represented in terms of the decomposition by E_0 and E_1, but never in terms of F_0 and F_1:

$$P = (E_0 \quad E_1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix},$$

i.e.
$$P \simeq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$Q = F_0F_0^* = (E_0 \quad E_1) \begin{pmatrix} E_0^*QE_0 & E_0^*QE_1 \\ E_1^*QE_0 & E_1^*QE_1 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix}$$

Thus

$$(1.3.3) \quad Q \simeq \begin{pmatrix} E_0^* \; F_0 \; F_0^* \; E_0 & E_0^* \; F_0 \; F_0^* \; E_1 \\ E_1^* \; F_0 \; F_0^* \; E_0 & E_1^* \; F_0 \; F_0^* \; E_1 \end{pmatrix}$$

Assuming that the dimension conditions (1.3.2) are satisfied, we conclude that there exists a unitary solution of (1.3.1). Actually, (1.3.2) implies the existence of two isometrics W_j, $j = 0,1$ from $K(E_j)$ onto $K(F_j)$, i.e. $W_jW_j^* = W_j^*W_j$. We then define $V = F_0 \, W_0 \, E_0^* + F_1 \, W_1 \, E_1^*$. V satisfies (1.3.1) and $W_j = F_j^* \, V \, E_j$, $j = 0,1$.

Now it follows that any two unitary operators taking $P\mathcal{H}$ onto $Q\mathcal{H}$ will differ only by a unitary transformation within the coordinate subspaces.

Let us name the entries of a unitary solution V of (1.3.1):

$$V \simeq \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix}$$

where

$$\begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix} = \begin{pmatrix} E_0{}^* \\ E_1{}^* \end{pmatrix} V \, (E_0 \quad E_1).$$

On the other hand,

$$\begin{pmatrix} E_0{}^*F_0 & E_0{}^*F_1 \\ E_1{}^*F_0 & E_1{}^*F_1 \end{pmatrix} \begin{pmatrix} W_0 & 0 \\ 0 & W_1 \end{pmatrix} = \begin{pmatrix} E_0{}^* \, V \, E_0 & E_0{}^* \, V \, E_1 \\ E_1{}^* \, V \, E_0 & E_1{}^* \, V \, E_1 \end{pmatrix}$$

Thus

$$(1.3.4) \quad \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix} = \begin{pmatrix} E_0{}^* \\ E_1{}^* \end{pmatrix} V \, (E_0 \quad E_1) = \begin{pmatrix} E_0{}^*F_0 & E_0{}^*F_1 \\ E_1{}^*F_0 & E_1{}^*F_1 \end{pmatrix} \begin{pmatrix} W_0 & 0 \\ 0 & W_1 \end{pmatrix}.$$

Since V is unitary, the relations between the entries are

$$(1.3.5) \quad V^*V \simeq \begin{pmatrix} C_0{}^*C_0 + S_0{}^*S_0 & -C_0{}^*S_1 + S_0{}^*C_1 \\ -S_1{}^*C_0 + C_1{}^*S_0 & S_1{}^*S_1 + C_1{}^*C_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$(1.3.6) \quad VV^* \simeq \begin{pmatrix} C_0C_0{}^* + S_1S_1{}^* & C_0S_0{}^* - S_1C_1{}^* \\ S_0C_0{}^* - C_1S_1{}^* & S_0S_0{}^* + C_1C_1{}^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Note that $C_j = E_j{}^* \, F_j \, W_j$, $j = 0,1$.

Thus $C_jC_j{}^* = E_j{}^* \, F_j \, W_j \, W_j{}^* \, F_j{}^* \, E_j = E_j{}^* \, F_j \, F_j{}^* \, E_j$

and $C_j{}^* \, C_j = W_j{}^* \, F_j{}^* \, E_j \, E_j{}^* \, F_j \, W_j$.

So as W_j varies, $C_jC_j^*$ does not change, while $C_j^*C_j$ changes by a unitary transformation. Similarly, as W_j varies, $S_jS_j^*$ does not change while $S_j^*S_j$ changes by a unitary transformation. This means that, as W_j varies, the singular values of C_j and S_j do not (Appendix B).

We now define

(1.3.7) $\Theta_j = \arc\cos (C_jC_j^*)^{1/2} \geq 0, \ j = 0,1,$

and we define an operator $\Theta \geq 0$ upon $\mathcal{H}$, by

(1.3.8) $$\Theta = (E_0 \quad E_1) \begin{pmatrix} \Theta_0 & 0 \\ 0 & \Theta_1 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix} = E_0\Theta_0E_0^* + E_1\Theta_1E_1^*$$

We take various norms of trigonometric functions of Θ or Θ_j as measures of separation between subspaces $P\mathcal{H}$ and $Q\mathcal{H}$. Note, from the previous discussion, that Θ_j is dependent only on P and Q, and independent of the choice of vectors within the subspaces.

<u>Definition 1.3.1</u> [11]

A unitary solution $V \simeq \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix}$ *of the equation*

$VP = QV$ *is called a "direct rotation" from* $P\mathcal{H}$ *to* $Q\mathcal{H}$, *if it satisfies the following additional conditions:*

(i) $C_0 \geq 0$ *and* $C_1 \geq 0$

(ii) $S_1 = S_0^*$

Definition 1.3.2 [11]

The subspaces $P\mathcal{H}$ and $Q\mathcal{H}$ are said to be in the "acute case", if

$$P\mathcal{H} \cap (I-Q)\mathcal{H} = (I-P)\mathcal{H} \cap Q\mathcal{H} = \{0\}.$$

Throughout, we will assume that relation (1.3.2) is satisfied.

Theorem 1.3.3 [11]

In the acute case, the direct rotation exists and is unique.

Proof:

From (1.3.2) it follows that there exist isometries $W_0\colon K(E_0) \to K(F_0)$ and $W_1\colon K(E_1) \to K(F_1)$. Setting $V = F_0 W_0 E_0^* + F_1 W_1 E_1^*$, V will be unitary and VP = QV. For the operator $C_0\colon K(E_0) \to K(E_0)$, the polar representation (Appendix A) is

$$C_0 = Z_0 (C_0^* C_0)^{1/2} = (C_0 C_0^*)^{1/2} Z_0$$

where Z_0 is a partial isometry uniquely determined from $\overline{R(C_0^*)}$ onto $\overline{R(C_0)}$. We now show that Z_0 is in fact unitary, i.e. $N(C_0)$ and $N(C_0^*)$ should be zero. Let $x_0 \varepsilon N(C_0)$, thus $x \simeq \begin{bmatrix} x_0 \\ 0 \end{bmatrix} \varepsilon P\mathcal{H}$ satisfies $Vx \varepsilon Q\mathcal{H}$, since VP = QV. On the other hand,

$$Vx \simeq \begin{bmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{bmatrix} \begin{bmatrix} x_0 \\ 0 \end{bmatrix} = \begin{bmatrix} C_0 x_0 \\ S_0 x_0 \end{bmatrix} = \begin{bmatrix} 0 \\ S_0 x_0 \end{bmatrix}$$

But $\begin{pmatrix} 0 \\ S_0 x_0 \end{pmatrix} \varepsilon \ (I-P)\mathcal{H}$.

Thus $Vx \varepsilon (I-P)\mathcal{H} \cap Q\mathcal{H} = \{0\}$

This means that the equation $Vx = 0$ implies $x = 0$ and hence $x_0 = 0$. Similarly, if $x_0 \varepsilon N(C_0^*)$, then $x \simeq \begin{pmatrix} x_0 \\ 0 \end{pmatrix} \varepsilon \ P\mathcal{H}$ and

$$V^*x = \begin{pmatrix} C_0^* & S_0^* \\ -S_1^* & C_1^* \end{pmatrix} \begin{pmatrix} x_0 \\ 0 \end{pmatrix} = \begin{pmatrix} C_0^* x_0 \\ -S_1^* x_0 \end{pmatrix} \varepsilon \ (I-P)\mathcal{H} .$$

Thus $x = VV^*x \ \varepsilon \ (I-Q)\mathcal{H}$, since $V(I-P) = (I-Q)V$, and $x \ \varepsilon \ P\mathcal{H} \cap (I-Q)\mathcal{H} = \{0\}$.

This implies that $N(C_0^*) = \{0\}$, and Z_0 is unitary. Similarly, by considering the polar representation of C_1, we get $C_1 = Z_1 (C_1^*C_1)^{1/2} = (C_1C_1^*)^{1/2} Z_1$ where Z_1 is an isometry from $\overline{R(C_1^*)}$ onto $\overline{R(C_1)}$, which is in fact unitary, since both $N(C_1)$ and $N(C_1^*)$ are $\{0\}$.

Let $U = VZ^{-1}$, where

$$Z \simeq \begin{pmatrix} Z_0 & 0 \\ 0 & Z_1 \end{pmatrix}, \quad Z^{-1} = Z^* \simeq \begin{pmatrix} Z_0^* & 0 \\ 0 & Z_1^* \end{pmatrix}$$

It is clear that

(1) U is unitary, since it is the product of two unitaries,

(2) P reduces Z^{-1} i.e. $PZ^{-1} = Z^{-1}P$,

(3) $UP = QU$, since

$$UP = VZ^{-1}P = VPZ^{-1} = QVZ^{-1} = QU.$$

So $U \simeq \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix} \begin{pmatrix} Z_0^* & 0 \\ 0 & Z_1^* \end{pmatrix} = \begin{pmatrix} C_0 Z_0^* & -S_1 Z_1^* \\ S_0 Z_0^* & C_1 Z_1^* \end{pmatrix}$

Then $C_0Z_0^* = (C_0C_0^*)^{1/2} \geq 0$

$C_1Z_1^* = (C_1C_1^*)^{1/2} \geq 0$

Thus, starting from an arbitrary V, we construct U. The uniqueness of U follows from the uniqueness of the polar representation of C_0 and C_1 (Appendix A). To show that $S_1 = S_0^*$, we put V = U,

i.e. $C_j \geq 0$, $Z_j = 1$, $j = 0,1$

From equations (1.3.5) and (1.3.6), we get $S_0^*C_1 = C_0^*S_1$, and this implies that $S_0^*C_1 = C_0S_1$. Similarly, we get $C_0S_0^* = S_1C_1$. Eliminating S_0^* from the last two equations, we get

$$C_0^2S_1 = S_1C_1^2.$$

Now $C_0^4S_1 = C_0^2C_0^2S_1 = C_0^2S_1C_1^2 = S_1C_1^4$. Thus $f(C_0^2)S_1 = S_1\ f(C_1^2)$ where f is any polynomial, hence it is true for any continuous real function f on [0,1]. Thus it is true for the square root function,

i.e. $C_0S_1 = S_1C_1$.

This implies that $S_1C_1 = S_0^*C_1$, which means that S_1 and S_0^* agree on the range of C_1. Since $R(C_1)$ is dense in $K(E_1)$ in the acute case, we finally get $S_1 = S_0^*$.

Theorem 1.3.4 [11]

In the non-acute case, a direct rotation exists, if and only if

(1.3.9) $\dim P\mathcal{H} \cap (I-Q)\mathcal{H} = \dim (I-P)\mathcal{H} \cap Q\mathcal{H}$.

In this case, the existing rotation is not unique.

Proof:

Suppose that (1.3.9) is satisfied, the proof goes similarly to that of the acute case, starting with a unitary

$V \simeq \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix}$, which is a solution of $VP = QV$.

The polar representation of C_0; $C_0 = Z_0(C_0{}^*C_0)^{1/2}$, where Z_0 is a partial isometry from $\overline{R(C_0{}^*)}$ to $\overline{R(C_0)}$. That is, Z_0 is determined except for $N(C_0)$.

We claim that $N(C_0)$ represents $V^{-1}((I-P)\mathcal{H} \cap Q\mathcal{H})$ in the sense described in §1.2. For that, suppose $x_0 \varepsilon N(C_0)$, $x \simeq \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$, so $x \varepsilon P\mathcal{H}$, and $Vx = VPx = QVx$, which implies that $Vx \varepsilon Q\mathcal{H}$; further

$$Vx \simeq \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix} \begin{pmatrix} x_0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ S_0x_0 \end{pmatrix},$$

so that $Vx \varepsilon (I-P)\mathcal{H}$.

Thus $Vx \varepsilon (Q\mathcal{H} \cap (I-P)\mathcal{H})$ and $x \varepsilon V^{-1}(Q\mathcal{H} \cap (I-P)\mathcal{H})$.

On the other hand, suppose that $x \varepsilon V^{-1}((I-P)\mathcal{H} \cap Q\mathcal{H})$, so $Vx \varepsilon (I-P)\mathcal{H} \cap Q\mathcal{H}$.

This means that $Vx \varepsilon Q\mathcal{H}$, i.e. $x \varepsilon V^{-1}Q\mathcal{H}$ or $x \varepsilon P\mathcal{H}$. Thus

$$Vx \simeq \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix} \begin{pmatrix} x_0 \\ 0 \end{pmatrix} = \begin{pmatrix} C_0x_0 \\ S_0x_0 \end{pmatrix},$$

which implies that $C_0x_0 = 0$, since $x\varepsilon(I-P)\mathcal{H}$. Thus $N(C_0)$ represents $V^{-1}((I-P)\mathcal{H} \cap Q\mathcal{H})$. Similarly $N(C_0^*)$ represents $P\mathcal{H} \cap (I-Q)\mathcal{H}$, $N(C_1)$ represents $V^{-1}(P\mathcal{H} \cap (I-Q)\mathcal{H})$ and $N(C_1^*)$ represents $(I-P)\mathcal{H} \cap Q\mathcal{H}$. But, by our assumption,

$$\dim((I-P)\mathcal{H} \cap Q\mathcal{H}) = \dim(P\mathcal{H} \cap (I-Q)\mathcal{H}),$$

so that Z_0 can be extended to a unitary, and it will take $N(C_0)$ onto $N(C_0^*)$. This extension is not unique.

By the same argument, the polar representation of C_1 is $C_1 = Z_1(C_1^*C_1)^{1/2}$, where Z_1: $\overline{R(C_1^*)} \to \overline{R(C_1)}$ is a partial isometry, determined except on $N(C_1)$. Since $\dim N(C_1) = \dim N(C_1^*)$, we can extend Z_1 to unitary, in such a way that the second requirement of the direct rotation will be satisfied. Now, since $N(C_0)$ represents $V^{-1}((I-P)\mathcal{H} \cap Q\mathcal{H})$ and $N(C_1^*)$ represents $(I-P)\mathcal{H} \cap Q\mathcal{H}$ and S_0: $K(E_0) \to K(E_1)$ where $S_0 = E_1^*VE_0$, we have that S_0 maps $N(C_0)$ isometrically onto $N(C_1^*)$. Similarly, we can show that S_1 takes $N(C_1)$ isometrically onto $N(C_0^*)$. Thus, we extend Z_1 by defining it to be $S_0Z_0^{-1}S_1$ on $N(C_1)$, and we claim that $S_0Z_0^{-1}S_1$ maps $N(C_1)$ isometrically onto $N(C_1^*)$. Since S_1 maps $N(C_1)$ isometrically onto $N(C_0^*)$ and Z_0^{-1} takes $N(C_0^*)$ isometrically onto $N(C_0)$, and S_0 takes $N(C_0)$ isometrically onto $N(C_1^*)$, the claim is justified.

As in the previous theorem, let

$$Z = \begin{pmatrix} Z_0 & 0 \\ 0 & Z_1 \end{pmatrix}, \text{ and let } U = VZ^{-1}.$$

As before,

(1) U is unitary,

(2) P reduces Z^{-1},

(3) U satisfies $UP = QU$.

In addition, we have

(4) For $x \varepsilon P\mathcal{H} \cap (I-Q)\mathcal{H}$ or $x \varepsilon (I-P)\mathcal{H} \cap Q\mathcal{H}$, we have $U^2x = -x$.

To prove that, let $x \varepsilon P\mathcal{H} \cap (I-Q)\mathcal{H}$, which represents $N(C_0{}^*)$, so $x \simeq \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$, where $x_0 \varepsilon N(C_0{}^*)$. Since $U^2x = (VZ^{-1})^2\, x$, we get

$$U^2x \simeq \begin{pmatrix} C_0Z_0{}^{-1} & -S_1Z_1{}^{-1} \\ S_0Z_0{}^{-1} & C_1Z_1{}^{-1} \end{pmatrix}^2 \begin{pmatrix} x_0 \\ 0 \end{pmatrix} =$$

$$\begin{pmatrix} (C_0Z_0{}^{-1})^2\, x_0 - S_1Z_1{}^{-1}\, S_0Z_0{}^{-1}\, x_0 \\ S_0Z_0{}^{-1}\, C_0Z_0{}^{-1}\, x_0 + C_1Z_1{}^{-1}S_0Z_0{}^{-1}\, x_0 \end{pmatrix}$$

Since $Z_0{}^{-1}\, x_0 \varepsilon N(C_0)$, then $C_0Z_0{}^{-1}\, x_0 = 0$ and $S_0Z_0{}^{-1}\, C_0Z_0{}^{-1}\, x_0 = 0$ and $S_0Z_0{}^{-1}\, x_0 \varepsilon N(C_1{}^*)$ implies that $C_1Z_1{}^{-1}\, S_0Z_0{}^{-1}\, x_0 = 0$. Now since $Z_1{}^{-1}$ maps $N(C_1{}^*)$ onto $N(C_1)$, and $S_0Z_0{}^{-1}\, x_0 \varepsilon N(C_1{}^*)$, and since $Z_1 = S_0Z_0{}^{-1}S_1$ on $N(C_1)$, we get $Z_1{}^{-1} = S_1{}^*\, Z_0\, S_0{}^*$ on $N(C_1{}^*)$, and $S_1Z_1{}^{-1}S_0Z_0{}^{-1}\, x_0 =$

$S_1S_1{}^*Z_0S_0{}^*S_0Z_0{}^{-1}\ x_0$. From equation(1.3.5), we know that $S_0{}^*S_0 = I$ on $N(C_0)$, and $S_1S_1{}^* = I$ on $N(C_0{}^*)$, so $S_1Z_1{}^{-1}S_0Z_0{}^{-1}\ x_0 = x_0$, and

$$U^2x \simeq \begin{pmatrix}-x_0\\ 0\end{pmatrix} \qquad \text{i.e. } U^2x = -x$$

Similarly, for $x\varepsilon(I-P)\mathcal{H} \cap Q\mathcal{H}$, we get the same result.

It is clear that $U = VZ^{-1}$ satisfies the first condition of the direct rotation. To prove that it satisfies the second condition, we reformulate the question as follows. If V satisfies (4) and $C_0 \geq 0$ and $C_1 \geq 0$, then $S_0{}^* = S_1$. In other words:

$$\text{Let } V = \begin{pmatrix}C_0 & -S_1\\ S_0 & C_1\end{pmatrix}, \quad C_0 \geq 0, \quad C_1 \geq 0$$

and let $V^2x = -x$ for any $x\varepsilon P\mathcal{H} \cap (I-Q)\mathcal{H}$, or $x\varepsilon(I-P)\mathcal{H} \cap Q\mathcal{H}$. Since V is unitary, we have $C_0S_1 = S_0{}^*C_1$ and $C_0S_0{}^* = S_1C_0$, by the previous arguments as in the acute case.

We have $S_1C_1 = S_0{}^*C_1$ which shows that S_1 and $S_0{}^*$ agree on $R(C_1)$. Since $\overline{R(C_1)} = N(C_1)^{\perp}$ and $K(E_1) = \overline{R(C_1)} \oplus N(C_1)$, the proof is complete if we show that $S_1 = S_0{}^*$ on $N(C_1)$.

For that, let $x_1 \ \varepsilon\ R(C_1)^{\perp} = N(C_1)$ so $x = \begin{pmatrix}0\\ x_1\end{pmatrix}$, $x\varepsilon(I-P)\mathcal{H} \cap Q\mathcal{H}$.

Thus $V^2x = -x$,

i.e. $Vx = -V^*x$, $Vx \simeq \begin{pmatrix} -S_1x_1 \\ C_1x_1 \end{pmatrix} = \begin{pmatrix} -S_1x_1 \\ 0 \end{pmatrix}$ and

$$V^*x \simeq \begin{pmatrix} S_0^*\ x_1 \\ C_1\ x_1 \end{pmatrix} = \begin{pmatrix} S_0^*\ x_1 \\ 0 \end{pmatrix}$$

Thus, for any $x_1 \in N(C_1)$, we have $S_1x_1 = S_0^*x_1$, and this means that $S_1 = S_0^*$ on $N(C_1)$.

To prove the converse, suppose that, there exists a direct rotation $U \simeq \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix}$, where $C_0 \geq 0$, $C_1 \geq 0$

and $S_0^* = S_1$. It is required to show that (1.3.9) is satisfied. We have to show that, to every $x \in P\mathcal{H} \cap (I-Q)\mathcal{H}$, $Ux \in (I-P)\mathcal{H} \cap Q\mathcal{H}$, and that for each $Uy \in P\mathcal{H} \cap (I-Q)\mathcal{H}$, $y \in (I-P)\mathcal{H} \cap Q\mathcal{H}$. To do that, let $x \in P\mathcal{H} \cap (I-Q)\mathcal{H}$.

i.e. $x \simeq \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$, where $x_0 \in N(C_0^*) = N(C_0)$, since $C_0^* = C_0$.

Now, since $N(C_0)$ represents $U^{-1}((I-P)\mathcal{H} \cap Q\mathcal{H})$, it follows that $Ux \in (I-P)\mathcal{H} \cap Q\mathcal{H}$. Let $Uy \in (I-P)\mathcal{H} \cap Q\mathcal{H}$, thus $Uy \simeq \begin{pmatrix} 0 \\ x_1 \end{pmatrix}$,

$x_1 \in N(C_1^*) = N(C_1)$. But $y = U^*\ Uy \simeq \begin{pmatrix} C_0 & S_0^* \\ -S_1 & C_1 \end{pmatrix} \begin{pmatrix} 0 \\ x_1 \end{pmatrix} =$

$\begin{pmatrix} S_0^*x_1 \\ C_1x_1 \end{pmatrix} = \begin{pmatrix} S_0^*x_1 \\ 0 \end{pmatrix}$. But $S_0^*x_1 \in N(C_0^*)$, thus $y \in P\mathcal{H} \cap (I-Q)\mathcal{H}$.

Now, unless otherwise stated, we will assume that (1.3.9) is satisfied. Thus the direct rotation will always exist, and rather than with the more general V, we will deal with its direct special case

$$U \simeq \begin{pmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{pmatrix} ; \qquad C_j \geq 0 \qquad j = 0,1,$$

Since UP = QU, we get Q = UPU*,

$$Q \simeq \begin{pmatrix} C_0{}^2 & C_0S_0{}^* \\ S_0C_0 & S_0S_0{}^* \end{pmatrix} .$$

By direct computation, we have

$$(2Q-I)\ (2P-I) \simeq \begin{pmatrix} 2C_0{}^2-1 & 2C_0S_0{}^* \\ 2S_0C_0 & 2S_0S_0{}^*-1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 2C_0{}^2-1 & -2C_0S_0{}^* \\ 2S_1C_0 & 1-2S_0S_0{}^* \end{pmatrix}$$

On the other hand,

$$U^2 = \begin{pmatrix} C_0{}^2 - S_0{}^*S_0 & -C_0S_0{}^* - S_0{}^*C_1 \\ S_0C_0 + C_1S_0 & -S_0S_0{}^* + C_1{}^2 \end{pmatrix}$$

$$= \begin{pmatrix} 2C_0{}^2 - 1 & -2C_0S_0{}^* \\ 2S_0C_0 & 1-2S_0S_0{}^* \end{pmatrix}$$

This follows from (1.3.5) and (1.3.6).

Thus

$$(1.3.10) \qquad U^2 = (2Q - I)\ (2P - I)$$

We remark that any direct rotation of $P\mathcal{H}$ to $Q\mathcal{H}$ is a principal square root of (2Q-I)(2P-I)

i.e. a unitary square root, with spectrum in the right half plane. This is because our constructed direct rotation of $P\mathcal{H}$ to $Q\mathcal{H}$ satisfies (1.3.10)

$$\text{Since } U \simeq \begin{pmatrix} C_0 & -S_0{}^* \\ S_0 & C_1 \end{pmatrix}, \qquad U^* \simeq \begin{pmatrix} C_0 & S_0{}^* \\ -S_0 & C_1 \end{pmatrix},$$

which gives

$$U + U^* \simeq \begin{pmatrix} 2C_0 & 0 \\ 0 & 2C_1 \end{pmatrix} \geq 0.$$

This implies that $\lambda + \bar{\lambda} \geq 0$ for any λ in the spectrum of U. But the spectrum of U lies on the unit circle, and this implies that it lies in the right half plane, (in general, the spectrum of a unitary lies in the right half plane if and only if $U + U^* \geq 0$). So, if P and Q are given, then U^2 is very easy to compute by the above given constructive definition of U. We now relate the operator angle Θ given by (1.3.8) to the direct rotation,

i.e. $\quad \Theta_j = \arc\cos C_j, \quad j = 0,1.$

From (1.3.5) and (1.3.6), it follows that $S_0{}^*S_0 = 1 - C_0{}^2$ and $S_0S_0{}^* = 1 - C_1{}^2$. Since $S_0{}^*S_0$ and $S_0S_0{}^*$ are isometrically equivalent if restricted to the orthogonal complement of their null spaces (Appendix B), it follows that $C_j{}^2$ must be isometrically equivalent except for their eigenspaces belonging to the eigenvalue 1. Since $C_j = \cos \Theta_j$, then the two operators Θ_j, $j = 0,1$ must be isometrically equivalent except perhaps for different dimensionalities of their null spaces. Let $\theta_1 \geq \theta_2 \geq \ldots$ be the singular values of Θ_0, then the nonzero singular values of Θ are the same, but each occuring twice.

i.e. $\theta_1, \theta_1, \theta_2, \theta_2, \ldots$

The polar representation of S_0: $K(E_0) \to K(E_1)$ is $S_0 = J_0 \; (S_0{}^*S_0)^{1/2}$ (where $S_0{}^*S_0 = 1 - C_0{}^2 = 1 - \cos^2 \Theta_0 = \sin^2 \Theta_0$), so that $S_0 = J_0 \sin \Theta_0$,

here J_0 is a partial isometry from $\overline{R((S_0{}^*S_0)^{1/2}}$ onto $\overline{R(S_0)}$,

i.e. From $\overline{R(S_0{}^*)}$ to $\overline{R(S_0)}$.

Since $N(S_0{}^*S_0) = N(\Theta_0)$, one has $\overline{R(S_0{}^*)} = \overline{R(\Theta_0)}$. Similarly, $\overline{R(S_0)} = \overline{R(\Theta_1)}$.

Now $S_0 = (S_0S_0{}^*)^{1/2} J_0$. (Appendix A), and we may write

$$
\begin{aligned}
S_0{}^* &= J_0{}^* \; (S_0S_0{}^*)^{1/2} \\
&= J_0{}^* \sin \Theta_1 \quad ,
\end{aligned}
$$

where $J_0{}^*$ is a partial isometry from $\overline{R(\Theta_1)}$ to $\overline{R(\Theta_0)}$.

If we put

$$J = \begin{pmatrix} 0 & -J_0{}^* \\ J_0 & 0 \end{pmatrix},$$

then J is defined uniquely on $\overline{R(\Theta)}$, and we put J = 0 on $N(\Theta)$.

Since $$U \simeq \begin{pmatrix} C_0 & -S_0{}^* \\ S_0 & C_1 \end{pmatrix} = \begin{pmatrix} \cos\Theta_0 & -J_0^* \sin\Theta_1 \\ J_0 \sin\Theta_0 & \cos\Theta_1 \end{pmatrix}$$

$$U \simeq \begin{pmatrix} \cos\Theta_0 & 0 \\ 0 & \cos\Theta_1 \end{pmatrix} + \begin{pmatrix} 0 & -J_0^* \sin\Theta_1 \\ J_0 \sin\Theta_0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\Theta_0 & 0 \\ 0 & \cos\Theta_1 \end{pmatrix} + \begin{pmatrix} 0 & -J_0{}^* \\ J_0 & 0 \end{pmatrix} \begin{pmatrix} \sin\Theta_0 & 0 \\ 0 & \sin\Theta_1 \end{pmatrix}$$

So $U = \cos\Theta + J\sin\Theta$. Now, it follows that

$$\cos^2\Theta = E_0 C_0{}^2 E_0{}^* + E_1 C_1{}^2 E_1{}^*,$$ while

from (1.3.10) we have

$$\cos^2\Theta = PQP + (I-P)(I-Q)(I-P)$$

$$\sin^2\Theta = P(I-Q)P + (I-P)Q(I-P)$$

$$= (P-Q)^2.$$

So, given P and Q, we know how to construct Θ_j.

§1.4 Unitary Invariants For a Pair of Subspaces

It has been known for many years that two m-dimensional subspaces of real n-dimensional inner product space have m angles as a complete set of unitary invariants [16].

By unitary invariants we mean a set of objects to be assigned to any pair of subspaces $P\mathcal{H}$ and $Q\mathcal{H}$, and such that $P\mathcal{H}$ and $Q\mathcal{H}$ can be carried to $(P\mathcal{H})'$ and $(Q\mathcal{H})'$ by an isometry of $\mathcal{H}$, if and only if the same set of objects was assigned to $(P\mathcal{H})'$ and $(Q\mathcal{H})'$ as to $P\mathcal{H}$ and $Q\mathcal{H}$.

We shall give a complete set of invariants for the subspaces $P\mathcal{H}$ and $Q\mathcal{H}$ in terms of the eigenvalues of Θ_0 and Θ_1 (multiplicity counted).

Theorem 1.4.1 [11]

Consider a pair of subspaces $P\mathcal{H}$ and $Q\mathcal{H}$ subject to $\dim P\mathcal{H} = \dim Q\mathcal{H}$, and

$$\dim[P\mathcal{H} \cap (I-Q)\mathcal{H}] = \dim[(I-P)\mathcal{H} \cap Q\mathcal{H}]$$

and such that $P(I-Q)P$ is compact. A complete system of invariants under isometric equivalence is afforded by the eigenvalues of Θ_0 and Θ_1 (multiplicity is counted). The eigenvalues θ_i, $i = 1,2,\ldots$ of Θ_0 are an arbitrary sequence, satisfying $\frac{\pi}{2} \geq \theta_1 \geq \theta_2 \geq \ldots$ and approaching zero, together with a possible eigenvalue 0. The eigenvalues of Θ_1 must be the same, except perhaps for the multiplicity of 0.

For proof see [11].

It is known [35] that, given two 2-dimensional subspaces $P\mathcal{H}$ and $Q\mathcal{H}$ of 4-dimensional space $\mathcal{H}$, intersecting in a single point 0, then there exist 2-dimensional perpendicular subspaces M_1 and M_2, intersecting at 0, each intersecting $P\mathcal{H}$

and $Q\mathcal{H}$ perpendicularly in a line. The angles θ_i $(0<\theta_i<\pi/2)$, between $M_i \cap P\mathcal{H}$ and $M_i \cap Q\mathcal{H}$ $(i=1,2)$ may have any values independently. These two numbers are determined uniquely, by the Figure of $P\mathcal{H}$ and $Q\mathcal{H}$. This determination is up to a congruence.

The previous theorem shows how this behaviour generalizes to higher dimensions. But in the general case, we have to modify it by the fact that Θ may have a continuous spectrum. (Note that for any normal operator, the residual spectrum is void). Other obvious properties are given by the following theorem, where by $\Omega(\cdot)$, we denote the spectral resolution of Θ, as defined in [33].

Theorem 1.4.2. [11]

Θ commutes with P, with Q, with J and with U described in Section 1.3. For every eigenvalue θ of Θ, the eigenvector x satisfies $\measuredangle(x, Ux) = \theta$. In the acute case, for every eigenvalue θ, the eigenspace $\Omega(\{\theta\})\mathcal{H}$ is the unique maximal subspace, with the following properties:

(a) It reduces P and Q.

(b) For every nonzero vector $x \in P\mathcal{H}$, lying in $\Omega(\{\theta\})\mathcal{H}$, $\measuredangle(x, Qx) = \theta$.

(c) For every nonzero vector x of $(I-P)\mathcal{H}$, lying in $\Omega(\{\theta\})\mathcal{H}$, $\measuredangle(x, (I-Q)x) = \theta$.

Proof. Since $\Theta \simeq \begin{pmatrix} \Theta_0 & 0 \\ 0 & \Theta_1 \end{pmatrix}$, $J = \begin{pmatrix} 0 & -J_0^* \\ J_0 & 0 \end{pmatrix}$

then J commutes with Θ, if and only if

$$\begin{pmatrix} 0 & -J_0{}^* \Theta_1 \\ J_0\Theta_0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\Theta_0 J_0{}^* \\ \Theta_1 J_0 & 0 \end{pmatrix},$$

That is $J_0\Theta_0 = \Theta_1 J_0$.

But $S_0 = J_0 \sin \Theta_0 = \sin \Theta_1 \ J_0$, where J_0 takes $\overline{R(\Theta_0)}$ onto $\overline{R(\Theta_1)}$.

By an argument similar to that given in the proof of theorem (1.3.3), we get $J_0\Theta_0 = \Theta_1 J_0$.

To show that Θ commutes with P, we have

$$\Theta P = (E_0\Theta_0E_0{}^* + E_1\Theta_1E_1{}^*) \ E_0E_0{}^*$$

$$= E_0\Theta_0E_0{}^*,$$

$$P\Theta = E_0E_0{}^* \ (E_0\Theta_0E_0{}^* + E_1\Theta_1E_1{}^*) = E_0\Theta_0E_0{}^*$$

Since $U = \cos \Theta + J \sin \Theta$, and noting that $J^3 = -J$, $J^2\Theta = -\Theta$, and J commutes with Θ we even can write $U = \exp J\Theta$.

Now that U commutes with Θ follows since J commutes with Θ.

$Q = UPU^*$, and both U and P commute with Θ, thus Q also commutes with Θ.

Suppose now that we are in the acute case. Let θ be an eigenvalue of Θ, and x a corresponding eigenvector. i.e. $\Theta x = \theta x$. Now Θ_0 and Θ_1 have the same nonzero eigenvalues, so θ is an eigenvalue of Θ_0 as well as of Θ_1, and $x \simeq \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$ where

x_0 and x_1 are the eigenvectors of Θ_0 and Θ_1 corresponding to θ. Thus $\sphericalangle(x,Ux) = \text{arc cos}\,\frac{\text{Re}(x^*Ux)}{\|x\|\;\|Ux\|} = \text{arc cos}\,\frac{[x^*(U+U^*)x]}{2\|x\|^2} =$

$$\text{arc cos}\,\frac{1}{2}\,\frac{[(x_0{}^*\,E_0{}^* + x_1{}^*\,E_1{}^*)\;2(E_0C_0E_0{}^* + E_1C_1E_1{}^*)\;(E_0x_0 + E_1x_1)]}{\|x\|^2} =$$

$$\text{arc cos}\,\frac{x_0{}^*C_0x_0 + x_1{}^*C_1x_1}{\|x\|^2} = \text{arc cos}\,(\cos\theta) = \theta.$$

Let $\Omega(.)$ be the spectral resolution of Θ. Since P and Q commute with Θ, each member of the spectral resolution of Θ commutes with P and Q, and in particular $\Omega(\{\theta\})$ commutes with P and Q. This proves part (a).

(b) If $x \neq 0$, $x \in P\mathcal{H} \cap \Omega(\{\theta\})\mathcal{H}$, then

$$x \simeq \begin{pmatrix} x_0 \\ 0 \end{pmatrix}, \text{ where } \Theta_0 x_0 = \theta x_0$$

Hence $(\cos\Theta_0)\; x_0 = (\cos\theta)\; x_0$

i.e. $C_0x_0 = (\cos\theta)\; x_0$

From $Q \simeq \begin{pmatrix} C_0{}^2 & C_0S_0{}^* \\ S_0C_0 & S_0S_0{}^* \end{pmatrix}$,

we have

$$\|Qx\|^2 = x^*Qx = x_0{}^*C_0{}^2x_0$$
$$= x_0{}^*\cos^2\theta\; x_0 = \cos^2\theta\;\|x\|^2,$$

i.e. $\|Qx\| = \cos\theta\;\|x\|$, and finally

$$\sphericalangle(x,Qx) = \text{arc cos}\,\frac{x^*\;Qx}{\|x\|\;\|Qx\|} = \text{arc cos}\,\frac{\|Qx\|}{\|x\|} = \theta.$$

(c) For $x \neq 0$, $x\varepsilon(I-P)\mathcal{H} \cap \Omega(\{\theta\})\mathcal{H}$, we have

$$x \simeq \begin{pmatrix} 0 \\ x_1 \end{pmatrix} \quad \text{and } \Theta_1 x_1 = \theta\, x,\ \cos\,\Theta_1\, x = \cos\,\theta\, x_1$$

Similarly as in proving (b), we can show here that $\measuredangle(x,(I-Q)x) = \theta$.

To prove that $\Omega(\{\theta\})$ is a maximal subspace satisfying the properties (a), (b) and (c), we assume that χ is a subspace of $\mathcal{H}$, which is not included in $\Omega(\{\theta\})\mathcal{H}$, and χ satisfies (a). We will show that χ satisfies neither (b) nor (c). Since $\chi \nsubseteq \Omega(\{\theta\})\mathcal{H}$, then there exists $x\varepsilon\chi$, having a nonzero component in $(\Omega(\{\theta\})\mathcal{H})^{\perp}$. Since χ reduces P and Q, and $\cos^2\,\Theta = PQP + (I-P)(I-Q)(I-P)$, then χ will reduce $\cos^2\,\Theta$, and thus reduce every spectral projector $\Omega(.)$ of Θ; in particular it reduces $\Omega(\{\theta\})$ and by our choice of x, we have $x - \Omega(\{\theta\})x\varepsilon\chi$.

The assumption about χ implies that there exist ϕ_1 and ϕ_2, where $\phi_1 \leq \phi_2 < \theta$ or $\theta < \phi_1 \leq \phi_2$, and such that $0 \neq \Omega([\phi_1,\phi_2])x = y\varepsilon\chi$. Now not both Py and (I-P)y are zero, and both are in χ. Since $\Omega(.)$ commutes with P, we can assume that there exists a unit vector $z = \Omega([\phi_1,\phi_2])z\varepsilon P\mathcal{H} \cap \chi$.

Therefore

$$z \simeq \begin{pmatrix} z_0 \\ 0 \end{pmatrix},\quad Qz \simeq \begin{pmatrix} C_0^{\,2} z_0 \\ S_0 C_0 z_0 \end{pmatrix},\quad Uz \simeq \begin{pmatrix} C_0 z_0 \\ S_0 z_0 \end{pmatrix}$$

and it follows that $Qz = (\cos \Theta)\ (Uz)$,

$$z^*Qz = z_0{}^*C_0{}^2z_0 \ \varepsilon\ [\cos^2 \phi_2, \cos^2 \phi_1]$$

$$\|Qz\|^2 = (Uz)^* \cos^2 \Theta\ Uz = z^*\ U^* \cos^2 \Theta\ Uz$$

$$= z^* \cos^2 \Theta\ z\ \varepsilon\ [\cos^2 \phi_2, \cos^2 \phi_1]$$

Now $\cos \sphericalangle(z,Qz) = \frac{z^*Qz}{\|Qz\|} = \frac{\|Qz\|^2}{\|Qz\|} \geq \frac{\cos^2 \phi_1}{\cos^2 \phi_2}$

This is true for any $\phi_1 \leq \phi_2 < \theta$, such that $\Omega([\phi_1,\phi_2])x \neq 0$. We can choose, a fixed ϕ; $\phi_1 \leq \phi \leq \phi_2$, with $\phi_2 - \phi_1$ arbitrarily small, such that $\frac{\cos^2 \phi_2}{\cos^2 \phi_1} > \cos \theta$.

The property (b) is then violated. Similarly, property (c) may be shown violated. This proves the theorem.

Remark.

If the roles of P and Q are interchanged, then the relation

$\cos^2 \Theta = (I-P-Q)^2$ shows that Θ remains the same, while $U^*\ (P,Q) = U\ (Q,P)$.

So $J(P,Q) = -J(Q,P)$.

§1.5 Extremal Properties of the Direct Rotation

In this section, we will study the properties of the direct rotation as introduced in definition 1.3.1. We will assume the hypothesis of theorem 1.4.1 to be satisfied, so that for any unitary, taking $P\mathcal{H}$ onto $Q\mathcal{H}$, the eigenvalues θ_i of Θ_0 and Θ_1 (where $\theta_1 \geq \theta_2 \geq \ldots$) will be invariant. We have already shown that

$$U = \cos\Theta + J\sin\Theta$$
$$U^2 = (2Q - I)\ (2P - I)$$

The first of these equations, gives the relation between Θ and the direct rotation, while the second one tells us how to construct U given P and Q. We should mention that [31,§105], a partial isometry also denoted U, was defined which maps $P\mathcal{H}$ onto $Q\mathcal{H}$. In fact, it coincides with the direct rotation on $P\mathcal{H}$. We refer the reader to [31, §136] for the application of using U in perturbation theory. From theorem 1.3.3 we have (back in our notation)

$$U \simeq \begin{pmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{pmatrix}, \quad V = UZ, \quad Z \simeq \begin{pmatrix} Z_0 & 0 \\ 0 & Z_1 \end{pmatrix}.$$

Remark 1: We have

(1.5.1) $$V \simeq \begin{pmatrix} C_0 Z_0 & -S_0^* Z_1 \\ S_0 Z_0 & C_1 Z_1 \end{pmatrix}$$

Thus $$I - V \simeq \begin{pmatrix} I - C_0 Z_0 & S_0^* Z_1 \\ -S_0 Z_0 & I - C_1 Z_1 \end{pmatrix}$$

and $(I-V)P \simeq \begin{pmatrix} I - C_0Z_0 & 0 \\ - S_0Z_0 & 0 \end{pmatrix}$

hence

$$[(I-V)P]^* \simeq \begin{pmatrix} I - Z_0{}^*C_0 & -Z_0{}^*S_0{}^* \\ 0 & 0 \end{pmatrix}$$

So, the singular values of $(I - V)\big|_{P\mathcal{H}} \equiv (1 - V)P\big|_{P\mathcal{H}}$ are the nonnegative square roots of the eigenvalues of

$$[(I-V)P]^*[(I-V)P] \simeq \begin{pmatrix} I - Z_0{}^*C_0 & -Z_0{}^*S_0^* \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I-C_0Z_0 & 0 \\ -S_0Z_0 & 0 \end{pmatrix}$$

i.e. the eigenvalues of $(1-Z_0{}^*C_0)(1-C_0Z_0) + Z_0{}^*S_0{}^*S_0Z_0$ on $K(E_0)$. Since $C_0{}^2 + S_0{}^*S_0 = I$ on $K(E_0)$, the singular values of $(I-V)\big|_{P\mathcal{H}}$ are the nonnegative square roots of the eigenvalues of $2\,I-C_0Z_0 - Z_0{}^*C_0$ on $K(E_0)$.

<u>Remark 2</u>. Since $\frac{1}{2}\,P(V + V^*)P$ is a Hermitian operator on $P\mathcal{H}$, it has a complete set of eigenvectors. Call them $v_1, v_2, \ldots$

Since $\frac{1}{2}\,P(V+V^*)P = \frac{1}{2}\,E_0(E_0{}^*V\,E_0 + E_0{}^*V^*E_0)\,E_0{}^*$

$$= \frac{1}{2}\,E_0(C_0Z_0 + Z_0{}^*C_0)\,E_0{}^*,$$

the operator $\frac{1}{2}(C_0Z_0 + Z_0{}^*C_0)$ has a complete set of eigenvectors $v_{01}, v_{02}, \ldots$ on $K(E_0)$, such that

$$v_k \simeq \begin{pmatrix} v_{0k} \\ 0 \end{pmatrix}$$

Since for any unit vector $x \varepsilon P\mathcal{H}$, $x \simeq \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$,

and since $\cos \measuredangle (x, Vx) = \operatorname{Re} x^*Vx = \frac{1}{2}x^*P(V+V^*)Px$

$$= \frac{1}{2}x^* (E_0E_0^*)(V+V^*)(E_0E_0^*)x$$

$$= \frac{1}{2}(E_0^*x)^* (C_0Z_0 + Z_0^*C_0)(E_0^*x)$$

$$= \frac{1}{2} x_0^* (C_0Z_0 + Z_0^*C_0)x_0 = x_0^* \frac{1}{2}(C_0Z_0 + Z_0^*C_0)x_0,$$

the vectors v_{0k} are the eigenvectors belonging to the eigenvalues $\cos \phi_1 \le \cos \phi_2 \le \ldots$ of $\frac{1}{2}(C_0Z_0 + Z_0^*C_0)$, where $\phi_k = \measuredangle (v_k, Vv_k)$.

Now, if $U = V$ then $Z_0 = I$ and $\phi_k = \theta_k$ and v_{0k} will be orthonormal eigenvectors u_{0k} of C_0. Now, the eigenvalues of $2I - C_0Z_0 - Z_0^*C_0 = 2[1 - \frac{1}{2}(C_0Z_0 + Z_0^*C_0)]$ are $2(1-\cos \phi_k)$ (Spectral mapping theorem [12]).

Thus, by remark 1, the singular values $\lambda_1 \ge \lambda_2 \ge \ldots$ of $(I-V)_{|P\mathcal{H}}$ are related to ϕ_k by

(1.5.2) $\lambda_k^2 = 2(1 - \cos \phi_k)$,

that is,

$$\lambda_k = 2 \sin \frac{1}{2} \phi_k.$$

Also, if $V = U$, then $\lambda_k = 2 \sin \frac{1}{2} \theta_k$.

Using the above remarks, we can prove the following theorem.

Theorem 1.5.1 [11]. *Given any unitary transformation, which maps* $P\mathfrak{H}$ *onto* $Q\mathfrak{H}$*, then there exists an orthonormal basis* $v_1, v_2, \ldots$ *of* $P\mathfrak{H}$*, such that for all k,* $\Re\,(v_k, Vv_k) \geq \theta_k$.

Proof. By the minimax principle

$$\lambda_k = \inf_{\chi} \sup_{x} \|(I-V)x\|, \tag{1.5.3}$$

where the inf is taken over the (k-1)-dimensional subspace χ of $P\mathfrak{H}$, and the sup is taken over unit vectors $x \in P\mathfrak{H} \ominus \chi$ i.e. those elements of $P\mathfrak{H}$ which are orthogonal to χ.

Fixing χ for which the minimum is attained (this is guaranteed under the hypothesis of Theorem 1.4.1), there is at least one unit vector $x \in P\mathfrak{H} \ominus \chi$, which is a linear combination of the first k eigenvectors

$$u_1 \simeq \begin{pmatrix} u_{01} \\ 0 \end{pmatrix}, \quad u_2 \simeq \begin{pmatrix} u_{02} \\ 0 \end{pmatrix}, \ldots, u_k \simeq \begin{pmatrix} u_{0k} \\ 0 \end{pmatrix}, \text{ of } PUP\,|_{P\mathfrak{H}}.$$

Note that $\frac{1}{2} P(U+U^*)P = PUP$ since $U^* = (2P-I)U(2P-I)$.

Since λ_k is related to ϕ_k by equation (1.5.2), one has

$$\phi_k = \sup_{\substack{y \in P\mathfrak{H} \ominus \chi \\ \|y\| = 1}} \Re\,(y, Vy) \geq \Re\,(x, Vx).$$

$$\phi_k \geq \Re\,(x, Vx), \quad x = \sum_{i=1}^{k} \alpha_i u_i. \tag{1.5.4}$$

Now, it is enough to show that $\Re\,(x,Vx) \geq \theta_k = \Re\,(u_k, Uu_k)$

Suppose $Qx \neq 0$ (otherwise x will be orthogonal to $Q\mathcal{H}$. i.e. x is orthogonal on Vx, and $\measuredangle(x,Vx) = \pi/2$, and by Theorem 1.4.1 it follows that $0 < \theta_k \leq \frac{\pi}{2}$, which gives $\measuredangle(x,Vx) \leq \theta_k$). From (1.5.2), it follows that ϕ_k will be minimized if λ_k is minimized. i.e. if $\|x-y\|$ is minimized where $y \in Q\mathcal{H}$, $\|y\| = 1$. But since $\inf_{\substack{y \in Q\mathcal{H} \\ \|y\| = 1}} \|x-y\|$ is attained at $y = Qx/\|Qx\|$, it follows that $\|x - Vx\| \geq \|x - \frac{Qx}{\|Qx\|}\|$. This implies that

$$(1.5.5) \qquad \measuredangle(x,Vx) \geq \measuredangle(x,Qx)$$

We now relate the right hand side of the above inequality to θ_k; this will depend upon our particular choice of x:

$$\cos \measuredangle(x,Qx) = \operatorname{Re} \frac{x^*Qx}{\|Qx\|} = \frac{x^*Qx}{\|Qx\|} = (x^*Qx)^{1/2}.$$

Since $x \in P\mathcal{H}$, then $x = E_0x_0$, and $\cos \measuredangle(x,Qx) = (x_0{}^*E_0{}^*Q\, E_0x_0)^{1/2}$.

But $$Q \simeq \begin{pmatrix} C_0{}^2 & C_0S_0{}^* \\ S_0C_0 & S_0S_0{}^* \end{pmatrix}$$

Thus $$\cos \measuredangle(x,Qx) = (x_0{}^*C_0{}^2x_0)^{1/2}$$

Since $u_1,\ldots,u_k$ are the eigenvectors of $PUP|_{P\mathcal{H}}$, corresponding to the eigenvalues $\cos\theta_1 \leq \ldots \leq \cos\theta_k$, and since $x \in [u_1, u_2, \ldots, u_k]$, thus $x_0 \in [u_{01}, u_{02}, \ldots, u_{0k}]$ where $u_{01}, \ldots, u_{0k}$ are the eigenvectors of C_0 corresponding to the eigenvalues $\cos\theta_1 \le \cos\theta_2 \le \ldots \le \cos\theta_k$. Since $x_0 = \sum_{j=0}^{k} \xi_j u_{0j}$, where $\sum_{j=1}^{k} |\xi_j|^2 = \|x_0\|^2 = 1$, then

$$\cos \measuredangle (x, Qx) = \left(\sum_{j=1}^{k} |\xi_j|^2 \cos^2 \theta_j\right)^{1/2} \le \left[\left(\sum_{j=1}^{k} |\xi_j|^2\right)\cos^2\theta_k\right]^{1/2}$$

$= \cos\theta_k$.

Combining (1.5.4) and (1.5.5) with the last inequality, we get $\phi_k \ge \theta_k$ for any k, and this means that there exists an orthonormal system which is efficiently moved by U, or equivalently the singular values of $(I-V)|_{P\mathcal{H}}$ are minimized when V = U, or by observing from (1.5.3) that λ_k is the minimax value of the distance a unit vector in $P\mathcal{H}$ is moved by V. This distance is minimized when V = U.

Corollary. *For every unitary invariant norm, $\|(I-V)P\|$ is minimized when $V = U$.*

Proof. Since for every unitary invariant norm, $\|(I-V)P\|$ is a monotone function of the nonzero singular values of $(I-V)P$, and by the previous theorem, the singular value λ_k of $(I-V)P$ is minimized when V = U. The corollary follows.

Theorem 1.5.2. *Given any unitary operator V which maps $P\mathcal{H}$ onto $Q\mathcal{H}$ and given any orthonormal basis $\{v_1, v_2, \ldots\}$ of $P\mathcal{H}$, we have*

$$\sum_{k=1}^{\infty} \sin^2 \measuredangle (v_k, Vv_k) \geq \sum_{k=1}^{\infty} \sin^2 \theta_k \tag{1.5.6}$$

Proof. Since $v_k \in P\mathcal{H}$, $v_k \simeq \begin{pmatrix} v_{0k} \\ 0 \end{pmatrix}$ we have

$$\sum_k \sin^2 \measuredangle (v_k, Vv_k) = \sum_k [1-\cos^2 \measuredangle (v_k, Vv_k)]$$

$$= \sum_k [1 - (\text{Re } v_k{}^*Vv_k)^2] = \sum_k [1-(\text{Re}v_{0k}{}^*E_0{}^*VE_0v_{0k})^2]$$

$$\geq \sum_k [1-|v_{0k}{}^* \, C_0Z_0 \, v_{0k}|^2].$$

Now, since $|v_{0k}{}^* \, C_0Z_0v_{0k}|^2 \leq \sum_\ell |v_{0k}{}^* \, C_0Z_0v_{0\ell}|^2$ it follows that $\sum_k \sin^2 \measuredangle (v_k, Vv_k) \geq \sum_k [1 - \sum_\ell | \, v_{0k}{}^* \, C_0Z_0v_{0\ell}|^2]$.

But $\sum_\ell |v_{0k}{}^* \, C_0Z_0v_{0\ell}|^2 = \|Z_0{}^*C_0v_{0k}\|^2 = \|C_0v_{0k}\|^2$, thus

$$\sum_k \sin^2 \measuredangle (v_k, Vv_k) \geq \sum_k [1-v_{0k}{}^*C_0{}^2v_{0k}] = \sum_k [v_{0k}{}^*(1-C_0{}^2)v_{0k}].$$

But from $I-C_0{}^2 = S_0{}^*S_0$ on $K(E_0)$, it follows that

$$\sum_k \sin^2 \measuredangle (v_k, Vv_k) \geq \sum_k v_{0k}{}^*S_0{}^*S_0v_{0k} = \text{tr } S_0{}^*S_0$$

$$= \sum_k (\text{eigenvalues of } S_0{}^*S_0) = \sum_k (\text{singular values of } \sin \Theta_0)^2 = \sum_k \sin^2 \theta_k.$$

We observe that in case V = U (so that Z = I) we obtain equality in (1.5.6) by choosing the orthonormal basis $u_1, u_2, \ldots$ of $P\mathcal{H}$ to be the eigenvectors of $\Theta|_{P\mathcal{H}}$ corresponding to the eigenvalues $\theta_1 \geq \theta_2 \geq \ldots$; in this case, $\measuredangle (u_k, U u_k) = \theta_k$ by Theorem (1.4.2).

Remark. In theorem 1.5.1, we explained that, if u_{0k} are the orthonormal eigenvectors of Θ_0, then $u_k = \begin{pmatrix} u_{0k} \\ 0 \end{pmatrix}$ are

the eigenvectors of $(I-U^*)(I-U)_{|P\mathcal{H}}$, corresponding to the eigenvalues $\lambda_j = 2 \sin \frac{1}{2} \theta_j$. But from theorem 1.4.2, we know that J commutes with Θ, that $J_0 \Theta_0 = \Theta_1 J_0$, and that Θ_1 has the same nonzero eigenvalues as Θ_0. Since $\Theta_0 u_{01} = \theta_1 u_{01}$, then $J_0 \Theta_0 u_{01} = \theta_1 J_0 u_{01}$, and thus $\Theta_1 J_0 u_{01}$ $\theta_1 J_0 u_{01}$. This means that $J_0 u_{01}$ is the eigenvector of Θ_1 corresponding to θ_1. But $Ju \simeq \begin{pmatrix} 0 & -J_0^* \\ J_0 & 0 \end{pmatrix} \begin{pmatrix} u_{01} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ J_0 u_{01} \end{pmatrix}$,

so, the eigenvalues of $(I-U^*)(I-U)$ will be $2 \sin \frac{1}{2} \theta_1$, $2 \sin \frac{1}{2} \theta_1$, $2 \sin \frac{1}{2} \theta_2$, $2 \sin \frac{1}{2} \theta_2$, ... and the corresponding eigenvectors are u_1, Ju_1, u_2, Ju_2, ...

Theorem 1.5.3 [11]. *For every unitary invariant norm,* $\|(I-V^*)(I-V)\|$ *is minimized when* $V = U$.

Proof. It is enough to show that

$$\|(I-V^*)(I-V)\|_\nu \geq \|(I-U^*)(I-U)\|_\nu; \quad \nu = 1,2,\ldots \quad \text{(Appendix B)}$$

For the compact operator A, we have equivalently

$$(1.5.7) \qquad \|A\|_\nu = \sup_{\Omega} \|A\Omega\|_\nu$$

Where Ω is the projector onto the ν-dimensional subspace of $\mathcal{H}$, or

$$(1.5.8) \qquad \|A\|_\nu = \sup_{\Omega, T} \|\Omega K T\|_\nu$$

Where Ω and T are projectors onto ν-dimensional subspaces i.e. over pair of ν-projectors.

Thus

$$(1.5.9)\begin{cases} \|(I-V^*)(I-V)\|_\nu \geq \sum_{k=1}^{\nu/2} \|\Omega_k(I-V^*)(I-V)\Omega_k\|_2; & \nu \text{ even} \\ \|(I-V^*)(I-V)\|_\nu \geq \sum_{k=1}^{[\nu/2]} \|\Omega_k(I-V^*)(I-V)\Omega_k\|_2 + & \end{cases}$$

$$\|\Omega_{\frac{\nu+1}{2}}(I-V^*)(I-V)\Omega_{\frac{\nu+1}{2}}\|_1; \quad \nu \text{ odd.}$$

Here $\Omega\mathcal{H} = [x_1, x_2, \ldots, x_\nu]$ where x_1 and x_2 lie in $\Omega_1\mathcal{H}$, x_3 and x_4 lie in $\Omega_2\mathcal{H}, \ldots$ where $\Omega_k\mathcal{H} = [u_k Ju_k]$. Thus, it is sufficient to prove that $\|\Omega_k(I-V^*)(I-V)\Omega_k\|_2$ and $\|\Omega_{\frac{\nu+1}{2}}(I-V^*)(I-V)\Omega_{\frac{\nu+1}{2}}\|_1$ are minimized when $V = U$. Let $\Omega = \Omega_k$, $\theta = \theta_k$ and $u = u_k$. Since u_{0k} is the eigenvector of Θ_0 corresponding to θ_k, Ju_{0k} will be the eigenvector of Θ_1 corresponding to the eigenvalue θ_k.

Thus $Uu = (\cos\Theta + J\sin\Theta)\,u = \cos\theta\, u + J\sin\theta\, u =$

$$\cos\theta\, u + \sin\theta\, Ju,$$

And $UJu = (\cos\Theta + J\sin\Theta)\,Ju = -\sin\theta\, u + \cos\theta\, Ju$,

since J commutes with Θ and $J^2\,\Theta = -\Theta$.

Since $Vu \in Q\mathcal{H}$, $V = (UZU^{-1})U$ and UZU^{-1} maps $Q\mathcal{H}$ into $Q\mathcal{H}$ and $(I-Q)\mathcal{H}$ into $(I-Q)\mathcal{H}$. Thus we can write

$$Vu = a_0\ Uu + b_0 w;\ w \in Q\mathcal{H} \ominus [Uu],\ \|w\| = 1,$$

$$|a_0|^2 + |b_0|^2 = 1.$$

$$VJu = a_1 UJu + b_1 x;\ x \in (I-Q)\mathcal{H} \ominus [JUu],\ \|x\| = 1$$

$$|a_1|^2 + |b_1|^2 = 1.$$

Since Ω commutes with Q, then $\Omega w = \Omega x = 0$.

We consider operators, reduced by the 2-dimensional subspace $\Omega\mathcal{H}$, and which are zero on the orthonormal complement. We represent the part of such an operator in $\Omega\mathcal{H}$ by its 2×2 matrix relative to the basis (u, Ju).

So $\Omega V\Omega = \{\alpha_{ij}\}$, where $\alpha_{11} = (\Omega V\Omega u, u) = (Vu, u)$ and $Vu = a_0 \cos\theta\ u + a_0 \sin\theta\ Ju + b_0 w$.

Thus $\alpha_{11} = a_0 \cos\theta$

and $\alpha_{12} = (\Omega V\Omega Ju, u) = -a_1 \sin\theta,$

$\alpha_{21} = (\Omega V\Omega u, Ju) = a_0 \sin\theta,$

$\alpha_{22} = (\Omega V\Omega Ju, Ju) = a_1 \cos\theta.$

In matrix representation

$$\Omega V\Omega:\ \begin{pmatrix} a_0 \cos\theta & -a_1 \sin\theta \\ a_0 \sin\theta & a_1 \cos\theta \end{pmatrix}.$$

The eigenvalues of $\Omega(I-V^*)(I-V)\Omega$ are μ_1^2 and μ_2^2, where μ_1 and μ_2 are the singular values of $(I-V)\Omega$. Hence, the eigenvalues of $\frac{1}{2}\Omega(V+V^*)\Omega$ are $1 - \frac{1}{2}\mu_1^2$ and $1 - \frac{1}{2}\mu_2^2$ (since $\Omega(I-V^*)(I-V)\Omega = \Omega(2I-V^* - V)\Omega)$.

Thus $\frac{1}{2}\Omega(V+V^*)\Omega$: $\frac{1}{2}\begin{pmatrix} a_0 \cos\theta & -a_1 \sin\theta \\ a_0 \sin\theta & a_1 \cos\theta \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \bar{a}_0 \cos\theta & \bar{a}_0 \sin\theta \\ -\bar{a}_1 \sin\theta & \bar{a}_1 \cos\theta \end{pmatrix}$

$$= \begin{pmatrix} (\mathrm{Re}\ a_0)\cos\theta & \frac{1}{2}(\bar{a}_0 - a_1)\sin\theta \\ \frac{1}{2}(a_0 - \bar{a}_1)\sin\theta & (\mathrm{Re}\ a_1)\cos\theta \end{pmatrix}$$

The calculated eigenvalues of $\frac{1}{2}\Omega(V+V^*)\Omega$, from the above matrix are

$$1 - \frac{1}{2}\mu_1^2 = c\cos\theta - \sqrt{d^2 + e^2\sin^2\theta},$$
(1.5.10)
$$1 - \frac{1}{2}\mu_2^2 = c\cos\theta + \sqrt{d^2 + e^2\sin\theta}$$

where c,d,e and f are real constants, defined by $a_0 + a_1 = 2c + 2ie$, $a_0 - a_1 = 2d - 2if$.

Since $|a_j^2| \le 1$, we have $(c+d)^2 + (e-f)^2 \le 1$ and $(c-d)^2 + (e+f)^2 \le 1$, so that $c^2 + d^2 + e^2 + f^2 \le 1$

But $\|\Omega(I-V^*)(I-V)\Omega\|_1 = \mu_1^2 \ge 2 - 2c\cos\theta \ge 2 - 2\cos\theta$, $(c \le 1)$ and since $\|(I-V)\Omega\|_1^2 = \|\Omega(I-V^*)(I-V)\Omega\|_1 = \mu_1^2 \ge 2 - 2\cos\theta = \|(I-U)\Omega\|_1^2$, then

$$\|\Omega(I-V^*)(I-V)\Omega\|_1 \ge \|\Omega(I-U^*)(I-U)\Omega\|_1$$
$$\|(I-V)\Omega\|_1 \ge \|(I-U)\Omega\|_1$$

But $\|\Omega(I-V^*)(I-V)\Omega\|_2 = \mu_1^2 + \mu_2^2$

And from (1.5.9), $\mu_1^2 + \mu_2^2 = 4\,(1 - c \cos\theta)$.

The right hand side will be minimized when $c = 1$, i.e. $e = d = f = 0$, which reads in original terms $V = U$.

So $\|(I-V^*)(I-V)\|_\nu \geq \|(I-U^*)(I-U)\|_\nu$ for any ν. Thus $\|(I-V^*)(I-V)\|$ is minimized when $V = U$.

From the proof, we also get $\|(I-V)\|_1 \geq \|(I-U)\|_1$, since $\|(I-V^*)(I-V)\|_1 = \|(I-V)\|^2_1$, and $\|(I-U^*)(I-U)\|_1 = \|I-V\|^2_1$. This conclusion is true for the bound norm, and for the square norm, but is not valid for other ν-norms and we will provide an example for the last case.

For the square norm

$$\|I-V\|^2_{sq} = tr[(I-V^*)(I-V)] = tr[P(I-V^*)(I-V)P] + tr[(I-P)(I-V^*)(I-V)(I-P)]$$

From the corollary to theorem (1.5.1), the right hand side will be minimized when $V = U$, thus

$$\|I-V\|^2_{sq} \geq \|I-U\|^2_{sq} .$$

<u>Example</u>.

Take $V = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$

as the unitary operator, taking $P\mathcal{H}$ to $Q\mathcal{H}$. The eigenvalues of V are 1 and -1. So the singular value of $(I-V)$

are 2 and 0.

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

The singular values of (I-U), are the positive square roots of the eigenvalues of $(I-U^*)(I-U) = 2 - U - U^*$, i.e. the square roots of $2 - 2\cos\theta$ and $2 - 2\cos\theta$. Thus the singular values of (I-U) are $2\sin\frac{\theta}{2}$ and $2\sin\frac{\theta}{2}$.

$$\|I-U\|_2 = 4\sin\theta/2, \quad \|I-V\|_2 = 2$$

So $\|I-V\|_2 \geq \|I-U\|_2$ if and only if $\theta \leq \pi/3$.

We conclude this chapter by quoting a positive result in this direction. We refer the reader to [11] for the proof.

Theorem 1.5.4 [11]. *Assume V is a unitary operator, taking $P\mathcal{H}$ onto $Q\mathcal{H}$ in a real space $\mathcal{H}$. Assume also that $\theta \leq \pi/3$. Then $\|I-V\|$ is minimized, for every unitary invariant when $V = U$.*

The previous example shows that if $\theta > \frac{\pi}{3}$, then the conclusion of the theorem fails.

CHAPTER 2

The Operator Equation BX - XA = Q

We consider a Banach algebra $\mathcal{B}$, with two particular elements A and B . T is an operator on $\mathcal{B}$, such that T(X) = BX - XA for every X in $\mathcal{B}$.

§2.1 The Matrix Equation BX-XA = Q in the Banach Algebra of n×n Matrices

Definition 2.1.1. *If $A = (a_{ij})$ is an m×n matrix and B is an s×t matrix, the ms×nt Kronecker product $A \otimes B$ is defined as the block matrix*

$$A \otimes B = (a_{ij}\, B).$$

One of the most important properties of this product is that it enables us to convert matrices into column vectors.

Definition 2.1.2. *If A_j denotes the jth column of an m×n matrix A, the mn vector vec A is then defined as*

$$vec\ A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$$

Theorem 2.1.3 [30]. *Let A be an m×n matrix, and B be an n×p matrix, then*

$$vec\ AB = (I_p \otimes A)\ vec\ B = (B' \otimes I_m)\ vec\ A$$

where B' is the transpose of B.

We now state the standard properties of Kronecker products. The proofs of these properties are given in [3].

1. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$,

2. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$,

3. $(A+B) \otimes (C+D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D$.

4. If A has eigenvalues α_i, $i = 1,2,\ldots,m$ and B has eigenvalues β_j, $j = 1,2,\ldots,s$, then $A \otimes B$ has eigenvalues $\alpha_i\beta_j$. Further, $I_s \otimes A + B \otimes I_m$ has eigenvalues $\alpha_i + \beta_j$.

The matrices involved here have the appropriate orders. In Property 4, it is assumed that A and B are square matrices of orders m and s respectively.

Theorem 2.1.4. *Let A,B,X and Q be square matrices of order n. Then a necessary and sufficient condition for the equation $BX - XA = Q$ to have a unique solution is that the eigenvalues of A are distinct from the eigenvalues of B.*

Proof. $BX - XA = Q$ can be written as follows:

$$\text{vec } BX - \text{vec } XA = \text{vec } Q.$$

Using theorem 2.1.3 we get

$$(I_n \otimes B)\text{vec } X - (A' \otimes I_n)\text{vec } X = \text{vec } Q,$$

that is,

$$[(I_n \otimes B) - (A' \otimes I_n)] \text{ vec } X = \text{vec } Q.$$

This equation has the unique solution

$$\text{vec } X = [(I_n \otimes B) - (A' \otimes I_n)]^{-1} \text{ vec } Q$$

if and only if the matrix $(I_n \otimes B) - (A' \otimes I_n)$ is nonsingular. Using property 4. of the Kronecker products, we conclude that $\beta_j - \alpha_i \neq 0$ is a necessary and sufficient condition for the equation $BX - XA = Q$ to have a unique solution.

Remarks.

1. Definition 2.1.2 and Theorem 2.1.3 can be applied to a more general class of linear matrix equation [30].
2. The theorem may be restated as follows: For the operator T on $\mathcal{B}$ defined by $T(X) = BX - XA$, this operator is invertible if and only if the eigenvalues of A and B are distinct. The solution X may be derived using definition 2.1.2; note that $\sigma(T) = \sigma(B) - \sigma(A)$. This follows from property 4. of Kronecker products.
3. The equation $BX - XA = 0$ has a non-zero solution if and only if $\alpha_i - \beta_j = 0$ for some i and j.

§2.2 The operator equation $BX - XA = Q$ where $\mathcal{B}$ is the space of bounded operators on a Hilbert space.

Theorem 2.2.1. [20]. *If there exist real numbers a and b such that $a>b$, $B+B^* \leq b$ and $A+A^* \geq a$, then the operator T^{-1} exists as a bounded operator and has the representation*

$$T^{-1}(Q) = -\int_0^{\infty} e^{Bt}\, Q\, e^{-AT}\, dt. \tag{2.2.1}$$

By $\rho(A)$ we denote the resolvent set of an element A of the Banach algebra $\mathcal{B}$, i.e. the set of all complex numbers z such that $(zI - A)^{-1}$ is in $\mathcal{B}$, while $\sigma(A)$, the complement of

$\rho(A)$ is the spectrum of A.

Definition 2.2.2. *A set D in the complex plane is said to be a Cauchy domain, if the following conditions are satisfied.*

1. *D is bounded and open*
2. *D has a finite number of components, the closures of any two of which are disjoint.*
3. *The boundary of D is composed of a finite number of closed rectifiable jordan curves, no two of which intersect. We denote the positively oriented boundary of D by b(D).*

Theorem 2.2.3 [39]. *Let F be closed, and G a bounded open subset of the complex plane, such that $F \subset G$. Then there exists a Cauchy domain D such that $F \subset D$ and $\bar{D} \subset G$.*

Definition 2.2.4 [39]. *Let f be a complex valued function, holomorphic in a bounded region G which includes $\sigma(T)$, the spectrum of the operator T. The function f(T) of the operator T is defined by*

$$(2.2.2) \qquad f(T) = \frac{-1}{2\pi i} \int_{b(D')} f(w)\ (T - wI)^{-1}\, dw,$$

where D' is a Cauchy domain, such that $\sigma(T) \subset D' \subset G$.

Theorem 2.2.5. [32]. *If $w \notin \sigma(B) - \sigma(A)$, then*

(1) $w \in \rho(T)$,

(2) There exists a Cauchy domain D such that $\sigma(A) \subset D$ and $\sigma(B - wI) \cap \bar{D} = \phi$,

(3) For any Cauchy domain D, which satisfies (2),

$$(T-wI)^{-1} Q = \frac{1}{2\pi i} \int_{b(D)} (B-wI - zI)^{-1} Q (zI - A)^{-1} dz.$$

<u>Proof</u>. Since $w \notin \sigma(B) - \sigma(A)$, then $\sigma(B) - w \cap \sigma(A) = \phi$.

Since $\sigma(B) - w = \sigma(B-wI)$, then $\sigma(B-wI) \cap \sigma(A) = \phi$.
But A and B are bounded operators, thus $\sigma(B-wI)$ and $\sigma(A)$ are compact disjoint sets

i.e. there exists a bounded open set G containing $\sigma(A)$ and disjoint from $\sigma(B-wI)$.

From theorem 2.2.3, it follows that there exists a Cauchy domain D, such that $\sigma(A) \subset D$, and $\bar{D} \subset G$. Thus $\sigma(B-wI) \cap \bar{D} = \phi$.

Now suppose X is a solution of the operator equation

$$(T-wI)X \equiv BX-XA - wX = Q.$$

If $z \in b(D)$, then $z \in \rho(A)$ since $\sigma(A) \subset D$. Also $z \in \rho(B-wI)$ since $\sigma(B-wI) \cap \bar{D} = \phi$. Next, $z \in b(D)$ implies that $z \in \rho(A) \cap \rho(B-wI)$, and

$$(T-wI)X = (B-wI-zI)X + X(zI-A) = Q.$$

Since $(zI-A)^{-1}$ and $(B-wI-zI)^{-1}$ exist, then

$$X(zI-A)^{-1} + (B-wI-zI)^{-1}X = (B-wI-zI)^{-1} Q (zI-A)^{-1}$$

i.e. $$\frac{1}{2\pi i} \int_{b(D)} X(zI-A)^{-1} dz + \frac{1}{2\pi i} \int_{b(D)} (B-wI-zI)^{-1}X dz =$$

$$\frac{1}{2\pi i} \int_{b(D)} (B-wI-zI)^{-1} Q (zI-A)^{-1} dz.$$

Now, from equation (2.2.2), it follows that

$$\frac{1}{2\pi i} \int_{b(D)} X(zI-A)^{-1}\, dz = X \left\{ \frac{1}{2\pi i} \int_{b(D)} (zI-A)^{-1}\, dz \right\} = X.$$

Since $\rho(B-wI) \supset \bar{D}$ and $(B-wI-zI)^{-1}$ is an analytic vector valued function on $\rho(B-wI)$, then

$$\left[\frac{1}{2\pi i} \int_{b(D)} (B-wI-zI)^{-1}\, dz\right] X = 0.$$

By Cauchy's theorem [32], it follows that

$$X = \frac{1}{2\pi i} \int_{b(D)} (B-wI-zI)^{-1}\, Q\, (zI-A)^{-1}\, dz,$$

and the proof is complete.

In an analogous way, we can obtain the following.

Theorem 2.2.6 [32]. *If $w \notin \sigma(B) - \sigma(A)$, then*

$$(T-wI)^{-1}Q = \frac{1}{2\pi i} \int_{b(D_1)} (B-zI)^{-1}\, Q\, (A+wI-zI)^{-1}\, dz$$

For any Cauchy domain D_1, such that $\sigma(B) \subset D_1$, $\sigma(A+wI) \cap \bar{D}_1 = \phi$.

Corollary 2.2.7.

(1) $\sigma(T) \subset \sigma(B) - \sigma(A)$

(2) If $\sigma(B) \cap \sigma(A) = \phi$, then T^{-1} exists as a bounded operator, and this generalizes the results of §2.1.

Proof.

(1) Follows from theorem 2.2.5.

(2) Follows from theorem 2.2.5, part (3) by putting $w = 0$.

G. Lumer and M. Rosenblum [27] proved the following unpublished theorem of D.C. Kleinceke, and generalized it.

Theorem 2.2.8. *Given A and B from $\mathcal{B}$, where is the Banach algebra of all bounded operators on $\mathcal{B}$. Let T be defined on $\mathcal{B}$, by $T(X) = BX - XA$, then*

$$\sigma(T) = \sigma(B) - \sigma(A)$$

We now present an operational calculus for T in terms of elements of $\mathcal{B}$. For this we need the following lemma.

Lemma 2.2.9 [32]. *Let G be a bounded open set containing $\sigma(B) - \sigma(A)$. Then there exist Cauchy domains D and D', such that $\sigma(B) - \sigma(A) \subset D'$ and $\sigma(A) \subset D$. Furthermore:*

(1) If $w \in b(D')$, then $w \notin \sigma(B)-\sigma(A)$, and $\sigma(B-wI) \cap \bar{D} = \phi$.

(2) If $z \in b(D)$, then $\sigma(B-zI) \subset D'$.

Theorem 2.2.10. [32]. *If $f(z)$ is a complex valued function, holomorphic, in a region which include $\sigma(B) - \sigma(A)$, then*

$$(2.2.3)\quad f(T)Q = \frac{1}{2\pi i} \int_{b(D)} f(B-zI)\, Q\, (zI-A)^{-1}\, dz$$

where D is as in lemma 2.2.9.

Proof. Since $\sigma(T) = \sigma(B) - \sigma(A) \subset D'$, where D' is defined as in lemma 2.2.9, then using equation (2.2.2), we get

$$f(T)Q = \frac{-1}{2\pi i} \int_{b(D')} f(w) \ (T-wI)^{-1} \ Q \ dw.$$

By theorem 2.2.5, we have

$$f(T)Q = \frac{-1}{2\pi i} \int_{b(D')} f(w) \left\{ \frac{1}{2\pi i} \int_{b(D)} (B-wI-zI)^{-1} \ Q \ (zI-A)^{-1} \ dz \right\} dw$$

Interchanging the order of integration, we get

$$f(T)Q = \frac{-1}{2\pi i} \int_{b(D)} [\frac{1}{2\pi i} \int_{b(D')} f(w) \ (B-wI-zI)^{-1} \ dw] \ Q \ (zI-A)^{-1} \ dz$$

Now, from lemma 2.2.9, it follows that $\sigma(B-zI) \subset D'$ for $z \varepsilon b(D)$, and thus

$$\frac{1}{2\pi i} \int_{b(D')} f(w) \ (B-wI-zI)^{-1} \ dw = -f(B-zI)$$

and

$$f(T)Q = \frac{1}{2\pi i} \int_{b(D)} f(B-zI) \ Q \ (zI-A)^{-1} \ dz.$$

This proves the theorem.

We can similarly prove that if $f(z)$ is a complex-valued function, holomorphic in a region G that includes $\sigma(B) - \sigma(A)$, then

$$f(T)Q = \frac{1}{2\pi i} \int_{b(D')} (zI-B)^{-1} \ Q \ f(zI-A) \ dz,$$

where D' is a certain Cauchy domain that contains $\sigma(B)$.

Theorem 2.2.11. [20]. *Let $\mathcal{B}$, A, B, a, b, be as in theorem 2.2.1. Then T^{-1} exists, and is defined everywhere in $\mathcal{B}$, and*

$$T^{-1}(Q) = -\int_0^\infty e^{tT}(Q)\,dt = -\int_0^\infty e^{Bt}\,Q\,e^{-At}\,dt.$$

Proof. Let $f(z) = e^{tz}$ in theorem 2.2.10, thus

$$f(T)\ (Q) = e^{tT}\ (Q) = \frac{1}{2\pi i}\int_{b(D)} e^{t(B-zI)}\ Q\ (zI-A)^{-1}\,dz =$$

$$= \frac{1}{2\pi i}\ e^{Bt}\ Q\int_{b(D)} e^{-zt}\ (zI-A)^{-1}\,dz$$

But $e^{-At} = \frac{1}{2\pi i}\int_{b(D)} e^{-zt}\ (zI-A)^{-1}\,dz$.

Thus $e^{tT}\ (Q) = e^{Bt}\ Q\ e^{-At}$.

Let $B_1 = \frac{1}{2}(B+B^*)$ and $B_2 = \frac{1}{2i}(B-B^*)$. B_1 and B_2 are hermitian operators and $B = B_1 + iB_2$. Since $B+B^* \leq b$, we have $B_1 \leq \frac{1}{2}b$. Using $e^B = I + \sum_{k=1}^{\infty} \frac{B^k}{k!}$ we conclude that there exists a number $m>0$ such that for every positive integer n,

$$\|e^{\frac{1}{n}(B_1+iB_2)}f\|^2 = \|f + \frac{1}{n}(B_1+iB_2)f + \frac{1}{2n^2}(B_1+iB_2)^2 f + \ldots\|^2$$

$$\leq (f,f) + \frac{b}{n}(f,f) + \frac{m}{n^2}(f,f)$$

$$= (1 + \frac{b}{n} + \frac{m}{n^2})\ (f,f).$$

Thus $\|e^{\frac{1}{n}(B_1+iB_2)}\| \leq (1 + \frac{b}{2n})\ (1 + \frac{m}{n^2})$,

$$\|e^B\| \leq \|e^{\frac{1}{n}(B_1+iB_2)}\|^n \leq (1 + \frac{b}{2n})^n (1 + \frac{m}{n^2})^n.$$

Taking the limits as $n \to \infty$, we get $\|e^B\| \leq e^{b/2}$.

Now $\|e^{tT}(Q)\| = \|e^{Bt}\, Q\, e^{-At}\| \le \|e^{Bt}\|\ \|Q\|\ \|e^{-At}\|$

$$\le e^{-t(a-b)/2}\ \|Q\| \text{ for } t \ge 0.$$

Also, $\int_0^\infty e^{tT}\,(Q)\,dt = \int_0^\infty e^{Bt}\,Q\,e^{-At}\,dt$, and these integrals are absolutely convergent, we then get

$$\int_0^\infty \|e^{tT}\,(Q)\|\,dt = \int_0^\infty \|e^{Bt}\,Q\,e^{-At}\,dt\| \le \left(\int_0^\infty e^{-t(a-b)/2}\,dt\right)\|Q\|$$

$= \frac{2}{a-b}\|Q\|$, for any Q, which finally gives

$$(2.2.4) \qquad \int_0^\infty \|e^{tT}\|\,dt \le \frac{2}{a-b}.$$

We now complete the proof, by showing that

$$-\int_0^\infty e^{tT}\,dt = T^{-1}$$

Actually, $-T\int_0^\infty e^{tT}\,dt = -\int_0^\infty e^{tT}\,T\,dt$

$$= -\int_0^\infty \frac{d}{dt}e^{tT}\,dt = -\lim_{t\to\infty} e^{tT} + I$$

$= I$ (This follows from relation (2.2.4)).

Thus $T^{-1} = -\int_0^\infty e^{tT}\,dt$,

i.e. $T^{-1}(Q) = -\int_0^\infty e^{tT}\,(Q)\,dt = -\int_0^\infty e^{Bt}\,Q\,e^{-At}\,dt.$

Thus, the operator equation $BX - XA = Q$, has solution X, and

$$\|X\| = \|T^{-1}(Q)\| \leq \frac{2}{a-b}\|Q\|.$$

§2.3 The Operator Equation BX - XA = Q in a More General Setting

Theorem 2.3.1 [11]. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, let the operators A on $\mathcal{X}$ and B on $\mathcal{Y}$, satisfy $\|A\|_1 \leq \alpha$ and $\|B^{-1}\|_1 \leq (\alpha+\delta)^{-1}$, for some $\alpha \geq 0$ and $\delta > 0$. $\|\cdot\|_1$ denotes the bound norms on the respective spaces. For any transformation from $\mathcal{X}$ to $\mathcal{Y}$, we may use any norm compatible with the bound norms (See App. B). Assume $BX - XA = Q$, then $\|Q\| \geq \delta\|X\|$.*

Proof. Compatibility implies that

$$\|XA\| \leq \|X\| \, \|A\|_1 \leq \alpha\|X\|, \text{ and } \|X\| = \|B^{-1}BX\|$$

$$\leq \|BX\| \, \|B^{-1}\|_1 \leq (\alpha+\delta)^{-1} \|BX\|,$$

i.e. $$\|BX\| \geq (\alpha+\delta)\|X\|$$

From $BX - XA = Q$, it follows that

$$\|Q\| \geq \|BX\| - \|XA\| \geq (\alpha+\delta)\|X\| - \alpha\|X\| = \delta\|X\|$$

This result is similar to theorem 2.2.11, but the separation of the spectrum of A and B; $\sigma(A) \cap \sigma(B) = \phi$ does not give as sharp a result as theorem 2 .3 .1 or theorem 2.2.11.

Further generalizations of theorem 2.3.1 for unbounded operators A and B may be found in [11].

Theorem 2.3.2. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hilbert spaces, let B on $\mathcal{Y}$ and A on $\mathcal{X}$ be semi-bounded self adjoint operators, satisfying*

$$B \geq \gamma+\delta \geq \gamma \geq A$$

for some scalars γ and δ. Assuming $BX-XA = C$, where X and C are bounded operators from $\mathcal{X}$ to $\mathcal{Y}$ then $\|C\| \geq \delta\|X\|$ for every unitary-invariant norm.

CHAPTER 3

Rotation of Eigenvectors by a Perturbation

§3.1 Rotation of Eigenvectors by a Perturbation in a Finite Dimensional Space.

We discuss here how the eigenvalues and the eigenvectors (or eigenprojections) change with the change of the operator, in particular when the operator depends analytically on a parameter. The discussion of the finite dimensional case is analogous to that of the general case when the eigenvalues are isolated. However it is easy to treat the finite dimensional case separately, without being bothered by complications arising from the infinite dimensionality of the underlying space. Another reason for treating the finite dimensional case separately is that the finite dimensional theory has its direct applications for example, in connection with the numerical analysis of matrices. The method used is based on a function-theoretic study of the resolvent, in particular on the expression of eigenprojections as contour integrals of the resolvent.

Let X be a finite dimensional normed space, and let $T \in \mathcal{B}(X)$ be an operator having eigenvalues λ_h; $h = 1,2,\ldots,s$ with multiplicities m_h; $h = 1,2,\ldots,s$. It is known that T has the canonical form

(3.1.1) $$T = \sum_h \lambda_h P_h + D_h \quad , \text{ where}$$

(3.1.2) $$P_h = \frac{1}{2\pi i} \int_{\Gamma_h} (zI-T)^{-1}\, dz.$$

Here each Γ_h, $h = 1,\ldots,s$ is a positively oriented small circle enclosing λ_h and lying outside other such circles. Finally, D_h and P_h, $h = 1,2,\ldots,s$ satisfy

$$(3.1.3) \qquad P_h P_k = \delta_{hk} P_h; \quad \sum_{h=1}^{s} P_h = I,$$

$$P_h T = TP_h; \quad P_h(T-\lambda_h I) = (T-\lambda_h I)P_h = D_h.$$

P_h is called the eigenprojection, and D_h is the eigennilpotent, and $M_h = P_h X$ is called the algebraic eigenspace of the eigenvalue λ_h of T, where dim $M_h = m_h$ is the algebraic multiplicity of λ_h. T is called diagonable if and only if all $D_h = 0$, $h = 1,2,\ldots,s$, and simple if $m_h = 1$ for $h = 1,2,\ldots,s$. Now

$$(3.1.4) \qquad T = S+D; \quad S = \Sigma\lambda_h P_h; \quad D = \Sigma D_h$$

S is the diagonable operator, D is the nilpotent, S commutes with D since $P_h D_h = D_h P_h = D_h$, $h = 1,2,\ldots,s$, $P_h D_k = 0$ $h \neq k$. Equation (3.1.4) is called the spectral representation of T. This representation is unique, in the sense that if T is the sum of S and D where S is diagonable and D is nilpotent, and S and D commute, then S and D would be given by (3.1.4)

To see the effect of a perturbation on a linear operator T, we consider a family of operators of the form

$$T(\chi) = T + \chi T'$$

$T(0) = T$ is the unperturbed operator, and $\chi T'$ is the perturbation. Now, if we can express the eigenvalues and the eigenvectors of $T(\chi)$ as power series in χ, then they will be of at least the same order of magnitude as the perturbation $\chi T'$ for small $|\chi|$. This is not always the case. More details are given in [24].

If $T(\chi) \varepsilon \mathcal{B}(X)$ is a family, holomorphic in a domain D_0 of the complex χ-plane. By representing $T(\chi)$ as a matrix with respect to a basis of X, then the eigenvalues of $T(\chi)$ satisfy the characteristic equation

$$\det(T(\chi) - \lambda(\chi)) = 0$$

This is an algebric function in $\lambda(\chi)$ with coefficients holomorphic in χ. It is known [25] that the roots of this equation are branches of analytic functions of χ with only algebraic singularities in D_0; such points are called exceptional points. So at an exceptional point there is always splitting of the eigenvalues. As an illustration, consider the two-dimensional example where $T(\chi)$ is represented by a matrix with respect to a basis $T(\chi) = \begin{pmatrix} 1 & \chi \\ \chi & -1 \end{pmatrix}$. The eigenvalues of $T(\chi)$ are $\lambda_{\pm}(\chi) = \pm(1+\chi^2)^{1/2}$. The exceptional points are $\chi = \pm i$, $T(\pm i)$ have only the eigenvalue 0. Now the number s of eigenvalues of $T(\chi)$ is constant if χ is not one of the exceptional points, of which there are only a finite number in each compact subset of D_0. In each simple subdomain (simply connected subdomain containing no exceptional points) D of D_0, the eigenvalues of $T(\chi)$ can be expressed as s holomorphic functions $\lambda_h(\chi)$, $h = 1,2,..,s$. The eigenvalue $\lambda_h(\chi)$ has constant multiplicity m_h. The eigenprojections $P_h(\chi)$ and the eigennilpotents $D_h(\chi)$ for the eigenvalues $\lambda_h(\chi)$ of $T(\chi)$ are also holomorphic in each simple subdomain D. In this case there is exactly one eigenvalue $\lambda(\chi)$ of $T(\chi)$ in the neighbourhood of λ, and $P(\chi)$ is itself the eigenprojection for this eigenvalue $\lambda(\chi)$. Note that $\dim P_h(\chi) = \dim P_h = m_h$, the multiplicity of the eigenvalue $\lambda_h(\chi)$. Most of the results in error estimates are much simplified when X is a unitary space and T is normal. We have the following

Theorem 3.1.1 ([24] p. 95)

Let X be a unitary space, let $T(\chi) = T+\chi T^{(1)}$, and let T be normal. Then, the power series for $P(\chi)$ and $\lambda(\chi)$ are convergent if the magnitude of the perturbation $\|\chi T^{(1)}\|$ is smaller than half the isolation distance of the eigenvalues λ of T.

So far, we speak about eigenprojections. Since the eigenvectors are not uniquely determined, there are no definite formulas for the eigenvectors of $T(\chi)$ as functions of χ. However, they vary analytically under analytic perturbations. In some situations, we may need sharp bounds on the distance between eigenvectors, and those approximating them. We will discuss this case for Hermitian matrices, or equivalently for Hermitian operators. Let A and A+H be Hermitian operators, acting on n-dimensional complex (or real) Hilbert space $\mathcal{H}$. We denote the eigenvalues of A by λ_i $i = 1,2,\ldots,n$ where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ and let the corresponding normalized eigenvectors be x_i $i = 1,\ldots,n$. By λ'_i; $i = 1,\ldots,n$ we denote the eigenvalues of A+H where $\lambda'_1 \geq \lambda'_2 \geq \ldots \geq \lambda'_n$, and let the corresponding normalized eigenvectors be x'_i, $i = 1,2,\ldots,n$. Sometimes we will speak about the spectral projectors $E(I)$ and $E'(I)$ the argument I of which is a subset of the real line. Now given a specified perturbation H, how much may x_i be rotated to become x'_i? For that, suppose the spectrum of A is confined to m intervals of length $\leq 2\beta$, with gaps $\geq \gamma > 0$, so that we can write

$$P_j = E([\nu_j, \mu_j]) \qquad j = 1,2,\ldots,m$$

$$0 \leq \mu_j - \nu_j \leq 2\beta$$

$$\nu_j - \mu_{j+1} \geq \gamma$$

$$\Sigma P_j = I$$

Let $\|H\| = \delta < \gamma/2$, then $P'_j = E'([\nu_j - \delta, \mu_j + \delta])$ is of the same dimensionality as the corresponding P_j. Generalizing what has been done in §1.3, we try to find a unitary W which for all j satisfies $WP_j = P'_jW$. [Note that W will not necessarily take eigenvectors to eigenvectors]. Every vector x in $P_j\mathcal{H}$ is nearly an eigenvector, in the sense that

$\|Ax - \frac{1}{2}(\mu_j + \nu_j)x\| \le \beta \|x\|$. To see this, let $\lambda_{j1}, \lambda_{j2}, \ldots, \lambda_{js}$ be the eigenvalues of A in $[\nu_j, \mu_j]$ and $x \in P_j\mathcal{H}$. Then $x = \sum_{i=1}^{s} \alpha_i e_{ji}$, e_{ji} being the eigenvectors of A corresponding to λ_{ji}, and

$$\|Ax - \frac{1}{2}(\nu_j + \mu_j)x\| = \| \sum_{i=1}^{s} \alpha_i \lambda_{ji} e_{ji} - \frac{1}{2}(\mu_j + \nu_j) \alpha_i e_{ji}\|$$

$$\le \max_i |\lambda_{ji} - \frac{1}{2}(\mu_j + \nu_j)| \, \|x\|.$$

Consequently,

$$\text{(3.1.5)} \quad \|Ax - \frac{1}{2}(\nu_j + \mu_j)x\| \le \beta\|x\|.$$

The method of constructing a canonical unitary map, which carries $P_j\mathcal{H}$ to $P'_j\mathcal{H}$, is carried over from the special case. Let $B = \sum_j P_jP'_j$. It is easy to check that B is normal. Let $C = B\,B^*$ then $C \ge 0$ and

$$C = (\sum_j P_jP'_j)\ (\sum_j P'_jP_j) = \sum_j P_jP'_jP_j$$

$$= (\sum P'_jP_j)\ (\sum_j P_jP'_j) = \sum_j P'_jP_jP'_j$$

From the definition of C, it follows that

$$CP_j = P_jP'_jP_j = P_jC \text{ for all } j,$$

$$CP'_j = P'_jP_jP'_j = P'_jC \text{ for all } j.$$

Define $U(\{P_j\}, \{P'_j\}) = (\sum_j P_j P'_j P_j)^{-1/2} (\sum_j P'_j P_j)$

$$= (\sum_j P'_j P_j) (\sum_j P_j P'_j P_j)^{-1/2},$$

It follows that $U U^* = U^* U = I$, as well as

$$U(\{P_j\}, \{P'_j\})^* = U(\{P'_j\}, \{P_j\}),$$

$$UP_j = (\sum_j P_j P'_j P_j)^{-1/2} P'_j P_j = P'_j (\Sigma P_j P_j' P_j)^{-1/2} P_j$$

$$= P'_j P_j (\Sigma P_j P'_j P_j)^{-1/2} = P'_j U, \text{ in short,}$$

$$UP_j = P'_j U = P'_j (P_j P'_j P_j)^{-1/2} P_j \text{ for all } j \text{ where}$$

$(P_j P'_j P_j)^{-1/2}$ is the pseudo inverse.

A sufficient condition for the existence of such U is $\|P_j - P'_j\| < 1$ for all j, or equivalently $x = P_j x \neq 0$ implies $P'_j x \neq 0$ for all j. (See theorem 1.1.2 and recall that we are in a finite-dimensional space). This condition will be satisfied if P_j and P'_j arise from A and A+H as described above. One can get results similar to those in [8]. For this, let $P_1, P_2, \ldots, P_m$ be a complete orthogonal set of projectors (one may take them to be the spectral projectors of A). We define the pinching of B by P_j as

$$\mathcal{C}B = \sum_j P_j B P_j.$$

$\mathcal{C}$ has the following properties:

Lemma 3.1.2

In the real Hilbert space of Hermitian operators on $\mathcal{H}$ *under the Frobenius norm,* $\mathcal{E}$ *is a projector, and it is trace preserving.*

Proof.

Let F denote the real Hilbert space of the Hermitian operators on $\mathcal{H}$ under the Frobenius norm $\|\cdot\|_F$, and let $\mathcal{E}$: $F \to F$ be defined by

$$\mathcal{E} B = \sum_j P_j B P_j \quad ,$$ then we have

$$\mathcal{E}^2 B = \mathcal{E}(\mathcal{E} B) = \sum_j P_j (\sum_j P_j B P_j) P_j = \sum_j P_j B P_j$$

i.e. $\mathcal{E}^2 B = \mathcal{E} B$ for any $B \varepsilon F$;

$$(\mathcal{E} B, A) = \mathrm{tr}(\mathcal{E} B) A = \sum_j \mathrm{tr}\, P_j B P_j A = \sum_j \mathrm{tr}\, B (P_j A P_j)$$

$$= \mathrm{tr}\, B \mathcal{E} A = (B, \mathcal{E} A)$$

i.e. $(\mathcal{E} B, A) = (B, \mathcal{E} A)$ for any A and $B \varepsilon F$.

Thus $\mathcal{E} = \mathcal{E}^*$, and $\mathcal{E}$ is a projector. Let the orthogonal complement of $\mathcal{E}$ be denoted by $\tilde{\mathcal{E}}$, so $B = \mathcal{E} B + \tilde{\mathcal{E}} B$, and hence

$$\|B\|^2_F = \|\mathcal{E} B\|^2_F + \|\tilde{\mathcal{E}} B\|^2_F,$$ where $\|\cdot\|_F$ denotes the Frobenius norm.

We now prove that $\mathcal{E}$ is trace preserving:

$$\operatorname{tr} \mathcal{E} B = \sum_j \operatorname{tr} P_j B P_j = \sum_j \operatorname{tr} B P_j .$$

Let $\{x_j\}_{j=1}^{n}$ be a complete orthonormal set of vectors adapted to the decomposition of $\mathcal{H}$ by $\{P_j\}$

$$\operatorname{tr} \mathcal{E} B = \sum_{j=1}^{m} \operatorname{tr} B P_j = \sum_{j=1}^{m} \left(\sum_{i=1}^{mj} (B P_j x_i, x_i) \right)$$

where $m_j = \dim P_j$, $\sum_{j=1}^{m} m_j = n = \dim \mathcal{H}$

$$\operatorname{tr} \mathcal{E} B = \sum_{j=1}^{m} \left(\sum_{i=1}^{mj} (B x_i, x_i) \right) = \sum_{i=1}^{n} (B x_i, x_i) = \operatorname{tr} B.$$

So $\mathcal{E}$ is trace preserving, as claimed.

Theorem 3.1.3 [9].

Let $P_1, P_2, \ldots, P_m$ and $P'_1, P'_2, \ldots, P'_m$ be two complete sets of orthogonal projectors, such that $x = P_j x \neq 0$ implies that $P'_j x \neq 0$. Let $U = U(\{P_j\}, \{P'_j\})$ be defined as before. Let W be any unitary, such that $WP_j = P'_j W$. Then

$$\| \mathcal{E}((I-W^*)(I-W)) \|_{\phi} \geq \| \mathcal{E}((I-U^*)(I-U)) \|_{\phi},$$

for any unitary invariant norm.

Corollary 3.1.4.

Under the hypothesis of theorem 3.1.3, we have

$$\|I-W\|_F \geq \|I-U\|_F .$$

Proof: Note that

$$\|I-W\|^2_F = \text{tr } (I-W)^*(I-W) = \sum_{i=1}^{n} ((I-W^*) \ (I-W) y_i, y_i)$$

where the sum is taken over orthonormal y_i. The right hand side is equal to the sum of the eigenvalues of $(I-W^*)(I-W)$; we will denote it by

$$\|(I-W^*)(I-W)\|_1$$

We know that $\|\cdot\|_1$ is unitary invariant, and hence we can apply the previous theorem:

$$\|\mathcal{E}((I-U^*)(I-U))\|_1 \le \|\mathcal{E}((I-W^*)(I-W))\|_1.$$

But $\|\mathcal{E}((I-W^*)(I-W))\|_1 = \text{tr } \mathcal{E}((I-W^*)(I-W))$

$$= \text{tr } (I-W^*)(I-W).$$

This implies that $\|\mathcal{E}(I-W^*)(I-W))\|_1 = \|I-W\|^2_F$

From this, the result follows.

We now get a bound for the rotation of a single spectral subspace. Let $P = P_j = E([\nu_j, \mu_j])$ of A, where $\mu_j - \nu_j \le 2\beta$, $\beta \ge 0$, and the intervals $(\nu_j - \gamma, \nu_j)$ and $(\mu_j, \mu_j + \gamma)$ contain no eigenvalues of A. For a unit vector x, x = Px, we estimate now how large x-P'x is, where $P' = P'_j$ is the corresponding spectral projector of A+H. Without loss of generality, one can take $-\nu_j = \beta = \mu_j$.

Theorem 3.1.5 [9]

If $\|H\| \leq \delta < \gamma/2$, *then*

$$\|(I-P')P\| \leq (\beta+\delta)/(\beta+\gamma-\delta)$$

Proof.

$P = E([-\beta,\beta])$, and the intervals $(-\beta-\gamma,-\beta)$ and $(\beta,\beta+\gamma)$ contain no eigenvalues of A, hence $P' = E([-\beta-\delta,\beta+\delta])$ and the intervals $(-\beta-\gamma+\delta,-\beta-\delta)$ and $(\beta+\delta,\beta+\gamma-\delta)$ do not intersect the spectrum of A+H. For $x \in P\mathcal{H}$ i.e. $x = Px$, $\|x\| = 1$ we have

$$((A+H)^2x, x) = (A^2x,x) + 2\,\mathrm{Re}(AHx, x) + (H^2x, x).$$

Since $PAP \leq \beta$, it follows that $PA^2P \leq \beta^2$, and

$$((A+H)^2x, x) \leq \beta^2 + 2|(AHx, x)| + (H^2x, x)$$

$$\leq \beta^2 + 2\ \|Ax\| . \|Hx\| + \delta^2$$

Since $\|Ax\| \leq \beta\|x\|$, we finally get

$$((A+H)^2x, x) \leq (\beta+\delta)^2. \tag{3.1.6}$$

Since $P'\mathcal{H}$ is the subspace spanned by the eigenvectors corresponding to the eigenvalues of A+H lying in $(-\beta-\gamma+\delta,\beta+\gamma-\delta)$, we obtain

$$(\beta+\gamma-\delta)^2\ \|x-P'x\|^2 \leq \|(A+H)x\|^2. \tag{3.1.7}$$

This follows from [36], theorem 2, with $\alpha=0$, $\varepsilon=\beta+\gamma-\delta$. Equations (3.1.6) and (3.1.7) imply

$$\|(I-P')P\| \le \sup_{\substack{x=Px \\ \|x\|=1}} \|x-P'x\| \le (\beta+\delta)/(\beta+\gamma-\delta)$$

In the above, we gave a bound for the rotation of a single subspace. We now give an estimate of the total amount of rotation, i.e. an estimate of $\|I-U\|^2_F$. From our construction of the unitary canonical mapping, we know that (we recall that $(P_j\, P'_j\, P_j)^{-1/2}$ is the pseudo inverse)

$$UP_j = (P_j\, P'_j\, P_j)^{-1/2}\, P'_j\, P_j, \text{ thus}$$

$$P_j U P_j = (P_j\, P'_j\, P_j)^{1/2}.$$

Since $P_j\, P'_j\, P_j > 0$ on $P_j \mathcal{H}$, $P_j U P_j$ is positive definite, and this implies that $P_j U P_j$ has an orthonormal basis of eigenvectors in $P_j \mathcal{H}$. Let us choose within each $P_j \mathcal{H}$ the x_i's as unit eigenvectors of $P_j\, P'_j\, P_j$ and let $\theta_i = \arccos\,(Ux_i, x_i)$, $\theta_i > 0$. Thus $(P_j\, P'_j\, P_j\, x_i, x_i) = \cos^2 \theta_i$,

$$(P_j(I-P'_j)P_j\, x_i, x_i) = \sin^2 \theta_i, \text{ so that}$$

$$\sin^2 \theta_i = \|(I-P_j')P_j\, x_i\|^2 \le \|(I-P'_j)P_j\|^2$$

and $\sin \theta_i \le \|(I-P'_j)P_j\|$ for some j, $j=1,2,\ldots,m$

Theorem 3.1.6 [9].

Assume that any two eigenvalues of A differ by at least γ, *and suppose that* $\|H\| = \delta<\gamma/2$ *then*

$$\|I-U\|^2_F \le \frac{2}{1+\cos\alpha}\,\frac{\|H\|^2_F}{\gamma(\gamma-2\delta)}$$

where $\alpha = \arcsin \frac{\delta}{\gamma-\delta}$.

Proof.

$$\|I-U\|^2_F = tr((I-U^*)(I-U))$$

$$= \sum_{i=1}^{n} ((I-U^*)(I-U)x_i,\ x_i)$$

where $\{x_i\}_{i=1}^{n}$ is an orthonormal set. Taking $\{x_i\}$ to be the eigenvectors of $P_j\ P'_j\ P_j$, we get

$$\|I-U\|^2_F = \sum_{i=1}^{n} ((2I-U-U^*)x_i,\ x_i) = \sum_{i=1}^{n} (2-2\cos\theta_i)$$

$$= 2\sum_{i=1}^{n} (1-\cos\theta_i) = 2\sum \frac{\sin^2\theta_i}{1+\cos\theta_i}$$

$$\le 2\,\frac{\Sigma \sin^2\theta_i}{1+\min\cos\theta_i}$$

For any i, $\sin\theta_i \le \|(I-P'_j)P_j\|$, for some j. From theorem 3.1.5, setting $\beta=0$, we get

$$\sin\theta_i \le \frac{\delta}{\gamma-\delta},$$

i.e.

$$\theta_i \le \arcsin\frac{\delta}{\gamma-\delta} = \alpha, \text{ for all } i.$$

Thus $\min \cos \theta_i \geq \cos \alpha$, and

$$\sum_{i=1}^{n} \sin^2 \theta_i = \sum_{j=1}^{m} ((I-P'_j)P_j x_{j_i}, x_{j_i}) = \sum_{j=1}^{m} \operatorname{tr}(I-P'_j)P_j$$

From equation (3.1.6), we have

$$((A+H)^2 x_i, x_i) \leq (H^2 x_i, x_i).$$

Using equation (3.1.7), we get

$$\sum_{i=1}^{n} ((A+H)^2 x_i, x_i) \geq (\gamma^2 - 2\gamma\delta) \sum ((I-P'_j)P_j x_i, x_i)$$

$$\sum_{j=1}^{m} \operatorname{tr} H^2 P_j \geq (\gamma^2 - 2\gamma\delta) \sum_{j=1}^{m} \operatorname{tr} (I-P'_j) P_j.$$

Thus $\sum_{i=1}^{n} \sin^2\theta_i = \sum_{j=1}^{m} \operatorname{tr}(I-P'_j)P_j \leq \frac{1}{\gamma(\gamma-2\delta)} \sum_{j=1}^{m} \operatorname{tr} H^2 P_j$

$$= \frac{1}{\gamma(\gamma-2\delta)} \operatorname{tr} H^2 = \frac{1}{\gamma(\gamma-2\delta)} \|H\|^2_F.$$

Thus $\| I-U\|^2_F \leq \frac{2}{1+\cos\alpha} \frac{\|H\|^2_F}{\gamma(\gamma-2\delta)}$, as claimed.

From the proof, we see that the better way to estimate $\|I-U\|_F$ is via $\sum_{i=1}^{n} \sin^2 \theta_i$, which suggests that θ_i is the most natural way of measuring the direct rotation. One can get better estimates if only one spectral projector and its orthogonal complement are involved. Of course to get any conclusion, there should be some information about the size of H compared to the length γ of the gap in the spectrum of A. Without loss of generality, we take the gap to be between -1 and 1.

Theorem 3.1.7 [9].

Let $P = P_1$ *and* $I-P = P_2$ *be the spectral projectors* $E([1,\infty))$ *and* $E((-\infty,-1])$ *of A respectively, so that A has no spectrum in* $(-1,1)$. *Assume* $\|H\| = \delta<1$, *and let x be any eigenvector of A+H corresponding to an eigenvalue* $\lambda\geq 0$. *Then the acute angle between x and Px satisfies* $\sin 2\theta\leq\delta$. *Assuming instead* $PHP + (I-P)H(I-P) = 0$, *(off-diagonality of H) then* $\tan 2\theta\leq\delta$. *Both inequalities are sharp.*

Proof.

From the assumption, we have $P(A-I)P\geq 0$ and $(I-P)(A+I)(I-P)\leq 0$, thus the spectral projector which should be compared with P is $P' = E'([0,\infty))$ where $P'(A+H)P'\geq 0$ and $(I-P')(A+H)(I-P')\leq 0$. Now $x\varepsilon(I-P)\mathfrak{H}$ is impossible, since if it is true, it would imply that $(Ax,x) \leq -1$, and since $P'x = x$, we get $((A+H)x, x) \geq 0$, and hence $(Hx, x) \geq 1$, and this contradicts $\|H\| < 1$. Now, $x\varepsilon P\mathfrak{H}$ is a trivial case, since it implies $\theta = 0$. If, on the other hand, $PHP + (I-P)H(I-P) = 0$, then for $x = (I-P)x$ we have

$$(Hx,x) = ((I-P)H(I-P)x,x) = -(PHPx,x) = 0$$

again a contradiction. Thus, we assume that x, Px and (I-P)x span a 2-dimensional subspace $Q\mathfrak{H}$, and we represent vectors and operators of $Q\mathfrak{H}$ with respect to the bases vectors $Px = \begin{pmatrix}\cos\theta\\0\end{pmatrix}$ and $(I-P)x = \begin{pmatrix}0\\\sin\theta\end{pmatrix}$. We have

$$QHQ = \begin{bmatrix} h_{11} & h_{12} \\ \bar{h}_{12} & h_{22} \end{bmatrix}, \text{ since } H \text{ is Hermitian}$$

$$\|QHQ\| = \sup_{\|x\|=1} |(QHQx, x)| \le \sup_{\|x\|=1} |(Hx, x)| = \delta$$

Thus $\left\| \begin{pmatrix} h_{11} & h_{12} \\ \bar{h}_{12} & h_{22} \end{pmatrix} \right\| \le \delta$

Now $y \varepsilon Q\mathcal{H}$ implies that $Py \varepsilon Q\mathcal{H}$, so that $Q\mathcal{H}$ is an invariant subspace of P and hence a reducing subspace. This implies that P commutes with Q.

Thus $QAQ = \begin{pmatrix} a_{11} & a_{12} \\ \bar{a}_{12} & a_{22} \end{pmatrix}$, and since A commutes with P,

we get

$$QAQ = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \ a_1 \ge 1 \text{ and } a_2 \le -1.$$

Now $Q(A+H)x = \lambda Qx$ means that

$$\begin{pmatrix} a_1+h_{11} & h_{12} \\ \bar{h}_{12} & a_2+h_{22} \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} \lambda\cos\theta \\ \lambda\sin\theta \end{pmatrix}$$

Thus $0 \le \lambda = a_1 + h_{11} + h_{12}\tan\theta = \bar{h}_{12}\cot\theta + a_2 + h_{22}$ and h_{12} is real, so

$$h_{12}(\cot\theta - \tan\theta) = a_1 - a_2 + h_{11} - h_{22} \ge 2 - 2\delta > 0$$

So, if θ could be $\geq \pi/4$, then $\cot\theta - \tan\theta < 0$ and $h_{12} < 0$. Thus

$$0 \leq \lambda = h_{12}\cot\theta + a_2 + h_{22} < a_2 + h_{22} \leq -1 + \delta,$$

a contradiction. So $\theta < \frac{\pi}{4}$ and $h_{12} > 0$, consequently,

$$\cot\theta - \tan\theta = \frac{a_1 - a_2 + h_{11} - h_{22}}{h_{12}} .$$

For fixed h_{12}, the requirement

$$-\delta \leq \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix} \leq \delta \quad \text{implies that}$$

$$|h_{ii}| \leq (\delta^2 - h_{12}^2)^{1/2}, \text{ thus } h_{11} - h_{22} \geq -2(\delta^2 - h_{12}^2)^{1/2}$$

$$\cot\theta - \tan\theta \geq \frac{a_1 - a_2 - 2\sqrt{\delta^2 - h_{12}^2}}{h_{12}} .$$

The minimum of the right hand side is attained at

$$h_{12} = \delta\sqrt{1 - \frac{4\delta^2}{(a_1 - a_2)^2}} \in (0, \delta], \text{ thus}$$

$$\cot\theta - \tan\theta \geq \frac{a_1 - a_2 - 2\sqrt{\delta^2 - h_{12}^2}}{h_{12}}$$

$$= \delta^{-1}\sqrt{(a_1 - a_2)^2 - 4\delta^2} .$$

Since $a_1 - a_2 \geq 2$, we get $\cot\theta - \tan\theta \geq 2\delta^{-1}\sqrt{1-\delta^2}$

i.e. $\dfrac{\sec^2\theta}{\tan\theta} \le 2/\delta$

Thus $\sin 2\theta \le \delta$

In the other case when $\mathscr{E}\, H = 0$,

i.e. $h_{11} = h_{22} = 0$, we proceed as above:

$$(A+H)x = \lambda x,$$

$$0 \le \lambda = a_1 + h_{12}\tan\theta = \bar{h}_{12}\cot\theta + a_2$$

and again h_{12} is real. From $a_1 - a_2 = h_{12}(\cot\theta - \tan\theta)$, we get $h_{12} < 0$ under the assumption that, $\theta \ge \pi/4$, and in this case

$$0 \le \lambda = h_{12}\cot\theta + a_2 < a_2 < -1, \text{ a contradiction.}$$

Thus $\theta < \pi/4$ and $h_{12} > 0$, and $\cot\theta - \tan\theta = \dfrac{a_1 - a_2}{h_{12}}$.

Since $\|H\| \le \delta$, then $h_{12} \le \delta$, and $\cot\theta - \tan\theta \ge 2/\delta$.

i.e. $\dfrac{2\tan\theta}{1-\tan^2\theta} \le \theta$, from which we finally get $\tan 2\theta \le \delta$.

The proof is complete.

This theorem does not say that the angle θ between x and Px satisfies $\sin 2\theta \le \delta$ for all $x \varepsilon P'\mathscr{Y}$. The following theorem gives similar results as the previous theorem i.e. it gives a bound on the amount of rotation of P. In fact, this theorem coincides with the previous theorem in the 2-dimensional case, but it has a more general setting since the restriction to be finite dimensional is removed.

Suppose that A is a bounded self-adjoint operator on a Hilbert space $\mathcal{H}$. Let P and I-P be complementary projectors reducing A, and let the spectrum of A restricted to $P\mathcal{H}$ be from $[1,\infty)$ and the spectrum of A restricted to $(I-P)\mathcal{H}$ be from $(-\infty,-1]$,

i.e. $PAP \geq P$ and $(I-P)A(I-P) \leq -P$. Let H be a bounded self-adjoint perturbation such that $\|H\| = \delta$. Then A+H will have the spectral projectors P' and (I-P') where P' is the spectral projector of A+H corresponding to $[0,\infty)$, so that

$$P'(A+H)P' \geq 0, \quad (I-P')(A+H)(I-P') \leq 0.$$

We use the following measure of separation between $P\mathcal{H}$ and $P'\mathcal{H}$:

$$\sin^2\theta = \sup\{\|(I-P)x\|^2;\ x = P'x,\ \|x\| = 1\}$$

$$= \|P'(I-P)P'\| = \|P'(I-P)P' + (I-P')P(I-P)\|;$$

the last equality holds, since both sides have the same spectrum ([8], lemma 5.2).

<u>Theorem 3.1.8.</u> [10]

Let A, P, δ and θ be defined as above. Assuming $\delta < 1$, then $\sin 2\theta \leq \delta$. Assuming instead that $PHP + (I-P)H(I-P) = 0$, then $\tan 2\theta \leq \delta$. Both inequalities are sharp.

Proof.

The general case can be reduced to the case where all operators have only point spectrum. This can be done as follows.

By an approximate eigenvalue of an operator $T \in \mathcal{B}(\mathcal{H})$ we mean a complex number μ, such that there exists a sequence x_n such that $\|x_n\| = 1$ and $\|Tx_n - \mu x_n\|$ tends to zero, or equivalently, there does not exist a number $\varepsilon>0$ such that

$$(T - \mu I)^*(T - \mu I) \geq \varepsilon I.$$

By $\sigma_a(T)$ we denote the approximate point spectrum of T, the set of all approximate eigenvalues. Clearly $\sigma_p(T) \subset \sigma_a(T) \subset \sigma(T)$. Now, if T is a normal operator, then it can be shown that $\sigma(T) = \sigma_a(T)$ (C.F. [18], theorem 3.1.2). Now $\mathcal{H}$ will be extended to another Hilbert space $\mathcal{H}'$, in which we shall speak about "approximate eigenvectors". So, if T is a normal operator, and μ and ν are distinct approximate eigenvalues of T, then there exists sequences of unit vectors $\{x_n\}$ and $\{y_n\}$ such that $\|Tx_n - \mu x_n\| \to 0$ and $\|Ty_n - \nu y_n\| \to 0$ hence

$$|(\mu-\nu)(x_n, y_n)| \leq \|\mu x_n - Tx_n\| + \|Ty_n - \nu y_n\|,$$

generalizing the known fact for the eigenvectors of a normal operator for distinct eigenvalues. So we may think of $\{x_n\}$ and $\{y_n\}$ as approximate eigenvectors, with their inner product defined to be $\operatorname{glim}(x_n, y_n)$, where glim denotes the Banach generalized limit defined in[12, P.37]. For the extension through the space of approximate eigenvectors see [4,§3]. Now to every operator $T \in \mathcal{B}(\mathcal{H})$ there corresponds an operator

$\rho(T) \varepsilon \mathcal{B}(\mathcal{H}')$ and the mapping ρ: $\mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ satisfies

1. $\rho(S+T) = \rho(S) + \rho(T)$, $\rho(\lambda T) = \lambda\rho(T)$,
2. $\rho(ST) = \rho(S)\,\rho(T)$, $\rho(T^*) = \rho(T)^*$, $\rho(I) = I$,
3. $\|\rho(T)\| = \|T\|$,
4. $\rho(T) \geq 0$ if and only if $T \geq 0$,
5. For every operator $T \varepsilon \mathcal{B}(\mathcal{H})$

$$\sigma_a(T) = \sigma_a(\rho(T)) = \sigma_p(\rho(T)) \qquad \text{[see [4] §4, Theorem 1].}$$

i.e. ρ preserves algebraic operations, spectra, adjoints, and order. From (2), it follows that $\rho(P)$ will be a projector and every $\rho(T)$ has only point spectrum. So given A,H,P,P' as in the theorem, then $\rho(A)$, $\rho(H)$, $\rho(P)$ and $\rho(P')$ will enjoy the same properties. Since $\sin^2\theta = \|P'(I-P)P'\|$

i.e. the bound on θ is the same as the bound of a norm of certain operator and this is preserved under ρ. Hence, proving the conclusion for $\rho(P'(I-P)P')$ proves it for $P'(I-P)P'$. Now, considering that all operators have only point spectrum then since $\sin^2\theta = \|P'(I-P)P'\|$, the result is a bound on the norm of the positive operator $P'(I-P)P'$. In this case the norm of $P'(I-P)P'$ is its largest eigenvalue.

Assume, then, that $x \varepsilon P'\mathcal{H}$ satisfies $\|x\| = 1$ and $P'(I-P)P'x = \sin^2\theta\, x$, so that $P'PP'x = \cos^2\theta\, x$, $\|Px\| = \cos\theta$.

Let Q be the projector onto the two-dimensional subspace spanned by x, Px, and $(I-P)x$. The possibility that Q be one-dimensional can be ruled out as in the proof of Theorem 3.1.7. Since $Q\mathcal{H}$ is spanned by eigenvectors of P, then P commutes with Q. Similarly P' commutes with Q since $Q\mathcal{H}$ is spanned by x and (I-P')Px (This follows from (I-P')Px = Px - P'Px = Px - P'PP'x = Px - $\cos^2\theta$ x ε $Q\mathcal{H}$). It follows then that QPQ and QP'Q are projectors onto the one-dimensional subspace $Q\mathcal{H} \cap P\mathcal{H}$ and $Q\mathcal{H} \cap P'\mathcal{H}$ respectively.

As before, we represent vectors and operator of Q with respect to the basis vectors:

$$Px = \begin{pmatrix} \cos\theta \\ 0 \end{pmatrix} \quad \text{and} \quad (I-P)x = \begin{pmatrix} 0 \\ \sin\theta \end{pmatrix}$$

Since A = PAP + (I-P) A (I-P), and Q commutes with P, then $QAQ = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$, $a_1 \geq 1$ and $a_2 \leq -1$. Similarly, since A+H = P'(A+H)P' + (I-P')(A+H)(I-P') where $P'(A+H)P' \geq 0$ and $(I-P')(A+H)(I-P') \leq 0$, and since Q commutes with P', it follows that Q(A+H)Q has spectral projectors QP'Q and Q(I-P')Q, and QP'Q and Q(I-P')Q correspond to the nonnegative and nonpositive spectra of Q(A+H)Q. Since $QP'Q\mathcal{H}$ is spanned by x, then x is an eigenvector of P(A+H)A corresponding to an eigenvalue $\lambda \geq 0$. Let $QHQ = \begin{pmatrix} h_{11} & h_{12} \\ \overline{h_{12}} & h_{22} \end{pmatrix}$. Then, it follows that

$$\begin{pmatrix} a_1 + h_{11} & h_{12} \\ \overline{h_{12}} & a_2+h_{22} \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} \lambda\cos\theta \\ \lambda\sin\theta \end{pmatrix}$$

Since $\|QHQ\| \leq \|H\|$, then $\|QHQ\| \leq \delta$, and $PHP + (I-P)H(I-P) = 0$ and $PQHQP + (I-P)QHQ(I-P) = 0$. Since P commutes with Q, then if the bound in either part of the theorem is proved for the 2-dimensional case, it can be carried back from $Q\mathcal{H}$ to $\mathcal{H}$. So the proof is now reduced to the proof of the theorem in the 2-dimensional case which is the same as the proof carried out in Theorem 3.1.7.

§3.2 Rotation of eigenvectors by a perturbation in general.

Here we discuss the case when a Hermitian linear operator is slightly perturbed, and see how far its invariant subspaces will change. This discussion is an extension of the previous analysis in the finite dimensional case, and the main new idea here is the introduction of the operator angle Θ defined in §1.3. These angles unify the treatment of natural geometric, operator theoretic and error-analytic questions concerning those subspaces. Sharp bounds on trigonometric functions of these angles are obtained from the gap between appropriate parts of the spectra and from a bound on the perturbations. Similarly, sharp bounds will be obtained for arbitrary unitary invariant norms, as in [11]. In [9] and [10] such bounds could be asserted only upon the operator's bound-norms. Such theorems are of two types, single-angle theorems and double-angle theorems, and the last ones are extensions of Theorems 3.1.7 and 3.1.8. All the theorems are applicable for infinite as well as finite dimensional spaces. The chief new tool in the proofs is embodied in a simple inequality for binomials AX-XB which were discussed in §2.2 and §2.3.

Since the differences between the subspaces will be measured in terms of trigonometric functions of the angle Θ, we first give the various measures of differences between the subspaces $P\mathcal{V} = R(E_0)$ and $Q\mathcal{V} = R(F_0)$ mentioned in §1.1, in terms of Θ:

(1) $\sin^2 \Theta = P(I-Q)P + (I-P)Q(I-P) = (P-Q)^2$, thus

(3.2.1) $|\sin \Theta| = |P-Q|$, in all unitary invariant norms.

(2) Since $S_0 = J_0 \sin \Theta_0$, and $\sin \Theta_0 = (S_0{}^* S_0)^{1/2}$, then S_0 and $\sin \Theta_0$ have the same singular values and

$\|S_0\| = \|\sin \Theta_0\|$ (Appendix B,) and

(3.2.2) $\|\sin \Theta_0\| = \|S_0\| = \|E_1{}^* U^* E_0\| = \|(I-P)U^* P\|$

$= \|U^*(I-Q)P\| = \|(I-Q)P\| = \|(I-Q)E_0 E_0{}^*\|$

$= \|(I-Q)E_0\| = \|F_1{}^* E_0\| = \|E_0{}^* F_1\|$.

(3) $\text{Sup}\,\{\|Qp-p\|;\ \|p\| = 1,\ p = Pp\} = \|\sin \Theta\|_1$.

Proof.

L.H.S. $= \sup\,\{((I-Q)p, p),\ \|p\| = 1,\ p = Pp\}$

$= \sup\,\{((I-Q)Pp, Pp);\ \|p\| = 1\}$

$= \sup\,\{(P(I-Q)Pp, p);\ \|p\| = 1\}$

$= \|P(I-Q)P\|_1 = \|P(I-Q)P + (I-P)Q(I-P)\|_1$

$= \|\sin \Theta\|_1^2$

Thus

(3.2.3) $\sup\{\|Qp-p\|;\ \|p\| = 1,\ p = Pp\} = \|\sin \Theta\|_1 = \|\sin \Theta_0\|_1$.

(4) $\sup\{\inf\{\|q-p\|;\ \|q\| = 1,\ q = Qq\};\ \|p\| = 1,\ p = Pp\}$

$= 2\|\sin \frac{1}{2}\ \Theta\ \|_1.$

Proof.

Fixing p, we have

$$\inf_{\substack{q=Qq \\ \|q\|=1}}\{\|q-p\|^2,\ \|q\| = 1,\ q = Qq\} = \inf_{\substack{q=Qq \\ \|q\|=1}}\{\|Q(q-p)\|^2 + \|(I-Q)(q-p)\|^2\}$$

$$= \inf\{\|q\|^2 + \|Qp\|^2 - 2\mathrm{Re}(Qp,q) + (I-Q)p\|^2\}$$

$$= \inf\{1 + \|p\|^2 - 2\mathrm{Re}(Qp,q)\} \geq 1 + \|p\|^2 - 2\ \|Qp\|\ \|q\|$$

The equality holds, when $q = \frac{Qp}{\|Qp\|}$, and

$$\inf\{\|q-p\|^2;\|q\| = 1,\ q = Qq\} = 1 + \|p\|^2 - 2\|Qp\|,$$

$$\sup\{1 + \|p\|^2 - 2\|Qp\|,\ \|p\| = 1,\ p = Pp\}$$

$$= \sup\{2 - 2\|Qp\|\} = \sup\{2 - 2(PQPp,\ p)^{1/2}\}$$

$$= 2 - 2\inf\{(PQPp,\ p)^{1/2},\ \|p\| = 1,\ p = Pp\}$$

$$= 2 - 2\cos\theta_1 = 4\sin^2\theta_{1/2} = 4\|\sin \tfrac{1}{2}\ \Theta\ \|_1^2$$

where $\theta_1 \geq \theta_2 \geq \ldots$ are the singular values of Θ_0.

Thus we have

$$(3.2.4) \quad \sup\{\inf\{\|q-p\|;\ \|q\| = 1,\ q = Qq\};\ \|p\| = 1,\ p = Pp\} = 2\|\sin \tfrac{1}{2}\ \Theta\|_1 .$$

In the notation of §1.2 let $P\mathcal{H}$ be a reducing subspace of A and $Q\mathcal{H}$ be a reducing subspace of A+H, so in our decomposition of $\mathcal{H}$ onto $P\mathcal{H}$ and $(I-P)\mathcal{H}$, we have

$$(3.2.5) \quad A = (E_0 \quad E_1) \begin{pmatrix} A_0 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix}$$

$$(3.2.6) \quad H = (E_0 \quad E_1) \begin{pmatrix} H_0 & B^* \\ B & H_1 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix}$$

These equations define the new operators appearing in them e.g. $B = E_1^* HE_0$ is an operator from $K(E_0)$ to $K(E_1)$, and A_j and H_j are Hermitians. On the other hand, in the decomposition of $\mathcal{H}$ according to a reducing subspace $Q\mathcal{H}$ of A+H, the two ways of representing A+H are

$$(3.2.7) \quad A + H = (E_0 \quad E_1) \begin{pmatrix} A_0 + H_0 & B^* \\ B & A_1 + H_1 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix} =$$

$$= (F_0 \quad F_1) \begin{pmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_1 \end{pmatrix} \begin{pmatrix} F_0^* \\ F_1^* \end{pmatrix}$$

From (3.2.5), it is clear that A_0 is isometrically equivalent to a part of A, and instead of comparing A+H with

A and saying that the difference is small, we compare A+H with A_0 acting on a space of lower dimension, and say that the residual R defined by

$$(3.2.8) \qquad R = (A+H)E_0 - E_0A_0$$

(actually, $R = HE_0$ since P commutes with A) is small.

Note that if $E_0 = F_0$, $A_0 = \Lambda_0$, then $R = 0$.

Theorem 3.2.1 [11]

Assume there is an interval $[\beta,\alpha]$ *and a* $\delta>0$, *such that the spectrum of* A_0 *lies entirely in* $[\beta,\alpha]$, *while that of* Λ_1 *lies entirely outside* $(\beta-\delta,\ \alpha+\delta)$ *(or such that the spectrum of* Λ_1 *lies entirely in* $[\beta,\alpha]$, *while that of* A_0 *lies entirely outside* $(\beta-\delta,\ \alpha+\delta)$*). Then for every unitary invariant norm,* $\delta\|\sin\ \Theta_0\| \le \|R\|$.

Remarks.

1) In theorems 3.1.7 and 3.1.8, it has been usual to require a gap between parts of a single operator (e.g. A_0 and A_1). Here a part of A is separated from a part of A+H.

2) Here the spectrum of Λ_1 is also allowed to lie both above and below the spectrum of A_0.

Proof.

Without loss of generality, we may assume $\alpha = -\beta \ge 0$. From (3.2.8), we have

$R = (A+H)E_0 - E_0 A_0 = HE_0$, so for the unitary invariant norms, compatible with the bound norm, we have

$\|R\| = \|R^*\| \geq \|R^* F_1\|$, since $\|F_1\|_1 = 1$. From (3.2.8), we get

$$R^* = E_0^* (F_0 \Lambda_0 F_0^* + F_1 \Lambda_1 F_1^*) - A_0 E_0^* ,$$

$$R^* F_1 = E_0^* F_1 \Lambda_1 - A_0 E_0^* F_1 .$$

Applying theorem 2.3.1, with $\mathcal{X} = K(F_1)$, $\mathcal{Y} = K(E_0)$, $X = E_0^* F_1$, we have (since $||\Lambda_1|| \leq \alpha$ and $||A_0^{-1}||_1 \leq (\alpha+\delta)^{-1}$),

$$\text{(3.2.9)} \qquad \|R\| = \|R^*\| \geq \|R^* F_1\| \geq \delta\|E_0^* F_1\| .$$

From equation (3.2.2), we have $\|E_0^* F_1\| = \|\sin \Theta_0\|$ thus $\|R\| \geq \delta\|\sin \Theta_0\|$ in every unitary invariant norm.

In case of the bound norm, we can strengthen the conclusions, under the same hypothesis, since $\|\sin \Theta\|_1 = \|\sin \Theta_0\|_1$, namely $\|R\|_1 \geq \delta\|\sin \Theta\|_1$, and hence

$$\delta\|\sin \Theta\|_1 \leq \|R\|_1 = \|H E_0\|_1 \leq \|H\|_1, \quad (\|E_0\|_1 = 1)$$

On the other hand, if we allow some more hypotheses on the separation of the parts of the spectra, we may get the following conclusion:

<u>Theorem 3.2.2.</u> [11]

For a given $\delta>0$, assume that the spectra A_0 and Λ_1 are separated as in the hypothesis of theorem 3.2.1, and assume that *the spectra of* A_1 *and* Λ_0 *are also separated as in the hypothesis of the same theorem. Then, for every unitary invariant norm,* $\delta\|\sin\ \Theta\| \le \|H\|$.

Proof

Repeating what has been done in the proof of theorem 3.2.1, it follows from (3.2.2) and (3.2.9) that

(3.2.10) $$\|PH(I-Q)\| = \|E^*_0\ H\ F_1\| \ge \delta\|E^*_0\ F_1\|$$

$$= \delta\|P(I-Q)\| = \delta\|\sin\ \Theta_0\|.$$

Since $HE_1 = (A+H)E_1 - E_1A_1$, it follows that theorem 2.3.1 and from equation (3.2.2), that

(3.2.11) $$\|(I-P)HQ\| = \|E^*_1\ H\ F_0\| \ge \delta\|E^*_1\ F_0\|$$

$$= \delta\|(I-P)Q\| = \delta\|\sin\Theta_1\|.$$

Since (3.2.10) and (3.2.11) are true for all unitary invariant norms, it follows (see appendix B) that

$$\|(I-P)HQ + PH(I-Q)\| \ge \delta\|(I-P)Q + P(I-Q)\|$$

$$= \delta\|[(I-P)Q + P(I-Q)]\ [2Q-I]\| = \delta\|P-Q\|.$$

Thus $\|(I-P)HQ + PH(I-Q)\| \ge \delta\|\sin\ \Theta\|$; this follows from equation (3.2.1). Finally,

$$\delta\|\sin\ \Theta\| \le \|(I-P)HQ + PH(I-Q)\|$$

$$= \frac{1}{2}\|H + (I-2P)\ H(2Q-I)\|$$

$$\le \frac{1}{2}\|H\| + \|(I-2P)\ H\ (2Q-I)\| \le \|H\|,$$

since I−2P and 2Q−I are symmetric. We obtained

$$\delta\|\sin\ \Theta\| \le \|H\|.$$

In some applications of numerical analysis, concerning calculation only of some eigenvalues and eigenvectors of an operator A, this may be translated in our notation as follows: E_0 is used to approximate some of the eigenvectors, and hence the eigenvectors are not exactly orthonormal, and consequently E_0 is no longer an isometry, but we may suppose that $E^*_0 E_0 \ge \varepsilon$ where ε is very near to 1. The following theorem discusses, besides the above case, the case when it is required to compare an eigenspace of A+H with an eigenspace of A, with different dimension.

Theorem 3.2.2 [11]

Assume the Hermitian operator A+H satisfies (3.2.7) and that R is given by (3.2.8). Assume as before that F_0 and F_1 are isometrics with $F_0F_0^* + F_1F_1^* = 1$, but for E_0, assume only that $E_0^*E_0 \ge \varepsilon^2$ for some $\varepsilon>0$. Let P and Q be the projectors onto $R(E_0)$and $R(F_0)$as before, but without any hypothesis on the dimension of these subspaces. Let $\sin\Theta_0$ be *any operator with the same singular values as P(I-Q) which we assume to be compact. Assume there is an interval* $[\beta,\alpha]$ *and a* $\alpha>0$*, such that the spectrum of* A_0 *lies entirely in* $[\beta,\alpha]$ *while that of* Λ_1 *lies entirely outside* $(\beta-\delta,\ \alpha+\delta)$ *(or such that the spectrum of* Λ_1 *lies entirely in* $[\beta,\alpha]$ *while that of* A_0 *lies entirely outside* $(\beta-\delta,\ \alpha+\delta)$*). Then for every unitary-invariant norm,* $\delta\varepsilon\|\sin\Theta_0\| \le \|R\|$.

For some applications, the hypothesis in theorem 3.2.3 concerning the spectra of A_0 and Λ_1 is too restrictive. As a partial relief, we have the following theorem:

Theorem 3.2.4

Assume that all the hypotheses of theorem 3.2.3 are satisfied, except that the only restriction on the spectra is that $|\lambda-\alpha| \ge \delta>0$ *for all* λ *in the spectrum of* Λ_1 *and* α *in the spectrum of* A_0*. Assuming in addition that* Λ_1 *and* A_0 *are diagonable, then*

$$\delta\varepsilon\|\sin\Theta_0\|_{sq} \le \|R\|_{sq}\ .$$

Proof.

Note that the conclusion is trivial if $\|R\|_{sq}$ is infinite. Otherwise, the proof goes on the lines as for theorem 3.2.1, except instead of applying Theorem 2.3.1 we need to show that the equation $C = A_0X - X\Lambda_1$ has a solution X, which satisfies $\|C\|_{sq} \geq \delta\|X\|_{sq} = \delta(\mathrm{tr}\ X^*X)^{1/2}$. To show that, consider the following singular decomposition of A_0 and Λ_1; $A_0 = UD_{A_0}U^*$ and $\Lambda_1 = VD_{\Lambda_1}V^*$ where D_{A_0} and D_{Λ_1} are diagonal relative to suitable orthonormal bases and U,V are corresponding isometrics. The equation $C = A_0X - X_{\Lambda_1}$ reduces to $U^*CV = D_{A_0}U^*XV - U^*XVD_{\Lambda_1}$, $B = U^*CV$, $Y = U^*XV$, $b_{ij} = \alpha_i y_{ij} - y_{ij}\lambda_i$,

$$|b_{ij}|^2 = |\alpha_i - \lambda_i|^2 |y_{ij}|^2 \geq \delta|y_{ij}|^2,$$

$$\sum_{i,j} |b_{ij}|^2 \geq \delta \sum_{i,j} |y_{ij}|^2$$

$$\|U^*\ CV\|_{sq} \geq \delta\|U^*\ XV\|_{sq}$$

But $\|\cdot\|_{sq}$ is unitary invariant, thus

$$\|C\|_{sq} \geq \delta\|X\|_{sq}.$$

Now applying this inequality, to the equation

$$R^*\ F_1 = E_0^{\ *}\ F_1\ \Lambda_1 - A_0\ E_0^{\ *}\ F_1,$$ we get

$$\|R^*\ F_1\|_{sq} \geq \delta\|E_0^{\ *}\ F_1\|_{sq}.$$

But $\|P(I-Q)\| = \|\sin\ \Theta_0\|$ for any unitary-invariant norm, in particular

$$\|P(I-Q)\|_{sq} = \|\sin\ \Theta_0\|_{sq}.$$

Since $P = E_0(E_0^* E_0)^{-1} E_0^*$, one calculates

$$\|P(I-Q)\|_{sq} = \|E_0(E_0^* E_0)^{-1} E_0^* F_1 F_1^*\|_{sq}$$

$$\leq \|E_0(E_0^* E_0)^{-1}\|_1 \|E_0^* F_1\|_{sq} \|F_1^*\|_1.$$

From $E_0^* E_0 \geq \varepsilon^2 > 0$, we get $(E_0^* E_0)^{-1} \leq 1/\varepsilon^2$, and thus

$$\|P(I-Q)\|_{sq} \leq \frac{1}{\varepsilon}\|E_0^* F_1\|_{sq}, \text{ and}$$

$$\varepsilon\delta\|\sin \Theta_0\|_{sq} \leq \|E_0^* F_1\|_{sq}.$$

<u>Theorem 3.2.5</u> [11]

Assume there is an interval $[\beta,\alpha]$ *and a* $\alpha>0$ *such that the spectrum of* A_0 *lies entirely in* $[\beta,\alpha]$ *while that of* Λ_1 *lies entirely outside* $(\beta-\delta, \infty)$. *Assume further that* $H_0 = 0$, *then for every unitary-invariant norm,* $\delta\|\tan \Theta_0\| \leq \|R\|$ *and* $\delta\|\tan \Theta\| \leq \|H\|$.

<u>Remark</u>.

Note that the spectrum of Λ_1 should lie above that of A_0, in contrast with theorem 3.2.1, but we have gained an improved bound by a further assumption.

<u>Proof</u>.

In terms of the direct rotation U,

$$(3.2.12) \qquad U \simeq \begin{pmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{pmatrix} = \begin{pmatrix} \cos\ \Theta_0 & -J_0^*\ \sin\ \Theta_1 \\ J_0\ \sin\ \Theta_0 & \cos\ \Theta_1 \end{pmatrix}$$

where $J_0\ \Theta_0 = \Theta_1 J_0$, $c_j \geq 0$

We rewrite (3.2.7) in terms of (3.2.12) in the form

$$(3.2.13) \qquad \begin{pmatrix} A_0+H_0 & B^* \\ B & A_1+H_1 \end{pmatrix} \begin{pmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{pmatrix} = \begin{pmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{pmatrix} \begin{pmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_1 \end{pmatrix}$$

Thus, it follows that

$$(3.2.14) \qquad (A_0+H_0)(-S_0^*) + B^*C_1 = -S_0^*\Lambda_1.$$

But $B \leq A_0 \leq \alpha < \alpha+\delta \leq \Lambda_1$ and $H_0 = 0$, and $R = HE_0 = E_0H_0 + E_1B = E_1B$, thus

$\|R\| = \|E_1B\| = \|B\|$ for every unitary-invariant norm. From (3.2.14), we get

$$(3.2.15) \qquad C_1B = S_0A_0 - \Lambda_1S_0.$$

To simplify the proof, we assume that all the operators are bounded, and S_0 is compact. Since $\|A_0\|_1 \leq \alpha$, $\|\Lambda^{-1}{}_1\|_1 \leq \frac{1}{\alpha+\delta}$, then applying theorem 2.3.1 we get

$$\|C_1B\| \geq \delta\|S_0\|$$

To get our conclusion, we try to prove that $\|B\|_\nu \geq \delta\|\tan\ \Theta_0\|_\nu$, from this, it follows that $\|R\| = \|B\| \geq \delta\|\tan\ \Theta_0\|$

for all unitary invariant norms (Appendix B). For an operator K, we use the norm

$$\|K\|_\nu = \sup_{\Omega,T}\|TK\Omega\|_\nu = \sup \operatorname{Re} \sum_{k=1}^{\nu} y_k^* K x_k$$

The first sup is taken over pairs of ν-projectors Ω and T, and the second sup is taken over all orthonormal ν-tuples $\{x_1, x_2, \ldots, x_\nu\}$ and $\{y_1, y_2, \ldots, y_\nu\}$.

Since $(S_0^* S_0)^{1/2} = \sin\ \Theta_0$, and S_0 is compact, then $S_0^* S_0$ has the eigenvalues $\sin^2\theta_1 \geq \sin^2\theta_2 \geq \ldots$. We calculate $\|B\|_\nu$ for integers ν exceeding neither dim $K(E_0)$ nor dim $K(E_1)$. We choose orthonormal ν-eigenvectors $x_{01}, x_{02}, \ldots, x_{0\nu} \in K(E_0)$ corresponding to eigenvalues $\sin^2\theta_1 \geq \sin^2\theta_2 \geq \ldots \geq \sin^2\theta_\nu$, then we choose orthonormal vectors $y_{11}, y_{12}, \ldots, y_{1\nu} \in K(E_1)$ defined by $y_{1j} = -S_0\ x_{0j}/\sin\theta_j$, $\theta_j \neq 0$. If $\theta_j = 0$, we take y_{1j} to form an orthonormal set from $N(S_0^*)$ so y_{1j} satisfies $S_0^* y_{1j} = \sin\theta_j\ x_{0j}$ $(S_0^* S_0 x_{0j} = \sin^2\theta_j\ x_{0j})$. From $C_1 = (I - S_0 S_0^*)^{1/2}$ on $K(E_1)$, we get

$$S_0\ S_0^*\ y_{1j} = -\sin\ \theta_j\ S_0\ x_{0j} = \sin^2\ \theta_j\ y_{1j}$$

$$C_1\ y_{1j} = \cos\ \theta_j\ y_{1j}$$

Now, from (3.2.15), it follows that $y^*_{1j}\ (C_1 B)\ x_{0j} =$ $y^*_{1j}\ (S_0\ A_0 - \Lambda_1\ S_0)\ x_{0j}$ $\cos\ \theta_j\ y^*_{1j}\ B\ x_{0j} = -\sin\ \theta_j\ x^*_{0j}\ A\ x_{0j} +$ $\sin\ \theta_j\ y_{1j}\ \Lambda_1\ y_{1j} = \sin\ \theta_j\ (y^*_{1j}\ \Lambda_1\ y_{1j} - x^*_{0j}\ A_0\ x_{0j})$

Since $\Lambda_1 \geq \alpha+\delta$, $\alpha \geq A_0$ we find $y^*_{1j}\ \Lambda_1\ y_{1j} \geq \alpha+\delta$, $x^*_{0j}\ A_0\ x_{0j} \leq \alpha$, $\cos\ \theta_j\ y^*_{1j}\ B\ x_{0j} \geq \sin\ \theta_j\ (\alpha+\delta-\alpha) = \delta\ \sin\ \theta_j$;

Since $\delta>0$ implies that $\cos\theta_j > 0$, we have

$$y^*_{1j}\, B\, x_{0j} \geq \delta \tan\theta_j,$$

$$\|B\|_\nu = \sup \sum_{j=1}^{\nu} y^*_{1j}\, B\, x_{0j} \geq \delta \sum_{j=1}^{\nu} \tan\theta_j = \delta\|\tan\Theta_0\|_\nu .$$

Thus $\|R\| = \|B\| \geq \delta\,\|\tan\Theta_0\|$ for all unitary-invariant norms. Now, since $\|\tan\Theta\| = \|J\sin\Theta(\cos\Theta)^{-1}\|$, then in matrix notation we have

$$J\sin\Theta(\cos\Theta)^{-1} = \begin{pmatrix} 0 & -J^*_0\tan\Theta_1 \\ J_0\tan\Theta_0 & 0 \end{pmatrix}$$

and $\|J_0\tan\Theta_0\| = \|J_0^*\tan\Theta_1\| = \|\tan\Theta_0\| \leq \|B\|/\delta.$

It implies that

(3.2.16) $$\delta\|\tan\Theta\| = \|E_1J_0\tan\Theta_0E^*_0 - E_0J_0^*\tan\Theta_1E_1^*\| \leq \|E_1BE^*_0 + E_0B^*E^*_1\| = \|(I-P)HP + PH(I-P)\| \leq \|H\|$$

(For the 1st and the 2nd equality in equation (3.2.16) see Appendix B)

If we now assume that the gap is between A_0 and A_1 or between Λ_0 and Λ_1, we have the following:

<u>Theorem 3.2.6</u>. [11]

Assume that there is an interval $[\beta,\alpha]$ *and a* $\delta>0$, *such that the spectrum of* Λ_0 *lies entirely in* $[\beta,\alpha]$ *while that of* Λ_1 *lies entirely outside* $(\beta-\delta,\alpha+\delta)$, *then for every unitary-invariant* *norm,* $\delta\|\sin 2\ \Theta_0\| \le 2\|R\|$ *and* $\delta\|\sin 2\Theta\| \le 2\|H\|$.

Proof.

Let $X = 2P - I \simeq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and $Q_- = XQX$; clearly

$X^2 = I$, $X = X^* = X^{-1}$, $Q_-^2 = Q_- = (XF_0)(XF_0)^*$, and

$$(3.2.17) \qquad X(A+H)X = A+XHX \simeq \begin{pmatrix} A_0 + H_0 & -B^*{}_0 \\ -B & A_1+H_1 \end{pmatrix}$$

From equation (1.3.10), we have $U^2 = (2Q - I)(2P - I)$, thus $U^2X = 2Q - I$, and since Q commutes with A+H, we obtain

$$(A + H)U^2X = U^2X(A + H),$$
$$(A + H)U^2 = U^2(A + XHX),$$

and in matrix notation

$$(3.2.18) \qquad \begin{pmatrix} A_0 + H_0 & B^* \\ B & A_1+H_1 \end{pmatrix} \begin{pmatrix} C_0 & -S^*{}_0 \\ S_0 & C_1 \end{pmatrix}^2 = \begin{pmatrix} C_0 & -S^*{}_0 \\ S_0 & C_1 \end{pmatrix}^2 \begin{pmatrix} A_0+H_0 & -B^* \\ -B & A_1+H_1 \end{pmatrix}.$$

Since $U^2Q_- = U^2XQX = (2Q-I)QX = Q(2Q-I)X = QU^2$, we find that

$$U^2 \simeq \begin{pmatrix} \cos 2\ \Theta_0 & -J_0 \sin 2\ \Theta_1 \\ J_0 \sin 2\ \Theta_0 & \cos 2\ \Theta_1 \end{pmatrix}$$

is a unitary taking $Q_-\mathcal{H}$ to $Q\mathcal{H}$.

The intention is to apply theorem 3.2.2 by regarding A+H as a perturbation, of A + XHX i.e. the perturbation is H - XHX. The parts of A+H on $Q\mathcal{H}$ and $(I-Q)\mathcal{H}$ are represented by Λ_0 and Λ_1

where $\Lambda_j = F_j^*(A+H)F_j$, $j = 0,1$. Clearly Q_- commutes with $A+XHX$, hence the parts of $A+XHX$ in $Q_-\mathcal{H}$ and $(I-Q_-)\mathcal{H}$ are

$$A_j = (XF_j)^* (A+XHX)(XF_j) = \Lambda_j$$

and the hypothesis of theorem 3.2.2 is satisfied, replacing E_0 by XF_0 and $E_0^*F_1$ by $(XF_0)^*F_1 = F_0^*XF_1 = 2F_0^*E_0E_0^*F_1 = 2(F_0^*E_0)(E_0^*F_1)$ so

$$\|(XF_0)^* F_1\| = \|\sin 2\ \Theta_0\| ,$$

and from theorem 3.2.2, it follows that

$$\delta\|\sin 2\ \Theta\| \le \|H - XHX\| \le \|H\| + \|XHX\| = 2\|H\|,$$

so that for all unitary invariant norms, we have

$$\delta\|\sin 2\ \Theta\| \le 2\ \|H\|.$$

But $\delta\|-E_0 \sin 2\ \Theta_0\ J_0^*E_1^* + E_1J_0 \sin 2\ \Theta_0\ E_0^*\|$

$$\le \|E_0B^*E_1^* + E_1BE_0^*\|$$

This implies that $\delta\|\sin 2\ \Theta_0\| \le 2\ \|B\| \le 2\ \|R\|$.

CHAPTER 4

Error Bounds for Approximate Invariant Subspaces of Closed Operators

In chapter 3 we showed that, given an invariant subspace of a self-adjoint operator and the corresponding invariant subspace of the perturbed operator, then we can find a bound for the difference between the two subspaces in terms of the magnitudes of the perturbation and of the gap between appropriate parts of the spectra, and we measure the difference between the two subspaces in terms of a nonnegative operator Θ. It was shown that the rotation is small if Θ is small (§1.3, §1.4) and Θ is small if the perturbation is small (§3.1, §3.2).

Here we extend the above results to the case of non-Hermitian matrices or more generally, to closed operators on a Hilbert space. The result for this case depends on a measure of the separation of the spectra of the two operators, and for Hermitian matrices or self-adjoint operators the distance between the spectra is an adequate measure (this being the one used in chapter 3). However, in the general case, the spectra and hence the distance between them may vary violently with small perturbations in the operators, and hence we need a more stable measure of the separation. This measure and its properties will be discussed in §4.2.

§4.1 The Class of Hilbert-Schmidt Operators

Definition 4.1.1

Let $\{x_\alpha, \alpha\varepsilon A\}$ *be a complete orthonormal set in the Hilbert space* $\mathcal{H}$. *A bounded linear operator* T *is said to be a Hilbert-Schmidt operator if the quantity* $\|T\|_{HS}$ *defined by the equation*

$$\|T\|_{HS} = \{\sum_{\alpha\varepsilon A} \|Tx_\alpha\|^2\}^{1/2} \text{ is finite.}$$

$\|T\|_{HS}$ *is called the Hilbert-Schmidt norm. The class of Hilbert Schmidt operators will be denoted by* $HS(\mathcal{H})$.

Lemma 4.1.2

The Hilbert-Schmidt norm is independent of the orthonormal basis used in its definition. If T *is in* $HS(\mathcal{H})$ *and* U *is unitary operator on* $\mathcal{H}$, *then* $U^{-1}TU$ *is in* $HS(\mathcal{H})$ *and*

$\|T\|_{HS} = \|U^{-1}TU\|_{HS}$. *In addition,* $\|T\|_{HS} \geq \|T\|$ *and*

$\|T\|_{HS} = \|T^*\|_{HS}$.

Proof

Let $\|T\|_A$, $\|T\|_B$ be the Hilbert Schmidt operator norm when defined in terms of different complete orthonormal systems $\{x_\alpha, \alpha\varepsilon A\}$, $\{x_\beta, \beta\varepsilon B\}$.

From $\|x\|^2 = \sum_\beta |(x,y_\beta)|^2$ we have $\|T\|_A^2 = \sum_\alpha \|Tx_\alpha\|^2 = \sum_\alpha \sum_\beta |(Tx_\alpha,y_\beta)|^2$

$$= \sum_\beta \sum_\alpha |(x_\alpha, T^*y_\beta)|^2 = \sum_\beta \|T^*y_\beta\|^2 = \|T^*\|_B^2.$$

If we take the same complete orthonormal set, we get

$$\|T\|_B^2 = \|T^*\|_B^2 = \|T\|_A^2 \text{ which implies that } \|T\|_B^2 = \|T\|_A^2.$$

If U is unitary operator, then the set $\{Ux_\alpha, \alpha \varepsilon A\}$ is also a complete orthonormal set, since $\|x\| = \|U^{-1}x\|$,

$$\|U^{-1} TU\|_{HS} = \sum_{\alpha\varepsilon A} \|U^{-1} TUx_\alpha\|^2 = \sum_{\alpha\varepsilon A} \|TUx_\alpha\|^2 = \|T\|_{HS}.$$

By definition, $\|T\| = \sup_{\|x\|=1} \|Tx\|$, so given $\varepsilon>0$, let x_0 be any unit vector such that $\|T\|^2 < \|Tx_0\|^2 + \varepsilon$.

Since there exists a complete orthonormal system containing x_0,

$\|T\|^2 \leq \sum_\alpha \|Tx_\alpha\|^2 + \varepsilon$; since $\varepsilon>0$ is arbitrary, we conclude $\|T\| \leq \|T\|_{HS}$.

An equivalent definition of the Hilbert-Schmidt norm is as follows:

Let $\{x_\alpha, \alpha\varepsilon A\}$ be any complete orthonormal system in $\mathcal{H}$. Then

$$\|T\|_{HS} = (\Sigma_{\alpha,\beta\varepsilon A}(|(Tx_\alpha, x_\beta)|^2))^{1/2}.$$

Since

$\|Tx_\alpha\|^2 = \sum_{\alpha\varepsilon A} |(Tx_\alpha, x_\beta)|^2$, the equivalence is obvious.

Theorem 4.1.3 [12]

The set $HS(\mathcal{H})$ of all Hilbert-Schmidt operators is a Banach Space under the Hilbert-Schmidt norm. In addition $HS(\mathcal{H})$ is an algebra with $\|TS\|_{HS} \leq \|T\|_{HS} \|S\|_{HS}$ for every $T, S\varepsilon HS(\mathcal{H})$.

Corollary 4.1.4

The set of Hilbert-Schmidt operators is a two-sided ideal in the Banach algebra of all bounded linear operators in a Hilbert Space $\mathcal{H}$. Moreover, if T is in $HS(\mathcal{H})$ and $B\varepsilon\mathcal{B}(\mathcal{H})$ then $\|TB\|_{HS} \leq \|T\|_{HS}\ \|B\|$ *and* $\|BT\|_{HS} \leq \|B\|\ \|T\|_{HS}$.

Proof.

Let $T\varepsilon HS(\mathcal{H})$, $B\varepsilon\mathcal{B}(\mathcal{H})$, then

$$\|BT\|_{HS}^2 = \sum_{\alpha\varepsilon A} \|BTx_\alpha\|^2 \leq \sum_{\alpha\varepsilon A} \|B\|^2\ \|Tx_\alpha\|^2 = \|B\|^2\ \|T\|_{HS}^2,$$

hence $BT\varepsilon HS(\mathcal{H})$.

On the other hand, $\|TB\|_{HS} = \|(TB)^*\|_{HS} = \|B^*T^*\|_{HS}$

$$\leq \|B^*\|\ \|T^*\|_{HS} = \|B\|\ \|T\|_{HS}$$

So $TB\varepsilon HS(\mathcal{H})$.

Theorem 4.1.5 [12]

Every Hilbert-Schmidt operator is compact and is the limit in the Hilbert-Schmidt norm of a sequence of operators with finite dimensional range.

Remark

Not every compact operator is in $HS(\mathcal{H})$, for example if $\{x_n\}$ is an orthonormal set in a separable Hilbert space and if T is determined by $Tx_n = n^{-1/2}\ x_n$ $n = 1,\ldots$ Then T is compact but $\sum_n \|Tx_n\|^2 = \sum_{n=1}^{\infty} \frac{1}{n}$ is not finite and hence T is not in $HS(\mathcal{H})$.

The class of Hilbert-Schmidt operators is a Banach algebra without identity. In addition $HS(\mathcal{H})$ is a Hilbert space with the inner product defined by

$$(S,T) = \sum_{\alpha} (Sx_\alpha, Tx_\alpha)$$

$$= \sum_{\alpha,\beta\varepsilon A} (Sx_\alpha, x_\beta)(x_\beta, Tx_\alpha)$$

(by the general Parseval relation, where $\{x_\alpha\}_{\alpha\varepsilon A}$ is a complete orthonormal system).

§4.2 The Separation of Two Operators

Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces. Let $B\varepsilon\mathcal{B}(\mathcal{X})$, $C\varepsilon\mathcal{B}(\mathcal{Y})$. Let $T\varepsilon\mathcal{B}[\mathcal{B}(\mathcal{X},\mathcal{Y})]$ defined by

$$T(P) = PB - CP \quad P\varepsilon\mathcal{B}(\mathcal{X},\mathcal{Y}).$$

Also let $\tau\varepsilon\mathcal{B}[HS(\mathcal{X},\mathcal{Y})]$ defined by

$$\tau(P) = PB - CP, \; P\varepsilon HS(\mathcal{X},\mathcal{Y}).$$

It was shown in theorem 2.2.8 that

$$\sigma(T) = \sigma(B) - \sigma(C) = \{\beta-\gamma: \; \beta\varepsilon\sigma(B), \; \gamma\varepsilon\sigma(C)\}.$$

Also it has been shown in theorem 2.2.5 that for $\lambda\varepsilon\rho(T)$,

$$(4.2.1) \qquad (T-\lambda I)^{-1}(Q) = \frac{1}{2\pi i}\int (zI-C)^{-1} Q \,(B-\lambda I-zI)^{-1} dz$$

$$= \frac{1}{2\pi i}\int R(z;\, C)\, Q\, R(\lambda+z;\, B)\, dz$$

where $R(z;\, C) = (zI-C)^{-1}$ and the integral is taken over a suitable contour.

Now we extend the above results when C is an unbounded operator. For that, let C be a closed operator on $\mathcal{Y}$ whose domain $\mathcal{Y}_C$ is dense in $\mathcal{Y}$. If $P \varepsilon \mathcal{B}(\mathcal{X},\mathcal{Y}_C)$, then the mapping $P \to PB - CP$ defines a linear operator

$T: \mathcal{B}(\mathcal{X},\mathcal{Y}_C) \to \mathcal{B}(\mathcal{X},\mathcal{Y})$, note: since CP is closed, defined on $\mathcal{X}$, then $CP \varepsilon \mathcal{B}(\mathcal{X},\mathcal{Y})$.

Theorem 4.2.1 [36]

$\sigma(T) = \sigma(B) - \sigma(C)$.

Proof

To prove this, it is clearly equivalent to prove that $0 \varepsilon \sigma(T)$ iff $\sigma(B) \cap \sigma(C) \neq \phi$. Suppose, $\sigma(B) \cap \sigma(C) = \phi$. Since $\sigma(B)$, $\sigma(C)$ are closed,and the complex plane is connected, we have $\rho(B) \cap \rho(C) \neq \phi$; this implies that there exists a point $\lambda \varepsilon \rho(B) \cap \rho(C)$. Let $Q \varepsilon \mathcal{B}(\mathcal{X},\mathcal{Y})$ and consider the equation

(4.2.2) $\quad T_\lambda(P) = PR(\lambda;B) - R(\lambda;C)P = R(\lambda;C)\,QR(\lambda;B).$

Since $\sigma(B) \cap \sigma(C) = \phi$ and $\lambda \varepsilon \rho(B) \cap \rho(C)$, so $\sigma(B-\lambda) \cap \sigma(C-\lambda) = \phi$ and hence $\sigma(R(\lambda;B)) \cap \sigma(R(\lambda;C)) = \phi$ which in turn implies that T_λ^{-1} exists as a bounded operator.

Moreover, if P satisfies (4.2.2), then $R(P) = \mathcal{Y}_C$, and if we postmultiply by $(\lambda I-B)$ and premultiply by $(\lambda I-C)$, we get

$$PB - CP = Q, \text{ that is } T(P) = Q$$

which implies that T has a bounded inverse, and

$P = T^{-1}(Q) = T_\lambda^{-1}(R(\lambda;\ C)QR(\lambda;\ B))$. Moreover,

$$\|T^{-1}\| = \sup_{\|Q\|=1} \|T^{-1}(Q)\| \le \|T_\lambda^{-1}\|\ \|R(\lambda;\ C)\|\ \|R(\lambda;\ B)\|$$

so that $0 \varepsilon \rho(T)$.

For the other implication, let $\lambda \varepsilon \sigma(B) \cap \sigma(C)$, then $0 \varepsilon \sigma(B-\lambda) \cap \sigma(C-\lambda)$, and since $T(P) = P(B-\lambda I) - (C-\lambda I)P$, we may assume without loss of generality that $\lambda = 0$, i.e. $0 \varepsilon \sigma(B) \cap \sigma(C)$, the proof is adapted from [28]. The spectrum of the operator C has the following subdivisions:

$$\sigma(C) = \sigma_p(C) \cup \sigma_c(C) \cup \sigma_r(C).$$

Here $\sigma_p(C)$ denotes the point spectrum, $\sigma_c(C)$ denotes the continuous spectrum, and $\sigma_r(C)$ denotes the residual spectrum. If $\lambda \varepsilon \sigma_p(C) \cap \sigma_c(C)$, then there is a sequence of unit vectors $y_i \varepsilon \mathcal{Y}_C$ such that $\|(\lambda I-C)y_i\| \to 0$, similarly for $\sigma(B)$. Now for $0 \varepsilon \sigma(B) \cap \sigma(C)$ and by the above subdivisions of $\sigma(B)$ and $\sigma(C)$, we have the following cases to consider. (The star denotes, for convenience, the Banach space adjoint).

(1) $0 \varepsilon \sigma_p(B^*) \cup \sigma_c(B^*)$, $0 \varepsilon \sigma_p(C) \cup \sigma_c(C)$.

Then there are sequences of unit vectors x_i, y_i such that $B^*x_i^* = x_i^*B \to 0$ and $Cy_i \to 0$. Let $P_i = y_i x_i^*$ then

$$\|P_i\| = \sup_{\substack{\|x\|=1 \\ x \varepsilon \mathcal{X}}} \|y_i\ x_i^*(x)\| = 1 \text{ and } T(P_i) = y_i(X_i^*B) - (Cy_i)x_i^*.$$

Now, $\|T(P_i)\| \le \|x_i^*B\| + \|Cy_i\| \to 0$, so that $0 \varepsilon \sigma(T)$.

(2) $0 \in \sigma_r(B^*)$, $0 \in \sigma_p(C) \cup \sigma_c(C)$

For $0 \in \sigma_r(B^*)$ we have $0 \in \sigma_p(B)$ (since $0 \in \sigma_r(B^*)$ imply $\overline{R(B^*)} \neq \mathcal{X}$, which imply that $N(B) \neq \{0\}$). So we have unit vectors x, y_i such that $Bx = 0$ and $Cy_i \to 0$. If $0 \notin \sigma(T)$ (note $y_i x^* \in \mathcal{B}(\mathcal{X},\mathcal{Y})$) it follows that there are $P_i \in \mathcal{B}(\mathcal{X},\mathcal{Y}_C)$ such that $T(P_i) = y_i x^*$. Now $CP_i = P_iB - y_ix^*$ and $C^2P_i = CP_iB - Cy_ix^*$, which implies in turn that $CP_i \in (\mathcal{X},\mathcal{Y}_C)$ and C^2P_i is bounded. It follows that $T(CP_i) = CP_iB - C^2P_i = CT(P_i) = Cy_ix^* \to 0$, so that $CP_i \to 0$. But, $1 = y_i^*(y_ix^*)x = y_i^*T(P_i)x = y_i^*(P_iB - CP_i)x = y_i^*CP_ix \to 0$, a contradiction.

(3) $0 \in \sigma_p(B^*) \cup \sigma_c(B^*)$, $0 \in \sigma_r(C)$. This goes similar to (2) and implies $0 \in \sigma(T)$.

(4) $0 \in \sigma_r(B^*)$, $0 \in \sigma_r(C)$. Let x,y be unit vectors such that $Bx = 0$ and $y^*C = 0$. If $0 \notin \sigma(T)$, then there is a $P \in \mathcal{B}(\mathcal{X},\mathcal{Y}_C)$ such that $T(P) = yx^*$. But then $1 = y^*T(P)x = y^*(PB-CP)x = 0$, a contradiction. The proof is complete.

The same result holds for the operator τ:

<u>Theorem 4.2.2</u>

$$\sigma(\tau) = \sigma(B) - \sigma(C)$$

<u>Proof</u>

The same as in theorem 4.2.1.

After we extended theorem 2.2.3 to the case where C is a closed linear operator, we try to find a measure of the separation between the two operators. Now in case $\sigma(B) \cap \sigma(C) = \phi$, that is, $0 \notin \sigma(T)$, T^{-1} exists, and

$$\|T^{-1}\| \geq \sup_{\lambda \varepsilon \sigma(T)} \left|\frac{1}{\lambda}\right|,$$

so that $0 < \|T^{-1}\|^{-1} \leq \operatorname{Inf}_{\lambda\varepsilon\sigma(T)} |\lambda| = \operatorname{Inf} |\sigma(B) - \sigma(C)|$

$$= \operatorname{Inf}\{|B-\gamma|, \beta\varepsilon\sigma(B), \gamma\varepsilon\sigma(C)\}.$$

Definition 4.2.3

$$sep(B,C) = \begin{cases} \|T^{-1}\|^{-1} & if\ 0\notin\sigma(T) \\ 0 & if\ 0\varepsilon\sigma(T) \end{cases}$$

$$sep_{HS}(B,C) = \begin{cases} \|\tau^{-1}\|^{-1} & if\ 0\notin\sigma(\tau) \\ 0 & if\ 0\varepsilon\sigma(\tau) \end{cases}$$

Theorem 4.2.4

The separation of B and C satisfies the inequality

(4.2.3) $\quad sep(B,C) \leq \operatorname{Inf} |\sigma(B) - \sigma(C)|$, *and if* $sep(B,C) \neq 0$, *then*

$$sep(B,C) = \operatorname{Inf}_{\|P\|=1} \|T(P)\|$$

The Hilbert-Schmidt separation also satisfies (4.2.3) and if $sep_{HS}(B,C) \neq 0$ *then* $sep_{HS}(B,C) = \operatorname{Inf}_{\|P\|_{HS}=1} \|\tau(P)\|_{HS}$.

Proof

As we showed before, if $\sigma(B) \cap \sigma(C) = \phi$ then

$$\|T^{-1}\|^{-1} \leq \operatorname{Inf} |\sigma(B) = \sigma(C)|$$

Similarly $\|\tau^{-1}\|^{-1} \leq \operatorname{Inf} |\sigma(B) - \sigma(C)|$, and hence inequality (4.2.3) follows from definition 4.2.3.

But $\operatorname{Inf}_{\|P\|=1} \|T(P)\| = \|T^{-1}\|^{-1}$ if T is invertible, and

$$\operatorname{Inf}_{\|P\|_{HS}=1} \|\tau(P)\| = \|\tau^{-1}\|^{-1} \text{ if } \tau \text{ is invertible}$$

So it follows that if sep(B,C) $\neq$ 0, then

$$\text{sep}(B,C) = \operatorname{Inf}_{\|P\|=1} \|T(P)\|$$

$$\text{sep}_{HS}(B,C) = \operatorname{Inf}_{\|P\|_{HS}=1} \|\tau(P)\|_{HS}$$

The reason for using sep(B,C) as the measure of separation of the spectra of B & C, is that it is insensitive to small perturbations in B and C, as shown by the following theorem:

Theorem 4.2.5

If $E\varepsilon\mathcal{B}(\mathcal{X})$ *and* $F\varepsilon\mathcal{B}(\mathcal{Y})$, *then*

$$sep(B-E,\ C-F) \geq sep(B,C) - \|E\| - \|F\| \qquad \text{and}$$

$$sep_{HS}(B-E,\ C-F) \geq sep_{HS}(B,C) - \|E\| - \|F\|.$$

Proof

The proof is the same for sep and sep_{HS}, so we prove it for sep_{HS}.

If $\text{sep}_{HS}(B,C) - \|E\| - \|F\| \leq 0$, then the theorem is true since $\text{sep}_{HS}(B-E, C-F) \geq 0$. Now we suppose that $\text{sep}_{HS}(B,C) - \|E\| - \|F\| > 0$, that is $\text{sep}_{HS}(B,C) > \|E\| + \|F\|$.

Let $V\varepsilon\mathcal{B}[HS(\mathcal{X},\mathcal{Y})]$ be defined by $V(P) = PE - FP$

$$\|V\| = \sup_{\|P\|_{HS}=1} \|V(P)\|_{HS} \le \|E\| + \|F\| \quad \text{(by corollary 4.1.4)}$$

So $\|V\tau^{-1}\| \le \|V\| \; \|\tau^{-1}\| \le \dfrac{\|E + F\|}{\mathrm{sep}_{HS}(B,C)} < 1.$

Hence $(V\tau^{-1})$ is invertible and

$$\|(I-V\tau^{-1})^{-1}\| \le (1 - \|V\tau^{-1}\|)^{-1}.$$

But $$(\tau-V)^{-1} = \tau^{-1} \, (I-V\tau^{-1})^{-1},$$

which implies that $(\tau-V)^{-1}$ is bounded. But $\mathrm{sep}_{HS}(B-E, C-F) = \|(\tau-V)^{-1}\|^{-1}$,

$$\begin{aligned}\mathrm{sep}_{HS}(B-E, C-F) &= \|(\tau-V)^{-1}\|^{-1} = \|\tau^{-1}(I-V\tau^{-1})^{-1}\|^{-1} \\ &\ge \|\tau^{-1}\|^{-1} \; \|(I-V\tau^{-1})^{-1}\|^{-1} \\ &\ge \mathrm{sep}_{HS}(B,C) \; (1 - \|V\tau^{-1}\|) \\ &\ge \mathrm{sep}_{HS}(B,C) - \mathrm{sep}_{HS}(B,C)\|V\tau^{-1}\| \\ &\ge \mathrm{sep}_{HS}(B,C) - \|E\| - \|F\| .\end{aligned}$$

The importance of sep_{HS} rests in extra properties not satisfied by sep. We list some of the properties of sep_{HS}.

For proofs and more properties of sep_{HS} we refer to [36].

1. Let $\mathcal{Y} = \mathcal{Y}_1 \oplus \mathcal{Y}_2 \oplus \ldots \oplus \mathcal{Y}_m$, S_i the projector onto $\mathcal{Y}_i$ such that $S_iC = CS_i$, so we can write

   $C = C_1 \oplus C_2 \oplus \ldots \oplus C_m$ where C_i is the restriction of C to $S_i(\mathcal{Y}_C)$;

   Similarly, let $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \ldots \oplus \mathcal{X}_n$ where W_i is the projector on $\mathcal{X}_i$ and $W_iB = BW_i$ $i = 1,\ldots,n$. Then we have

$$\mathrm{sep}_{HS}(B_1 \oplus B_2 \oplus \dots \oplus B_n, C_1 \oplus C_2 \oplus \dots \oplus C_m)$$
$$= \min \{\mathrm{sep}_{HS}(B_i, C_j): \; i=1,2,..,n, \; j=1,2,..,m\}.$$

(2) If $B \in \mathcal{B}(\mathcal{X})$ and $C \in \mathcal{B}(\mathcal{Y})$, then

$$\mathrm{sep}_{HS}(B,C) = \mathrm{sep}_{HS}(C,B).$$

(3) If B and C are selfadjoint then $\mathrm{sep}_{HS}(B,C) = \mathrm{Inf}|\sigma(B) - \sigma(C)|$.

§4.3 The Error Bounds

Let A be a closed linear operator defined on a separable Hilbert space $\mathcal{H}$ whose domain $\mathfrak{D}(A)$ is dense in $\mathcal{H}$. Let $\mathcal{X} \subset \mathfrak{D}(A)$ be a subspace, let $\mathcal{Y}$ be the orthogonal complement of $\mathcal{X}$. Let $\mathcal{Y}_A$ be the projection of $\mathfrak{D}(A)$ into $\mathcal{Y}$.

We note that the linear manifold $\mathcal{Y}_A$ is contained in $\mathfrak{D}(A)$ and is dense in $\mathcal{Y}$. Because $y \in \mathcal{Y}_A$ implies $y = z - x$ for some $z \in \mathfrak{D}(A)$ and $x \in \mathcal{X} \subset \mathfrak{D}(A)$ which implies that $y \in \mathfrak{D}(A)$.

Since $\mathfrak{D}(A)$ is dense in $\mathcal{H}$; $\mathcal{Y}_A$ is dense in $\mathcal{Y}$.

Let X, Y and Y_A be the insertions of $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Y}_A$ into $\mathcal{H}$ respectively. (note that X,Y are isometrics).

Theorem 4.3.1

Let $P \in \mathcal{B}(\mathcal{X}, \mathcal{Y}_A)$. Let

$$X' = (X + Y_A P)(I + P^*P)^{-1/2}$$

$$Y_A' = (Y_A - XP^*)(I + PP^*)^{-1/2}$$

Let $\mathcal{X}' = R(X')$ and $\mathcal{Y}_A' = R(Y_A')$, then

(i) *X' and Y_A' are isometries*

(ii) *$\mathcal{X}' \subset \mathfrak{H}(A)$ is a subspace,*

(iii) *$\mathcal{Y}_A'$ is the projection of $\mathfrak{D}(A)$ onto the orthogonal complement of $\mathcal{X}'$,*

(iv) *the subspace $\mathcal{X}'$ is an invariant subspace of A iff $Y_A'^{*} AX' = 0$.*

<u>Proof</u>

(i) To prove that X' and Y_A' is an isometry, it is enough to prove that $X'^{*}X' = I$.

But $X'^{*}X' = (I + P^{*}P)^{-1/2} (X^{*} + P^{*} Y_A^{*}) (X + Y_A P) (I + P^{*}P)^{-1/2}$

Since $\mathcal{Y}_A \subset \mathcal{Y}$ and $R(X) = \mathcal{X}$, and since $N(X^{*}) = \mathcal{Y}$, we have $X^{*}Y_A = 0$ and $Y_A^{*} X = 0$. Als0 $X^{*}X = I, Y_A^{*} Y_A = I_A$ (The identity on $\mathcal{Y}_A$). Consequently, $X'^{*}X' = (I + P^{*}P)^{-1/2} (I + P^{*}P) (I + P^{*}P)^{-1/2} = I$. Similarly we can show that $Y_A'^{*} Y_A' = I$.

(ii) Since $X': \mathcal{X} \to \mathcal{H}$ and $R(X') = \mathcal{X}'$, the set $\mathcal{X}'$ is a subspace since $\mathcal{X}$ is, and X' is an isometry,

$$\mathcal{X}' = R(X') = R(X + Y_A P) \subset R(X) + R(Y_A P) \subset \mathcal{X} + \mathcal{Y}_A \subset \mathfrak{H}(A).$$

(iii) First we note that

$$Y_A' : \mathfrak{D}(Y_A') \to \mathcal{H}$$
$$\mathfrak{D}(Y_A') = (I + P^{*}P)^{+1/2} \mathcal{Y}_A.$$

Let Q' be the projection operator on $\mathcal{X}'$. Then $(I - Q')$ is the projection operator on $\mathcal{Y}'$, the orthogonal complement of $\mathcal{X}'$. Since $Y_A'^{*} X' = 0$, this implies that $\mathcal{Y}_A' \subset \mathcal{Y}'$.

To prove that $(I - Q')\mathfrak{D}(A) = \mathcal{Y}'_A$, we first note that clearly $\mathcal{Y}'_A \subset (I - Q')\mathfrak{D}(A)$. On the other hand, let $y' \varepsilon (I - Q')\mathfrak{D}(A)$; by the previous remark, $(I - Q')\mathfrak{D}(A) \subset \mathfrak{D}(A)$ and $(I - Q')\mathfrak{D}(A)$ is dense in $\mathcal{Y}'$ which implies that $y' \varepsilon \mathfrak{D}(A)$. Now $y' = x + y$ $x \varepsilon \mathcal{X}$, $y \varepsilon \mathcal{Y}_A$ but since $y' \varepsilon (I - Q')\mathfrak{D}(A)$ this implies that $X'^*y' = 0$, which implies

$$(I + P^*P)^{-1/2}(X^* + P^* y_A)\, y' = 0,$$

$$(I + P^*P)^{-1/2}(x + P^* y) = 0,$$

which implies $x = -P^* y$ $\quad y' = y + x = (Y_A - XP^*)y$, which implies $y' = Y'_A (I + PP^*)^{1/2} y$,

So finally, $y' \varepsilon \mathcal{Y}'_A$ and $(I - Q')\mathfrak{D}(A) = \mathcal{Y}'_A$.

(iv) $\mathcal{Y}'_A$ is the projection of $\mathfrak{D}(A)$ into $\mathcal{Y}'$, so that by the previous remark $\mathcal{Y}'_A \subset \mathfrak{D}(A)$ and $\mathcal{Y}'_A$ is dense in $\mathcal{Y}'$,

So $Y'^*_A AX' = 0$ iff $A\mathcal{X}' \subset \mathcal{X}'$.

<u>Lemma 4.3.2</u>

The operator $AY'_A : \mathfrak{D}(Y'_A) \to \mathcal{H}$ is closed.

<u>Proof</u>

Let $z_n \to z$ $\qquad z_n \varepsilon \mathfrak{D}(Y'_A)$

and $AY'_A z_n \to h$. We will show that $AY'_A z = h$.

Let $y'_n = Y'_A z_n$ where $y'_n \varepsilon \mathcal{Y}'_A \subset \mathfrak{D}(A)$.

Since Y'_A is an isometry, the sequence $y'_n \to y' \varepsilon \overline{\mathcal{Y}'_A}$.

Since A is closed, $Ay'_n \to h$, $y'_n \to y'$, hence $y' \varepsilon \mathfrak{D}(A)$ and $Ay' = h$. The fact that $y' \varepsilon \mathfrak{D}(A)$ and $y' \varepsilon \overline{\mathcal{Y}'_A}$ implies that $y' \varepsilon \mathcal{Y}'_A$. So, $y' = Y'_A \bar{z}$ for some $\bar{z} \varepsilon \mathfrak{D}(Y'_A)$.

By assumption, $z_n \to z$, which implies $Y'_A z_n \to Y'_A z$.

Since Y_A' is an isometry, it follows that z = z, $z \in \mathfrak{D}(Y_A')$

$\mathcal{X}$ will be an invariant subspace of A iff $G = Y_A^* AX = 0$, so if G is small, then $\mathcal{X}$ is hopefully near an invariant subspace of A. We will show in the next theorem that, under certain conditions, there exists an isometry $X': \mathcal{X} \to \mathcal{H}$ such that R(X') is an invariant subspace of A and $\|X-X'\|$ tends to zero as G tends to zero.

Theorem 4.3.3 [36]

Let $A: \mathfrak{D}(A) \to \mathcal{H}$ *be a closed linear operator with domain* $\mathfrak{D}(A)$ *dense in* $\mathcal{H}$. *Let* $\mathcal{X} \subset \mathfrak{D}(A)$ *be a subspace and* $\mathcal{Y}_A$ *the projection of* $\mathfrak{D}(A)$ *onto the orthogonal complement of* $\mathcal{X}$. *Let* X, *and* Y_A *be the injections of* $\mathcal{X}$, $\mathcal{Y}_A$ *into* $\mathcal{H}$, *respectively and let*

$$B = X^*AX, \qquad H = X^*AY_A,$$

$$G = Y_A{}^*AX, \qquad C = Y_A{}^*AY_A.$$

Set

$$\gamma = \|G\|, \quad \eta = \|H\|, \quad \delta = sep(B,C).$$

Then if

$$\kappa_1 = \gamma\eta/\delta^2 < 1/4 \tag{4.3.1}$$

then there is a $P \in \mathcal{B}(\mathcal{X},\mathcal{Y}_A)$ *satisfying*

$$\|P\| \le \frac{\gamma}{\delta}(1+\kappa) = \frac{\gamma}{\delta}\,\frac{1+\sqrt{1-4\kappa_1}}{1-2\kappa_1+\sqrt{1-4\kappa_1}} < 2\gamma/\delta. \tag{4.3.2}$$

such that $R(X + Y_AP)$ *is an invariant subspace of* A. *Moreover,* $\sigma(A)$ *is the disjoint union*

$$(4.3.3) \qquad \sigma(A) = \sigma(B + HP) \cup \sigma(C - PH).$$

Proof

Since B: $\mathcal{X} \to \mathcal{X}$, $\mathcal{X}$ is a subspace and A is closed then it follows that B is bounded; on the other hand, C is closed, so that δ is well defined. X', Y'_A are as before, so according to theorem 4.3.1, $\mathcal{X}'$ is an invariant subspace iff $G' = Y'_A AX' = 0$. We can calculate

$$G' = (I+PP^*)^{-1/2}(Y_A^* - PX^*)\ A(X+Y_AP)(I+P^*P)^{-1/2}$$
$$= (I+PP^*)^{-1/2}\ (CP - PB+G - PHP)(I+P^*P)^{-1/2}.$$

$$(4.3.4) \qquad T(P) = PB - CP = G - PHP.$$

Since $\delta>0$, T^{-1} exists and $\|T^{-1}\| = 1/\delta$.

To solve (4.3.4) by for P, we solve it by successive substitutions. Let

$$(4.3.5) \qquad P_0 = T^{-1}(G) \text{ so } \|P_0\| \le \|T^{-1}\|\ \|G\| = \gamma/\delta = \pi_0$$

Now given P_i, define P_{i+1} as follows:

$$(4.3.6) \qquad P_{i+1} = T^{-1}\ (G-P_iHP_i) = T^{-1}(G) - T^{-1}(P_iHP_i)\ \ i \ge 0$$
$$= P_0 - T^{-1}(P_iHP_i)$$

From (4.3.6), if $\|P_i\| \le \pi_i$ then

$$\|P_{i+1}\| \le \|P_0\| + \|T^{-1}\|\ \|P_iHP_i\|$$
$$\le \pi_0 + \delta^{-1}\eta\pi_i^2 = \pi_{i+1}.$$

Now π_i can be written as follows:

$$\pi_i = \pi_0(1 + \kappa_i) \text{ , where}$$

$$\kappa_1 = \pi_1/\pi_0 - 1 = \frac{\pi_0 + \delta^{-1} \eta \pi_0^2}{\pi_0} - 1 = \delta^{-1} \eta \pi_0.$$

Hence $\pi_{i+1} = \pi_0(1 + \kappa_{i+1}) = \pi_0 + \delta^{-1} \eta \pi_i^2$

$$= \pi_0 + \delta^{-1} \eta \pi_0^2 (1 + \kappa_i)^2,$$

which implies that κ_i has the recursion

$$\kappa_{i+1} = \kappa_1 (1 + \kappa_i)^2.$$

To find the limit of the numbers κ_i we solve the equations

$y = \kappa_1(1 + x)^2$ and $y = x$, then we have two roots r_1, r_2 given by

$$r_{1,2} = \frac{(1 - 2\kappa_1) \mp \sqrt{1 - 4\kappa_1}}{2\kappa_1} = \frac{2\kappa_1}{1 - 2\kappa_1 \mp \sqrt{1 - 4\kappa_1}}$$

Condition (4.3.1) guarantees that r_1, r_2 exist; also since $2\kappa_1(1+x)<1$, $x \varepsilon [0,1)$, the numbers κ_i will converge to

$$r_1 = \frac{2\kappa_1}{1 - 2\kappa_1 + \sqrt{1 - 4\kappa_1}} < 1.$$

so $$= \lim_i \kappa_i = \frac{2\kappa_1}{1 - 2\kappa_1 + \sqrt{1 - 4\kappa_1}} < 1$$

$$\text{and } \sup_{i \geq 0} \|P_i\| \leq \lim_{i \to \infty} \pi_i = \pi_0 (1 + \kappa),$$

so that the sequence $\{\|P_i\|\}$ is bounded.

To show that the iteration defined by (4.3.6) converges we show that the P_i converge.

Let $D_i = P_{i+1} - P_i$, then

$$\|D_i\| = \|P_{i+1} - P_i\| \le \delta^{-1} \|P_i H P_i - P_{i-1} H P_{i-1}\|$$
$$= \delta^{-1} \|P_i H P_i - P_{i-1} H P_i + P_{i-1} H P_i - P_{i-1} H P_{i-1}\|$$
$$= \delta^{-1} \|D_{i-1} H P_i + P_{i-1} H D_{i-1}\|$$
$$\le \delta^{-1} \|H\| \, \|D_{i-1}\| \, (\|P_i\| + \|P_{i-1}\|)$$
$$\le 2\delta^{-1} \|H\| \, \|P_i\| \, \|D_{i-1}\|$$
$$\le 2\delta^{-1} \eta \, \pi_i \, \|D_{i-1}\| \le 2\kappa_1 (1+\kappa_i) \, \|P_{i-1}\|.$$

We find that $\overline{\lim_{i\to\infty}} \dfrac{\|D_i\|}{\|D_{i-1}\|} \le 2\kappa_1(1+\kappa)$ unless D_i terminates at 0; in either case, $\sum_i \|D_i\| < \infty$ provided $2\kappa_1(1+\kappa)\kappa<1$, which is true since $\kappa<1$, $\kappa_1<1/4$.

So $\sum_{i=0}^{\infty} \|P_{i+1} - P_i\| = 0$, which implies that the iteration converges. Hence $P_i \to P$, $P \varepsilon \overline{\mathcal{B}(\mathcal{X},\mathcal{Y}_A)}$; but since $P = T^{-1}(G - AHP)$ where $T : \mathcal{B}(\mathcal{X},\mathcal{Y}_A) \to \mathcal{B}(\mathcal{X},\mathcal{Y})$, it follows $P \varepsilon \mathcal{B}(\mathcal{X},\mathcal{Y}_A)$ and

$$\|P\| \le \pi_0(1+\kappa) = \gamma/\delta \, \frac{1 + \sqrt{1-4\kappa_1}}{1-2\kappa_1 + \sqrt{1-4\kappa_1}} < 2\gamma/\delta.$$

Now to prove the statement about the spectrum of A, let Y' be the extension of Y_A' to $\mathcal{Y}'$. Then the transformation

$$U = X'X^* + Y'Y^* \text{ satisfies}$$

$$U^*U = UU^* = I \text{ and } U\mathfrak{H}(A) = \mathfrak{H}(A).$$

Hence if $A' = U^* AU$ then $\sigma(A) = \sigma(A')$.
With respect to the decomposition $\mathcal{H} = \mathcal{X} \oplus \mathcal{Y}$, the operator A' has the representation

$$A' = (X \quad Y) \begin{pmatrix} B' & H' \\ G' & C' \end{pmatrix} \begin{pmatrix} X^* \\ Y^* \end{pmatrix}$$

where $B' = X^*A'X$, $H' = X^*A'Y_A$, $G' = Y_A^* A'X$, $C' = Y_A^* A'Y_A$.
But $A' = (X\ X'^* + Y\ Y'^*)\ A\ (X'\ X^* + Y'\ Y^*)$.

So it follows that
$B' = X'^*\ A\ X'$, $H' = X'^*\ A\ Y_A'$, $C' = Y_A'^*\ A\ Y_A'$, $G' = Y_A'^*\ AX = 0$, the last equality holds since $\mathcal{X}'$ is an invariant subspace.
Therefore

$$A' \cong \begin{pmatrix} B' & H' \\ 0 & C' \end{pmatrix},$$

So that if $\lambda \varepsilon \rho(B') \cap \rho(C')$, then $R(\lambda;\ A')$ has the representation

$$R(\lambda;\ A') \cong \begin{pmatrix} R(\lambda;\ B') & R(\lambda;\ B')\ H'\ R(\lambda, C') \\ 0 & R(\lambda;\ C') \end{pmatrix}$$

Consequently $R(\lambda; A')$ is bounded if $R(\lambda; B')\ H\ R(\lambda; C')$ is bounded.

Since by Lemma 4.3.2 AY'_A is closed and $R(\lambda; C') \in \mathcal{B}(\mathcal{X}, \mathcal{Y}_A)$. it follows that $AY'_A R(\lambda; C')$ is bounded. Hence $R(\lambda; B')\ X'^*\ AY'_A\ R(\lambda; C') = R(\lambda; B') H'\ R(\lambda; C')$ is bounded.

This proved that $\sigma(A) = \sigma(A') \subseteq \sigma(B') \cup \sigma(C')$; the above representation also shows that the reverse inclusion, hence equality, holds.

From $B' = X'^* AX'$ we deduce

$$\begin{aligned} B' &= (I+P^*P)^{-1/2}\ (X^*+P^*Y_A^*)\ A(X+Y_AP)\ (I+P^*P)^{-1/2} \\ &= (I+P^*P)^{-1/2}\ (B+P^*G+HP+P^*CP)\ (I+P^*P)^{-1/2}. \end{aligned}$$

Since G satisfies (4.3.4), we have $P^*G = P^*PB - P^*CP+P^*PHP$. Hence $B' = (I+P^*P)^{1/2}\ (B+HP)\ (I+P^*P)^{-1/2}$, and $\sigma(B') = \sigma(B+HP)$

Also $C' = Y'^*_A\ A\ Y'_A$,

$$\begin{aligned} C' &= (I+PP^*)^{-1/2}\ (Y_A^*-PX^*)\ A\ (Y_A-XP^*)\ (I+PP^*)^{-1/2} \\ &= (I+PP^*)^{-1/2}\ (C-GP^*-PH-PBP^*)\ (I+PP^*)^{-1/2}. \end{aligned}$$

But from (4.3.4), $GP^* = PBP^* - CPP^* + PHPP^*$, so that $C' = (I+P^*P)^{1/2}\ (C-PH)\ (I+PP^*)^{-1/2}$,

$\sigma(C') = \sigma(C-PH)$. Consequently $\sigma(A') = \sigma(B+HP) \cup \sigma(C-PH)$.

Finally, since $\|HP\| \leq \|H\|\ \|P\| \leq 2\eta\gamma/\delta$, we conclude that

$$\begin{aligned} \mathrm{sep}(B+HP,\ C-PH) &\geq \mathrm{sep}(B,C) - \|HP\| - \|PH\| \quad \text{(Theorem 4.2.5)} \\ &\geq \delta - 4\ \gamma/\delta = \frac{\delta^2 - 4\ \gamma\eta}{\delta} > 0 \text{ (by (4.3.1))}, \end{aligned}$$

So that $\sigma(B+HP) \cap \sigma(F-PH) = \phi$.

CHAPTER 5

Algorithms

In this chapter, we discuss how to compute the direct rotation U, if we are given two subspaces of a Hilbert space, or equivalently two ortho projectors P and Q. We also discuss here how to compute the angles between the subspaces. These quantities are of interest in many applications, as in statistics [7], the generalized eigenvalue problem [15] and in the computation of invariant subspaces of matrices [40].

§5.1 Definition and Properties of the Bisector of P and Q

Let $\mathcal{H}$ be a Hilbert space, and let $P\mathcal{H}$ and $Q\mathcal{H}$ be two subspaces satisfying

$$(5.1.1) \qquad \begin{cases} \dim P\mathcal{H} = \dim Q\mathcal{H}, \\ \dim (I-P)\mathcal{H} = \dim (I-Q)\mathcal{H}. \end{cases}$$

From theorem 1.3.4, we recall that the direct rotation exists if and only if $P\mathcal{H}$ and $Q\mathcal{H}$ are equivalently positioned, i.e.

$$(5.1.2) \qquad \dim P\mathcal{H} \cap (I-Q)\mathcal{H} = \dim (I-P)\mathcal{H} \cap Q\mathcal{H}.$$

We can show that $P\mathcal{H}$ and $Q\mathcal{H}$ can be decomposed as follows.

$$P\mathcal{H} = PQ\mathcal{H} \oplus (P\mathcal{H} \cap (I-Q)\mathcal{H}),$$
$$Q\mathcal{H} = QP\mathcal{H} \oplus ((I-P)\mathcal{H} \cap Q\mathcal{H}).$$

Thus, it follows that if $\mathcal{H}$ is a unitary space, then equations (5.1.1) and (5.1.2) are equivalent. We should also remark that, even if $\mathcal{H}$ is infinite dimensional, these two equations will still remain equivalent, provided either P or (I-P) is a finite dimensional projector.

Using the notation adopted in chapter 1, we have

$P = E_0 E_0^*$, $\quad Q = F_0 F_0^*$, and

$$(5.1.3) \qquad U = [QP + (I-Q)(I-P)]\,[I - (P-Q)^2]^{-1/2}$$

(whenever the inverse is bounded), or in terms of the decomposition of $\mathcal{H}$, into $P\mathcal{H}$ and $(I-P)\mathcal{H}$,

$$U \simeq \begin{pmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{pmatrix}$$

Let

$$(5.1.4) \qquad T = T(P,Q) = [I - (P-Q)^2]^{-1/2}\,(P+Q-I).$$

It follows from equation (5.1.3) that

$$T = U\,(2P - I),$$

or equivalently

$$T \simeq \begin{pmatrix} C_0 & S_0^* \\ S_0 & -C_1 \end{pmatrix}$$

It is easy to check that $T^* = T$, $T^2 = I$ and $TP = QT$, so that T is an involution exchanging $P\mathcal{H}$ with $Q\mathcal{H}$.

We define the bisector of P and Q [8, 26] by

$$Z = Z(P,Q) = \frac{1}{2}(I + T);$$

This is the projector on a subspace, which may be named the angle bisector of $P\mathcal{H}$ and $Q\mathcal{H}$.

Remark 1

Since in the 2-dimensional space the angle bisector is not unique, but the one defined above is unique, it will be the bisector of the acute angle as we will show in theorem 5.1.1.

Remark 2

In the acute case, T will be unique, but in the non-acute case, with equation (5.1.2) satisfied, we define T on

$$(P\mathcal{H} \cap (I-Q)\mathcal{H}) \cup (Q\mathcal{H} \cap (I-P)\mathcal{H})$$

as an involution exchanging $P\mathcal{H} \cap (I-Q)$ with $Q\mathcal{H} \cap (I-P)\mathcal{H}$.

Theorem 5.1.1

In the acute case, T(P,Q) is the unique involution, satisfying

(i) $TP = QT$

(ii) $PTP \geq 0$

Proof

Clearly T(P,Q), as defined by equation (5.1.4), satisfies (i), and since

$$PTP \simeq \begin{pmatrix} C_0 & 0 \\ 0 & 0 \end{pmatrix} \geq 0$$

then (ii) is also satisfied.

To prove uniqueness, let

$$W = \begin{pmatrix} T_{00} & T_{01} \\ T_{10} & -T_{11} \end{pmatrix}$$

be an involution, satisfying (i) and (ii). Thus, we have the following relations between the entries T_{ij} of W:

$$T_{00}^2 + T_{10}^* T_{10} = I,$$

(5.1.5)
$$T_{00} T_{10}^* - T_{10}^* T_{11} = 0,$$
$$T_{10} T_{10}^* + T_{11}^2 = I.$$

From the assumptions, we have $T_{00} \geq 0$, $T_{11} \geq 0$ and $Q = WPW$,

i.e.
$$\begin{pmatrix} C_0^2 & C_0 S_0^* \\ S_0 C_0 & S_0 S_0^* \end{pmatrix} = \begin{pmatrix} T_{00}^2 & T_{00} T_{10}^* \\ T_{10} T_{00} & T_{10} T_{10}^* \end{pmatrix}.$$

Thus, we have

$C_0^2 = T_{00}^2$, which implies that $C_0 = T_{00}$ since $T_{00} \geq 0$

$C_0 S_0^* = T_{00} T_{10}^*$ i.e. $S_0 C_0 = T_{10} T_{00} = T_{10} C_0$

which implies that S_0 and T_{10} agree on $R(C_0)$. But in the acute case, $R(C_0)$ is dense, and hence $S_0 = T_{10}$.

From equations (5.1.5) we have,

$$T_{11}^{\ 2} = I - T_{10}T_{10}^* = I - S_0S_0^* = C_1^{\ 2}$$

which implies that $T_{11} = C_1$. This proves the theorem.

Remark

We should point out that if $P\mathcal{H}$ and $Q\mathcal{H}$ are in the acute case, then $Z\mathcal{H}$ and $P\mathcal{H}$ will also be in the acute case, (otherwise, on $(P\mathcal{H} \cap (I-Z)\mathcal{H}) \cup ((I-P)\mathcal{H} \cap Z\mathcal{H})$, we will have $P+Z-I = 0$, so that $T = I - 2P$ and $PTP = -P$ which contradicts $PTP \geq 0$). So there exists a unique direct rotation mapping $P\mathcal{H}$ onto $Z\mathcal{H}$, which we denote $U(P,Z)$. Let the corresponding angle operator be Φ, so we have the following theorem which generalizes the facts in the 2-dimensional case.

Theorem 5.1.2

If $P\mathcal{H}$ and $Q\mathcal{H}$ are in the acute case, and $Z\mathcal{H}$ is the angle bisector, then

(i) $\cos^2 \Phi = \frac{1}{2}(1 + \cos \Theta)$,

(ii) $[U(P,Z)]^2 = U(P,Q)$.

Proof

We have $\cos^2\Phi = PZP + (I-P)(I-Z)(I-P)$ and

$$\cos^2\Phi \simeq \frac{1}{2}\begin{pmatrix}1 + C_0 & 0\\ 0 & 0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0 & 0\\ 0 & 1 + C_1\end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}\cos\Theta_0 & 0\\ 0 & \cos\Theta_1\end{pmatrix},$$

So that

$$\cos^2\phi = \frac{1}{2}(1 + \cos\Theta).$$

(ii) $U^2(P,Z) = (2Z\ I)(2P-I) = T(2P-I) = U(P,Q).$

Remark

The inequality $PTP \geq 0$ implies that $QTQ = TPTPT \geq 0$, and hence $Z\mathcal{H}$ and $Q\mathcal{H}$ are in the acute case, by the same argument as in the case of $P\mathcal{H}$ and $Z\mathcal{H}$. Now, let $U(Z,Q)$ be the direct rotation mapping $Z\mathcal{H}$ onto $Q\mathcal{H}$, then

$$U^2(Z,Q) = U(P,Q).$$

§5.2 An Economical Expression for U.

For simplicity, we assume that $\dim \mathcal{H} = n$ is finite. Suppose that the subspaces $P\mathcal{H}$ and $Q\mathcal{H}$ are defined by their "bases" E_0 and F_0, so that

$$E_0\ E_0{}^* = P \qquad \text{and} \qquad F_0\ F_0{}^* = Q.$$

So, in terms of P and Q we have an expression for the square of the direct rotation, given by

$$U^2 = (2Q-I)\ (2P-I). \tag{5.2.1}$$

If we are in the acute case, then we find that the direct rotation is unique, and we have just to find the principal square root, i.e. unitary square root whose spectrum is in the right half plane. Since U^2 is represented by an $n \times n$ matrix, one can write

$$U^2 = A + iB$$

where $i^2 = -1$, and A and B are real matrices. U^2 being unitary, implies that

$$(5.2.2) \qquad AA' + BB' = A'A + B'B = I, \quad AB' - BA' = A'B - B'A = 0.$$

Hence, if we let W be the real symmetric matrix of order 2n, defined as

$$W = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

then relation (5.2.2) gives

$$WW' = W'W = I,$$

so that W is an orthogonal matrix. Furthermore, from $W = KSK^{-1}$

where $S = \begin{pmatrix} A+iB & 0 \\ B & A-iB \end{pmatrix}$

and $K = \begin{pmatrix} I & -iI \\ 0 & I \end{pmatrix}$

there follows that

$$\det W = \det S = \det(A + iB) \, . \ \det\,(\overline{A + iB})$$
$$= |\det\,(A + iB)|^2 = |\det U^2|^2 = 1.$$

We now refer to [22] for a detailed discussion of how to compute the principal square root of an orthogonal matrix, with determinant equal to +1. It turns out that the principal square root of W is a matrix R of the form

$$R = \begin{pmatrix} L & -M \\ M & L \end{pmatrix}$$

with $L - iM$ a unitary matrix, and $(L + iM)^2 = A + iB$.

Thus, in order to find U, one has to work with a matrix of double dimensions.

Sometimes, it is of interest to find the restriction of U on $P\mathcal{N}$ and the above procedure will be computationally inefficient, especially when the dimension of $P\mathcal{N}$ is relatively small compared to that of $\mathcal{N}$. In the following, we will provide an economical expression for U.

<u>Lemma 5.2.1</u>

Let E_0 and F_0 be bases for $P\mathcal{N}$ and $Q\mathcal{N}$ respectively. Then there exists an isometry from $K(E_0)$ onto $R(F_0)$, so that it gives a basis for $R(F_0)$ closest to E_0.

Proof

Since

$$E_0{}^*E_0 = I,\ E_0E_0{}^* = P,\ F_0{}^*F_0 = I \text{ and } F_0F_0{}^* = Q,$$

and since we assume the equation (5.1.1) to be satisfied, then we have

$$UE_0\ E_0{}^* = F_0F_0{}^*U$$

where U is the direct rotation mapping $P\mathcal{H}$ onto $Q\mathcal{H}$.
Let $W_0 : K(E_0) \to K(F_0)$ be defined by

$$W_0 = F_0{}^*\ U\ E_0$$

It is easy to check that $W_0{}^*\ W_0 = W_0\ W_0{}^* = I$. Let $F: K(E_0) \to$ be defined by

(5.2.3) $$F = F_0\ W_0$$

Then $R(F) = R(F_0)$ and $F^*\ F = I$, so that F is an isometry mapping $K(E_0)$ onto $R(E_1)$. But since

$$F = F_0\ W_0 = F_0\ F_0{}^*\ U\ E_0 = U\ E_0\ E_0{}^*\ E_0 = U\ E_0,$$

F will be a basis for $R(E_1)$ closest to E_0 as was shown in theorem (1.5.2). This proves the lemma.

We need now to find an expression for W_0 in terms of F_0 and E_0, so that F will be expressed also in terms of E_0 and F_0.

Lemma 5.2.2

$$F = F_0(F_0^* E_0) [(F_0^* E_0)^* (F_0^* E_0)]^{-1/2}$$

Proof

Since $W_0 = F_0^* U E_0$, using (5.1.4) we have

$$W_0 = F_0^* T E_0 = F_0^* (P+Q-I) [(I-P-Q)^2]^{-1/2} E_0$$

$$= F_0^* (E_0 E_0^* + F_0 F_0^* - I) [(I-P-Q)^2]^{-1/2} E_0$$

$$= F_0^* E_0 E_0^* [(I-P-Q)^2]^{-1/2} E_0.$$

Let $C = \cos^2 \Theta = (I-P-Q)^2$, then

$$CE_0 = (I-P-Q + PQ + QP) E_0$$

$$= (I - E_0E_0^* - F_0F_0^* + E_0E_0^* F_0F_0^* + F_0F_0^* E_0E_0^*) E_0$$

$$= -F_0F_0^* E_0 + E_0E_0^* F_0F_0^* E_0 + F_0F_0^* E_0$$

$$= E_0(E_0^* F_0F_0^* E_0) = E_0(F_0^* E_0)^* (F_0^* E_0)$$

Let $L = (F_0^* E_0)^* (F_0^* E_0) \geq 0$ then

$$CE_0 = E_0L \quad \text{and}$$
$$C^2E_0 = E_0L^2$$

and in general,

$C^n E_0 = E_0 L^n$, for any positive integer n. Thus, for all polynomials f(C), we have

$$f(C) E_0 = E_0 f(L)$$

Thus, this is true for all continuous functions on [0,1], so it is true for the inverse square root, provided that $0 \notin \sigma(C)$,

i.e. $$C^{-1/2} E_0 = E_0 L^{-1/2}, \text{ and}$$

$$W_0 = (F_0^* E_0)[(F_0^* E_0)^* (F_0^* E_0)]^{-1/2}$$

and

$$F = F_0(F_0^* E_0) [(F_0^* E_0)^* (F_0^* E_0)]^{-1/2}.$$

<u>Lemma 5.2.3</u>

Let

$$G = F + E_0,$$

then $$D = G(G^* G)^{-1/2}$$

is a basis for $Z\mathcal{H}$.

<u>Proof</u>

We have

$$G = F + E_0 = UE_0 + E_0 = TE_0 + E_0 = 2ZE_0;$$

by the remark on theorem 5.1.1, we have

$$ZP\mathcal{H} = Z\mathcal{H}$$

and it follows that

$$G\mathcal{H} = Z\mathcal{H}.$$

Now $D^*D = I$, $R(D) = Z\mathcal{N}$, so it follows that D is a basis for $Z\mathcal{N}$, and

$$DD^* = ZE_0 \, [(ZE_0)^* \, (ZE_0)]^{-1} \, (ZE_0)^* = Z.$$

Remark

The construction of D was inspired by the elementary fact that the diagonal of a rhombus bisects the angle from which it emanates.

Note that

$$T = 2Z - I = 2DD^* - I$$

where $$D = G(G^* \, G)^{-1/2}$$

$$G = E_0+F = E_0+F_0 \, (F_0{}^*E_0) \, [(F_0{}^*E_0)^* \, (F_0{}^*E_0)]^{-1/2}$$

Thus $$U = (2DD^* - I) \, (2E_0E_0{}^* - I)$$

But $$(2E_0E_0{}^* - I)|_{P\mathcal{N}} = I.$$

Thus, we have an economical expression for U (when restricted to $P\mathcal{N}$) in terms of E_0 and F_0.

As an illustration, if $\dim \mathcal{N} = 50$ and $\dim P\mathcal{N} = 5$, then $F_0{}^*E_0$ will be a 5×5 matrix, and the computation of Ux, $x \in P\mathcal{N}$ will be notably shortened.

We should remark here, that in the previous expression of U, we did not demand that E_0 be represented as $\begin{pmatrix} I \\ 0 \end{pmatrix}$ in which case we would have the nice matrix representation for

$$U \simeq \begin{pmatrix} C_0 & -S_0{}^* \\ S_0 & C_1 \end{pmatrix}$$

But on the other hand, $P = E_0E_0{}^*$ instead of $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$.

A different algorithm to find the direct rotation U was done by A. Björck and G.H. Golub [5]. Their main tool was the singular value decomposition of a matrix.

<u>Theorem 5.2.4 [6, P.134]</u>

Let A be an m×n matrix with rank r. Then there exists m×m and n×n unitary matrices U and V and r×r diagonal matrix D with strictly positive elements called the singular values of A, such that

$$A = USV^*, \quad S = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \qquad D = diag\ (s_1, \ldots, s_r)$$

The columns u_i and v_i of U and V satisfy

$$Av_i = s_iu_i,$$

$$A^*u_i = s_iv_i,$$

so that $A^*Av_i = s_i{}^2v_i,$

$$AA^*u_i = s_i{}^2u_i.$$

They are called the singular vectors. This leads to the singular value decomposition of A (shortly SVD):

$$A = \sum_{i=1}^{r} s_i\ u_i\ v_i{}^*.$$

Since the direct rotation is expressed as

$$U = (E_0 \quad E_1) \begin{pmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix}$$

$$= E_0C_0E_0^* + E_1S_0E_1^* - E_0S_0^*E_1^* + E_1C_1E_1^*$$

If we consider the SVD of C_0, S_0 and C_1, then it turns out that assuming dim $P\mathcal{X} = k$, $2k \leq n$

$$C_0 = Y_{E_0} \, C \, Y_{E_0}^* , \qquad S_0 = Y_{E_1} \begin{pmatrix} S \\ 0 \end{pmatrix} Y_{E_0}^* ,$$

$$S_0^* = Y_{E_0} \, (S \quad 0) \, Y_{E_1}^* , \quad C_1 = Y_{E_1} \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} Y_{E_1}^* .$$

Then $U = E_0 \, Y_{E_0} \, C \, Y_{E_0}^* \, E_0^* + E_1 \, Y_{E_1} \begin{pmatrix} S \\ 0 \end{pmatrix} Y_{E_0}^* \, E_0^*$

$$+ E_0 \, Y_{E_0} \, (-S \quad 0) \, Y_{E_1}^* \, E_1^* + E_1 \, Y_{E_1} \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} Y_{E_1}^* \, E_1^*$$

where $C = \text{diag} \ (\cos \theta_k)$, $S = \text{diag} \ (\sin \theta_k)$,

$$U = V_{E_0} \, C \, V_{E_0}^* + V_{E_1} \begin{pmatrix} S \\ 0 \end{pmatrix} V_{E_0}^* + V_{E_0} \, (-S \quad 0) \, V_{E_1}^*$$

$$+ V_{E_1} \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} V_{E_1}^*$$

Let $V_{E_1} = (W_{E_1} \quad Z_{E_1})$, where W_{E_1} is an $n \times k$ matrix and $W_{E_1}^* \, V_{E_0} = 0$, which is possible.

$$U = (V_{E_0}\ W_{E_1}\ Z_{E_1}) \begin{pmatrix} C & (-S \quad 0) \\ \begin{pmatrix} S \\ 0 \end{pmatrix} & \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} \end{pmatrix} \begin{pmatrix} V_{E_0}^* \\ W_{E_1}^* \\ Z_{E_1}^* \end{pmatrix}$$

where $V_{E_0} = E_0\, Y_{E_0}$ are the principal vectors in $P\mathcal{N}$ [1], associated with the pair of subspaces $P\mathcal{N}$ and $Q\mathcal{N}$, and W_{E_1} are the principal vectors in $Q\mathcal{N}^{\perp}$ associated with the pair of subspaces $P^{\perp}$ and Q. So, U will be determined if the quantities C, S, V_{E_0}, W_{E_1} are known. In [17], an efficient algorithm for computing the SVD of a matrix is shown. Now, we are given (as before) bases E_0 and F_0 for $P\mathcal{N}$ and $Q\mathcal{N}$. In [5], these quantities were calculated as follows:

Let $L = E_0^*\, F_0$,

So the SVD of L will be

$E_0^*\, F_0 = Y_{E_0}\, C\, Y_{F_0}^*$.

Now $PF_0 = E_0\, E_0^*\, F_0 = E_0 L$,

So the SVD of PF_0 is $PF_0 = E_0\, Y_{E_0}\, C\, Y_{F_0}^* = V_{E_0}\, C\, Y_{F_0}^*$

where $C = \text{diag}\ (\cos\theta_k)$.

Also the SVD of $(I-P)F_0$ is $(I-P)F_0 = (I-E_0E_0^*)F_0 = W_{E_1}\, S\, Y_{F_0}^*$

where $S = \text{diag}\ (\sin\theta_k)$.

We can choose W_{E_1} such that $W_{E_1}^*\, V_{E_0} = 0$, so we have S, C, V_{E_0} and W_{E_1} by doing 2 SVD, so to find U we just complete $(V_{E_0}\ W_{E_1})$ to be a basis for $\mathcal{N}$ (this is always possible), say

$(V_{E_0}\ W_{E_1}\ Z_{E_1})$ so

$$U = (V_{E_0}\ W_{E_1}\ Z_{E_1}) \left(\begin{array}{cc|c} C & -S & 0 \\ S & C & 0 \\ \hline 0 & 0 & I \end{array}\right) \begin{pmatrix} V_{E_0}{}^* \\ W_{E_1}{}^* \\ Z_{E_1}{}^* \end{pmatrix}$$

One should also note that

$$U\big|_{P\mathcal{H}} = (V_{E_0}\ W_{E_1}) \begin{pmatrix} C \\ S \end{pmatrix} V_{E_0}{}^* \ .$$

Comparing this algorithm, with that given before, we find that this algorithm will not be computationally efficient, although it provides other information implicitly, such as the angle between the two subspaces and the principal vectors.

APPENDIX A

Polar Representation of a Bounded Operator

For $A \in \mathfrak{B}(\mathcal{H})$, where $\mathcal{H}$ is a separable Hilbert space, N(A) denotes the null space of A and R(A) is the range of A. It is known that

$\mathcal{H} = \overline{R(A)} \oplus N(A^*) = \overline{R(A^*)} \oplus N(A)$. The bar denotes the norm closure of the corresponding linear manifolds.

Definition A.1

An operator $U \in \mathfrak{B}(\mathcal{H})$ is said to be a partial isometry if it maps $\mathcal{H} \ominus N(U)$ isometrically onto R(U). So for a partially isometric operator U, the linear manifold R(U) is a subspace. Let P_1 be the orthogonal projector onto $\mathcal{H} \ominus N(U)$, then U being a partial isometry is equivalent to $\|U\phi\| = \|P_1\phi\|$ for all $\phi \in \mathcal{H}$. Consequently,

$$\|U\phi\|^2 = \|P_1\phi\|^2,$$

$$((U^*U-P_1)\phi,\phi) = 0 \qquad (\phi \in \mathcal{H}),$$

so $U^*U = P_1$.

Since $(I-P_1)\phi \in N(U)$ (all $\phi \in \mathcal{H}$) one has $U(I-P_1)\phi = 0$ $(\phi \in \mathcal{H})$ and hence $U = UP_1$.

The relation $U^*U = P_1$ implies $\|U^* U\phi\| = \|P_1\phi\| = \|U\phi\|$,

so $\|U^*\phi'\| = \|\phi'\|$ for all $\phi' \in R(U)$ and U^* is a partial isometry. As above, let P_2 be the projector onto the subspace $\mathcal{H} \ominus N(U^*)$, then $\|U^*\phi\| = \|P_2\phi\|$ for all $\phi \in \mathcal{H}$, which implies $U U^* = P_2$ and $U^* = U^* P_2$.

Next, let A be any operator from $\mathcal{B}(\mathcal{H})$. Then it is well known that there exists a unique nonnegative operator H such that $H^2 = A^*A$ $(H = (A^*A)^{1/2})$.

It follows that $\|Af\|^2 = (Af, Af) = (A^*Af, f) = (H^2f, f)$

$$= (Hf, Hf) = \|Hf\|^2$$

so that $\|Af\| = \|Hf\|$ for all $f \varepsilon \mathcal{H}$, which implies that there exists an isometry $U: \quad R(H) \overset{\text{onto}}{\to} R(A)$ such that $Af = UHf$. Extending U to all of $\overline{R(H)}$ by continuity, and setting $U\phi = 0$ for $\phi \varepsilon N(H)$, we obtain a partial isometry. The fact $H \geq 0$ implies $N(H) = N(H^2)$; also $\|Af\| = \|Hf\|$ (all $f \varepsilon \mathcal{H}$) implies $\overline{R(H)} = \overline{R(A^*)}$, so all these give $\overline{R(H)} = \overline{R(H^2)} = \overline{R(A^*A)} = \overline{R(A^*)}$. That is, U is a partial isometry which maps $\overline{R(A^*)}$ onto $\overline{R(A)}$.

Hence every operator $A \varepsilon \mathcal{B}(\mathcal{H})$ admits a representation in the form

(1) $\qquad A = UH$

where $H = (A^*A)^{1/2}$ and U is a partial isometry which maps $\overline{R(A^*)}$ isometrically onto $\overline{R(A)}$. (1) is called the polar representation of A.

From (1), it follows that

(i) $U^*A = H$, since $U^*A = U^*UH = P_1H = H$;

(ii) $H_1 = UHU^*$, $H = U^*H_1U$, where $H_1 = (AA^*)^{1/2}$.

$(H_1f,f) = (UHU^*f,f) = (HU^*f,U^*f) \geq 0 \quad f \varepsilon \mathcal{H}$. Since $H \geq 0$, this implies that H_1 is non-negative, and since

$H_1^{\,2} = UHU^* \; UHU^* = UHP_1HU^* = UH^2U^* = AA^*$, then $H_1 = (AA^*)^{1/2}$ by the uniqueness of the non-negative square root.

Now $U^*H_1U = P_1HP_1 = H$,

(iii) $A = H_1U$, $H_1 = AU^*$.

It follows from (ii) that $H_1 = UHU^* = AU^*$.

Also, $A = UH = UU^*H_1U = P_2H_1U = H_1U$;

this implies that $A = H_1U$ and $A^* = U^*H_1$.

Remarks

(1) If $A \varepsilon \mathcal{B}(\mathcal{H})$ and if A is invertible, then there exists a unique unitary operator U and a positive operator H such that

$$A = UH.$$

The partial isometry will be unitary since A is invertible.

(2) If $A \varepsilon \mathcal{B}(\mathcal{H})$ and A is normal, then A has a polar decomposition $A = UH$ where U is a unitary and H is a non-negative operator. The operators U and H commute with each other and with A.

(3) By the dimension (also called rank) of the operator A, we mean the number $r(A)$ $(\leq \infty)$ equal to the dimension of the subspace $\overline{R(A)}$.

It is clear that $r(A) = r(H) = r(H_1) = r(A^*)$.

APPENDIX B

Singular Values and Unitary Invariant Norms

§B.1 The Singular Values of a Completely Continuous Operator

Let A be a completely continuous operator. The eigenvalues of H where $H = (A^*A)^{1/2}$ are called the singular values of A. We shall enumerate the non-zero singular values of A in decreasing order taking account of their multiplicities, so that

$$s_j(A) = \lambda_j(H) \qquad (j=1,2,\ldots).$$

If $\text{rank}(H) < \infty$ then $s_j(A) = 0$ where $j = r(H) + 1$. Also, we have

(i) $s_1(A) = \lambda_1(H) = \|H\| = \|A\|$

(ii) $s_j(A) = |\lambda_j(A)|$ when A is self adjoint,

(iii) $s_j(cA) = |c|\, s_j(A)$ $(j=1,2,3,\ldots)$, c is a constant.

We encounter two important properties of the singular values of a completely continuous operator.

Lemma B.1.1

For a completely continuous operator A, we have

(i) $s_j(A) = s_j(A^*)$ $(j=1,2,\ldots)$.

(ii) For any bounded operator B,

$$s_j(BA) \le \|B\|\, s_j(A) \qquad (j=1,2,\ldots)$$

$$s_j(AB) \le \|B\|\, s_j(A) \qquad (j=1,2,\ldots)$$

Proof

Since it is well known that for a self adjoint completely continuous operator A, all the eigenvalues are real and the operator has a uniformly convergent representation

$$A = \sum_{j=1}^{\nu(A)} \lambda_j(A) \ (\ ,\phi_j) \ \phi_j$$

where ϕ_j $(j=1,...\nu(A))$ is an orthonormal system of eigenvectors of A, complete in R(A), such that

$$A\phi_j = \lambda_j(A)\phi_j \qquad j=1,2,...\nu(A),$$

and where $\nu(A)$ is the sum of the algebraic multiplicities of all the non-zero eigenvalues of the operator A. Note that $\nu(A)$ is related to r(A) by the inequality

$$\nu(A) \leq r(A).$$

Hence, H has the representation

$$H = \sum_{j=1}^{r(H)} s_j(A) \ \phi_j \ \phi_j^*.$$

Now let A = UH be the polar representation of A, so it follows that

$$A = UH = \sum_{j=1}^{r(A)} s_j(A) \ U\phi_j \ \phi_j^* \qquad \phi_j \varepsilon R(A).$$

Since U is a partial isometry mapping $\overline{R(H)}$ onto $\overline{R(A)}$ then $U\phi_j = \psi_j$ constitutes an orthonormal system complete in R(A). Consequently,

(1) $$A = \sum_{j=1}^{r(A)} s_j(A)\ \psi_j\ \phi_j^*,$$

Hence

(2) $$A^* = \sum_{j=1}^{r(A)} s_j(A)\ \phi_j\ \psi_j^*.$$

Next, we prove that A^*A has the same eigenvalues as AA^*.

It follows from (1) and (2), that

$$A^*A\ \phi_j = s_j^2\ (A)\ \phi_j \qquad j=1,2,\ldots,r(A)$$

$$AA^*\ \psi_j = s_j^2\ (A)\ \psi_j \qquad j=1,2,\ldots,r(A)$$

So, we obtain

$$s_j(A) = s_j(A^*) \qquad j=1,2,\ldots,r(A)$$

This proves (i). For (ii), we have

$$s_j^2\ (BA) = \lambda_j(A^*\ B^*\ BA).$$

But we have

$$(A^*B^*BAf,f) = \|BAf\|^2 \leq \|B\|^2\ (Af,Af) \qquad f\varepsilon\,\mathcal{H}$$

which implies $A^*B^*BA \leq \|B\|^2\ A^*A$. The last inequality implies that $\lambda_j(A^*B^*BA) \leq \lambda_j(\|B\|^2\ A^*A) = \|B\|^2\ \lambda_j(A^*A)$. Therefore

$$s_j^2\ (BA) \leq \|B\|^2\ \ s_j^2\ (A), \text{ that is, } s_j\ (BA) \leq \|B\|\ s_j(A).$$

Statement (iii) follows directly from (ii) since

$$s_j(AB) = s_j(B^*A^*) \leq \|B^*\|\ \ s_j(A^*) = \|B\|\ \ s_j(A)$$

Remarks

(1) The expansion (1) is called the Schmidt expansion of a completely continuous operator where $\{\phi_j\}$, $\{\psi_j\}$ are certain orthonormal systems.

(2) For a non-negative completely continuous operator A, we have the following minimax properties of eigenvalues ([31], §95):

Theorem B.1.2

Let $A(\not\equiv 0)$ be a non-negative completely continuous operator and let $\phi_j (j=1,2,\ldots)$ be an orthonormal system of its eigenvectors which is complete in $R(A)$, so that

$$A\phi_j = \lambda_j(A)\phi_j \qquad (j=1,2,\ldots),$$

where $\lambda_1(A) \geqq \lambda_2(A) \geqq \ldots\ldots$ Then its eigenvalues have the following minimax properties:

$$(3) \qquad \lambda_1(A) = \max_{\phi \in \mathcal{H}} \frac{(A\phi,\phi)}{(\phi,\phi)}$$

where the maximum in (3) is attained only for those eigenvectors of the operator A which correspond to $\lambda_1(A)$.

$$(4) \qquad \lambda_{j+1}(A) = \min_{\chi} \max_{\phi \in \chi^{\perp}} \frac{(A\phi,\phi)}{(\phi,\phi)} \qquad (j = 1,2,\ldots)$$

where the minimum is taken over all j-dimensional subspaces of the space $\mathcal{H}$, and the minimum in (4) is attained when χ coincides with the linear subspace of the eigenvector $\phi_1, \phi_2, \ldots, \phi_j$.

B.1.3 Equivalent Definition of the Singular Values of a Completely Continuous Operator

We shall denote by B_n $(n = 0,1,2,\ldots)$ the set of all finite dimensional operators of dimension less or equal to n.

Let A be a completely continuous operator, then for any $n = 0,1,2,\ldots$

(5) $$s_{n+1}(A) = \min_{k\varepsilon B_n} \|(A-K)\|.$$

To prove the equivalence, let K be an n-dimensional operator. Then the subspace $\mathcal{H}\ominus N(K)$ is n-dimensional (recall that $r(K^*) = r(K)$). Now it follows for (4) that

$$s_{n+1}(A) \leq \max_{\phi\varepsilon N(K)} \frac{\|A\phi\|}{\|\phi\|} .$$

Since for all $\phi\varepsilon N(K)$, we have $\|A\phi\| = \|(A-K)\phi\|$, then $\|A\phi\| \leq \|A-K\| \; \|\phi\|$,

so $$s_{n+1}(A) \leq \|A-K\| \qquad k\varepsilon B_n.$$

Let $K_n = \sum_{j=1}^{n} s_j(A)\, \psi_j\, \phi_j{}^*$ be the n-th partial sum of the Schmidt expansion of A. Clearly K_n has dimension n and

$$\| (A-K_n)f\|^2 = \| \sum_{j=n+1}^{r(A)} s_j(A)\ \psi_j\ \phi_j^*\ (f)\|^2$$

$$= \sum_{j=n+1}^{r(A)} s_j^2\ (A)\,|(f,\phi_j)|^2$$

$$\leq s_{n+1}^2 \sum_{j=n+1}^{r(A)} |(f,\phi_j)|^2$$

$$\leq s_{n+1}^2\ \|f\|^2,$$

So that $\|A-K_n\| \leq s_{n+1}$, hence $\|A-K_n\| = s_{n+1}$, concluding that

$$s_{n+1}(A) = \min_{K\varepsilon B_n} \|A-K\| \qquad n = 0,1,2,...$$

In fact, (5) shows that $s_{n+1}(A)$ is the distance from the operator A to the set B_n.

From this equivalent definition of the singular values of a completely continuous operator, we have the following inequalities. The proof can be found in [18].

1. If A is a completely continuous operator, let T be any r-dimensional operator. Then

$$s_{n+r}(A) \leq s_n(A+T) \leq s_{n-r}(A).$$

2. (K. Fan [14]) If A,B are completely continuous operators, then

$$s_{m+n-1}(A+B) \leq s_m(A) + s_n(B) \qquad (m,n = 1,2,...),$$

$$s_{m+n-1}(AB) \leq s_m(A)\ s_n(A).$$

3. For any linear completely continuous operators A,B,

$$|s_n(A) - s_n(B)| \leq \|A-B\| \quad (n = 1,2,\ldots).$$

<u>Lemma B.1.4</u> [A. Horn [23], K. Fan [14])

For any completely continuous operators A,B,

$$\prod_{j=1}^{n} s_j(AB) \leq \prod_{j=1}^{n} s_j(A) \prod_{j=1}^{n} s_j(B) \qquad (n = 1,2,\ldots),$$

$$\sum_{j=1}^{n} s_j(A+B) \leq \sum_{j=1}^{n} s_j(A) + \sum_{j=1}^{n} s_j(B) \quad (n = 1,2,\ldots).$$

§B.2 Symmetric Norms

A functional $\|X\|_s$ defined on some two-sided ideal σ of the ring $\beta(\mathcal{H})$ is called a symmetric norm if it has the following properties:

(1) $\|X\|_s > 0 \qquad (X\varepsilon\sigma,\ X\neq 0)$,

(2) $\|\lambda X\|_s = |\lambda|\ \|X\|_s$ $(x\varepsilon\sigma)$, where λ is any complex number,

(3) $\|X+Y\|_s \leq \|X\|_s + \|Y\|_s \qquad (X,Y\varepsilon\sigma)$,

(4) $\|AXB\|_s \leq \|A\|\ \|X\|_s\ \|B\|$ $(A,B\varepsilon\ \beta(\mathcal{H}),\ X\varepsilon\sigma)$,

(5) for any one-dimensional operator X, $\|X\|_s = \|X\| = s_1(X)$.

Clearly, the bound norm is symmetric on any σ. If in the definition of a symmetric norm,(4) is replaced by

(4') $\|UX\|_s = \|XU\|_s = \|X\|_s$, $(X\varepsilon\sigma)$ where U is an arbitrary unitary operator, then we have the definition of a unitary invariant norm. Note that every symmetric norm is a unitary invariant norm.

(Since for a symmetric norm $\|UXV\|_s \leq \|X\|_s$, $\|X\|_s$

$$= \|U^{-1}UXV\,V^{-1}\|_s \leq \|UXV\|_s, \text{ hence } \|UXV\|_s = \|X\|_s).$$

The reverse will hold only under certain assumptions.

B.2.1 Important Properties of Symmetric Norm

1. Let σ be some two-sided ideal of the ring $(\mathfrak{X})$ and let a symmetric norm $\|\cdot\|_s$ be defined on σ. Then for any operator $X\varepsilon\sigma$,

$$\|X\|_s = \|X^*\|_s = \|(X^*X)^{1/2}\|_s = \|(X\;X^*)^{1/2}\|_s.$$

Indeed, if $X = UH$ is the polar representation of X, then

$$\|X\|_s = \|H\|_s;$$

on the other hand $U^*X = H$,

$$\|H\|_s = \|U^*X\|_s \leq \|X\|_s.$$

Consequently $\|X\|_s = \|H\|_s$.

Now starting from the equalities $X^* = HU^*$ and $X^*U = H$, we obtain $\|X^*\|_s = \|H\|_s$.

2. Let σ be some two-sided ideal of the ring $\beta\,(\mathfrak{X})$ and let a symmetric norm be defined on σ. Then for any operator $X\varepsilon\sigma$ and a completely continuous operator Y such that

$$s_j(Y) \leq c\, s_j(X) \qquad j = 1,2,\ldots,$$

where c is a positive constant, it follows that $Y \varepsilon \sigma$ and $\|Y\|_s \leq c\|X\|_s$.

Proof

If $H_x = (X^*X)^{1/2}$ and $H_y = (Y^*Y)^{1/2}$, then by the assumption $s_j(Y) \leq c\, s_j(X)$ one can find a unitary operator V and a non-negative operator $A \varepsilon \mathcal{B}(\mathcal{H})$ with $\|A\| \leq 1$ (A can be that operator with eigenvalues equal to

$$\frac{1}{c}\ \frac{s_j(Y)}{s_j(X)} \text{ or } 0 \text{ if } s_j(X) = 0),$$

So that $H_y = cAV\, H_x\, V^{-1}$

where V maps some orthonormal basis of eigenvectors of H_x into an appropriate orthonormal basis of eigenvectors of H_y.

It follows from $H_y = cAV\, H_x\, V^{-1}$ that $H_y \varepsilon \sigma$ and $\|H_y\|_s \leq c\, \|H_x\|_s$. Now it follows that $Y \varepsilon \sigma$ and $Y = U_y H_y$ (the polar representation) gives $\|Y\|_s \leq c\, \|X\|_s$.

3. For any symmetric norm $\|X\|_s$ defined on some two sided ideal σ we have $s_1(X) \leqq \|X\|_s$, and if dim $X < \infty$, then also $\|X\|_s \leqq \sum_j s_j(X)$.

Proof

In fact, let $Y = s_1(X)\phi\phi^*$, where ϕ is an arbitrary unit vector of $\mathcal{H}$. Then it follows that the property (2) is satisfied with c = 1, hence

$$s_1(X) = \|X\| = \|Y\| = \|Y\|_s \leq \|X\|_s;$$

on the other hand, if dim X < ∞, then we have

$$X = \sum_j s_j(X)\ \phi_j\ \psi_j^*$$

$$\|X\|_s = \|\sum_j s_j(X)\ \phi_j\ \psi_j^*\|_s.$$

Hence it follows from property (3) and (5) of a symmetric norm that

$$\|X\|_s \leq \sum_j s_j(X).$$

Remark

It follows from property (2) that the symmetric norm $\|X\|_s$ depends only on the singular values of X, that is, if the singular values of X_1, X_2 coincide, then $\|X_1\|_s$, $\|X_2\|_s$ also coincide.

So, for every symmetric norm we have

$$\|X\|_s = \Phi(s_1(X),\ s_2(X)\ldots)$$

where $\Phi\ (\xi_1, \xi_2, \ldots.)$ is a function of the non-negative variables ξ_i.

§B.3 Symmetric Norming Functions

The case when σ coincides with the ideal R of finite dimensional operators, the domain of the function Φ mentioned before consists of all non-increasing sequences $\{\xi_i\}$ of non-negative numbers of which only finitely many are different from zero.

Let c_0 be the space of all sequences $\xi = \{\xi_i\}_{i=1}^{\infty}$ of real numbers which tend to zero. We denote by $\hat{c}$ the linear manifold of c_0 consisting of all sequences with a finite number of non-zero terms.

Definition B.3.1

A real function $\Phi(\xi) = \Phi(\xi_1, \xi_2, ...)$ defined on $\hat{c}$ is called a norming (gauge) function if it has the following properties:

(i) $\Phi(\xi) > 0 \quad (\xi \varepsilon \hat{c},\ \xi \neq 0)$,

(ii) for any real α, $\Phi(\alpha\xi) = |\alpha|\ \Phi(\xi) \quad (\xi \varepsilon \hat{c})$,

(iii) $\Phi(\xi + \eta) \leq \Phi(\xi) + \Phi(\eta) \quad (\xi, \eta \varepsilon \hat{c})$,

(iv) $\Phi(1,0,0,...) = 1$.

A norming function $\Phi(\xi)$ is said to be symmetric if it has the property

(v) $\Phi(\xi_1, \xi_2, ..., \xi_n, 0, 0, ..) = \Phi(|\xi_{j_1}|, |\xi_{j_2}|, ..., |\xi_{j_n}|, 0, ...)$

where $\xi = \{\xi_i\}$ is any vector from $\hat{c}$ and $j_1, j_2, ... j_n$ is any permutation of the integers $1, 2, ..., n$.

We bring here various properties of symmetric norming functions.

1. Let $\xi = \{\xi_j\}\ \hat{c}$, let $0 \leq p_j \leq 1$. Then $\phi(p_1\ {}_1, p_2\ {}_2, ..) \leq \Phi(\xi)$.

Without loss of generality, we may assume $\xi_j \geq 0$ $j = 1,2,\ldots$ It is clear by induction that it is sufficient to prove the above conclusion when $p_j \neq 1$ occur only for one j, i.e.

$$\Phi(\xi_1, \xi_2, \ldots, p_i\xi_i, \xi_{i+1}) = \Phi(\xi).$$

For $0 \leq p \leq 1$, the conclusion follows from direct calculation:

$$\begin{aligned}\Phi(\xi_1,\xi_2,\ldots,p\xi_i,\ldots) &= \Phi(\tfrac{1+p}{2}\xi_1+\tfrac{1-p}{2}\xi_2,\ldots,\tfrac{1+p}{2}\xi_i+\tfrac{(1-p)}{2}(-\xi_i),\ldots)\\ &\leq \Phi(\tfrac{1+p}{2}\xi_1,\ldots,\tfrac{1+p}{2}\xi_i,\ldots)+\Phi(\tfrac{1-p}{2}\xi_1,\ldots+\tfrac{1-p}{2}(-\xi_i),\ldots)\\ &\leq \tfrac{1+p}{2}\,\Phi(\xi_1,\xi_2,\ldots,\xi_i,\ldots)+\tfrac{1-p}{2}\,\Phi(\xi_1,\xi_2,\ldots,-\xi_i,\ldots)\\ &\leq \tfrac{1+p}{2}\,\Phi(\xi) + \tfrac{1-p}{2}\,\phi(\xi)\\ &\leq \Phi(\xi)\end{aligned}$$

Lemma B.3.2 (K. Fan, L. Mirsky)

Suppose $\xi = \{\xi_i\}$ *and* $\eta = \{\eta_i\} \varepsilon \hat{c}$. *If*

$\xi_1 \geq \xi_2 \geq \ldots \geq 0,\ \eta_1 \geq \eta_2 \geq \ldots \geq 0$

then the set of inequalities

$$\sum_{j=1}^{k} \xi_j \leq \sum_{j=1}^{k} \eta_j \quad (k=1,2,\ldots)$$

is a sufficient and necessary condition for the relation

$$\Phi(\xi) \leq \Phi(\eta)$$

to hold for every symmetric norming function.

Proof

See [14], [29],

Theorem B.3.3

Let $\|\cdot\|_\sigma$ *be any unitary invariant norm on the ideal R of all finite dimensional operators. Then the equation*

$$\Phi(s(A)) = \|A\|_\sigma \quad (A\varepsilon R;\ s(A) = \{s_j(A)\})$$

defines a symmetric norming function $\phi(\xi)$*. Conversely, if* $\phi(\xi)$ *is any symmetric norming function, then the equality*

$$\|A\|_\phi = \Phi(s(A)) \ (A\varepsilon R)$$

defines an invariant norm on the ideal R.

Proof

See [18].

So for any two completely continuous operators A,B $\|A\| \leq \|B\|$ holds for any unitary invariant norm if and only if it holds for ν-norms defined by

$$\|A\|_\nu = s_1(A) + s_2(A) + .. + s_\nu(A) \qquad \nu = 1,2,...$$

We state the following lemma without proof.

Lemma B.3.4

Let P and Q be projectors. If $\|PKQ\| \leq \|PLQ\|$ *and* $\|(I-P)K(I-Q)\| \leq \|(I-P)L(I-Q)\|$ *for all unitary invariant norms, then* $\|PKQ+(I-P)K(I-Q)\| \leq \|PLQ+(I-P)L(I-Q)\|$ *for all unitary* *invariant norms. The converse will hold whenever PKQ has the same singular values as (I-P)K(I-Q) and PLQ has the same singular values as (I-P)L(I-Q)..*

BIBLIOGRAPHY

1. Afriat, S.N., Orthogonal and oblique projectors and the characteristics of pairs of vector spaces, Proc. Cambridge Philos. Soc. 53, pp. 800-816,(1957).

2. Akhiezer, N.I., and Glazman, I.M., Theory of linear operators in Hilbert Space (English translation), Vol. I and II, Frederick Ungar, New York,(1961).

3. Bellman, R., Introduction to Matrix Analysis, McGraw-Hill, New York, (1960).

4. Berberian, S.K., Approximate propervectors, Proc. Am. Math. Soc. 13, pp. 111-114, (1962).

5. Björck, A. and Golub, G.H., Numerical methods for computing angles between linear subspaces, Math. Comp. 27, pp. 579-594, (1973).

6. Björck, A. and Dahlquist, G., Numerical Methods, Prentice-Hall, (1974).

7. Cohen, C., An investigation of the geometry of subspaces for some multivariate statistical models, Thesis, Dept. of Indust. Eng., University of Illinois, Urbana, Ill., (1969).

8. Davis, C., Separation of two linear subspaces, Acta. Sci. Math. Szeged 19, pp. 172-187, (1958).

9. Davis, C., Rotation of eigenvectors by a perturbation, J. Math. Anal. Appl. 6, pp. 159-173,(1963).

10. Davis, C., The rotation of eigenvectors by a perturbation II, J. Math. Anal. Appl. 11, pp. 20-27, (1965).

11. Davis, C. and Kahan, W.M., The rotation of eigenvectors by a perturbation III, SIAM J. Num. Anal. 7, pp. 1-46, (1970).

12. Dunford, N. and Schwartz, J.T., Linear Operators, Interscience, New York, (1963).

13. Fan, K., On a theorem of Weyl concerning eigenvalues of Linear transformation, I, Proc. Nat. Acad. Sci. U.S.A. 35, pp. 652-655, (1949).

14. Fan, K. Maximum properties and inequalities for the completely continuous operators, Proc. Nat. Acad. Sci., U.S.A. 37, pp. 760-766, (1951).

15. Fix, G. and Heiberger, R., An algorithm for the ill conditioned generalized eigenvalue problem, Numer. Math. (To appear).

16. Flanders, D.A., Angles between flat subspaces of a real n-dimensional Euclidean space, Studies and essays presented to R. Courant on his 60th birthday, Jan. 8, 1948, ed. K.O. Friedrichs, O.E. Neugebauer, and J.J. Stocker, Interscience, New York, (1948).

17. Gloub, G.H. and Reinsch, C., Singular value decompositions and least squares solutions, Numer. Math. 14, pp. 403-420, (1970).

18. Gohberg, I.C. and Krein, M.G., Introduction to the Theory of linear Nonselfadjoint Operators, Nauka, Moscow (1965); English translation, Math. Monographs 18, Amer. Math. Soc., Providence, R.I., (1969).

19. Halmos, P.R., Introduction to Hilbert Space and the Theory of Spectral Multiplicity, Chelsea, New York, (1951).

20. Heinz, E., Beiträge zur Störungstheorie der Spektralzerlegung, Mathematische Annalen 123, pp. 415-438, (1951).

21. Jordan, C., Essai sur la géometrie à n dimensions, Bull. Soc. Math. France 3, pp. 103-174, (1875).

22. Hinds, E.W., Square root of an orthogonal matrix, Technical Report 37, Dept. of Computer Science, University of Toronto, Toronto (1971).

23. Horn, A., On the singular values of a product of completely continuous operator, Proc. Nat. Acad. Sci. U.S.A. 36, pp. 374-375, (1950).

24. Kato, T., Perturbation Theory for Linear Operators, Springer, Berlin and New York, (1966).

25. Knopp, K., Theory of Functions, Parts I and II Dover, New York, (1945).

26. Kovarik, Z.V., Similarity and interpolation between projectors, Mathematical Report 76, Dept. of Math., McMaster University, Hamilton, (1975).

27. Krein, M.G. and Krasnoselskii, M.A., Fundamental theorems on the extensions of Hermitian operators and some of their applications to the theory of orthogonal polynomials and the moment problem, Uspekhi mat. Nauk 2, (1947).

28. Lumer, G. and Rosenblum, M., Linear operator equations, Proc. Amer. Math. Soc. 10, pp. 32-41, (1959).

29. Mirsky, L., Symmetric gauge functions and unitarily invariant norms, Quart. J. Math., Oxford 11, pp. 50-59, (1960).

30. Neudecker, H., A note on Kronecker Matrix Products and Matrix equation system, SIAM J. Appl. Math. 17, pp. 603-606, (1969).

31. Riesz, F., and Sz.Nagy, B., Functional Analysis, (English translation), Frederick Ungar, New York, (1955).

32. Rosenblum, M., On the operator equation BX - XA = Q, Duke Math. J. 23, pp. 263-269, (1956).

33. Rudin, W., Functional Analysis, McGraw-Hill, New York, (1973).

34. Schatten, R., A theory of Cross-Spaces, Princeton, (1950).

35. Schoute, P.H., Mehrdimensionale Geometrie,I. Teil, Die linearen Räume, Leipzig, (1902).

36. Stewart, G.W., Error bounds for approximate invariant subspaces of closed linear operator, SIAM J. Numer. Anal. 8, pp. 796-808, (1971).

37. Swanson, C.A., An inequality for linear transformation with eigenvalues, Bull. Amer. Math. Soc. 67, pp. 607-608, (1961).

38. Sz.Nagy, B., On semi-groups of self adjoint transformations in Hilbert space, Proc. Nat. Acad. Sci. U.S.A. 24, pp. 559-560, (1938).

39. Taylor, A.E., Spectral theory of closed distributive operators, Acta. Math. 84, pp. 189-224, (1951).

40. Varah, J.M., Computing invariant subspaces of a general matrix when the eigensystem is poorly conditioned, Math. Comp. 24, pp. 137-149, (1970).

41. Wilkinson, J.H., The Algebraic Eigenvalue Problem, Clarendon, Oxford, (1965).