TRANSFORMATION AND PERTURBATION OF SUBSPACES OF A BANACH SPACE

By

© NAGWA A.E.H. SHERIF, B.Eng., B.Sc., M.Sc.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
September 1980

TRANSFORMATION AND PERTURBATION OF SUBSPACES OF A BANACH SPACE.

DOCTOR OF PHILOSOPHY (1980)
(Mathematics)

MCMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Transformation and Perturbation of Subspaces of a Banach Space.

AUTHOR: Nagwa A.E.H. Sherif, B.Eng (Cairo University)

B.Sc. (Hons.) (Cairo University)

M.Sc. (McMaster University)

SUPERVISOR: Professor Z.V. Kovarik

NUMBER OF PAGES: vii, 119

ABSTRACT

An n-frame on a Banach space X is $E=(E_1,\ldots,E_n)$ where the E_j's are bounded linear operators on X such that $E_j \neq 0$, $\sum_{j=1}^{n} E_j = I$ and $E_j E_k = \delta_{jk} E_k$ $(j,k=1,2,\ldots,n)$. This thesis is concerned with the study of pairs of such n-frames. It is shown that if two n-frames are close to each other then they are similar. A particular similarity, the direct rotation comes naturally in connection with the geodesic arc connecting the two frames when the set of n-frames is regarded as a Banach manifold. For a pair of 2-frames, the direct rotation is characterized. Another similarity, the balanced transformation which realizes the equivalence of the two frames is locally characterized and its closeness to the direct rotation is investigated. These results are used to obtain an error bound on invariant subspaces under perturbation. Our study, which is based on a functional calculus approach, involves techniques and results from operator theory, perturbation theory, and differential geometry. Some of the results are relevant to numerical spectral analysis.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my supervisor, Professor Z.V. Kovarik, for his help and patient guidance without which the completion of this thesis would have been impossible.

I also extend my appreciation to McMaster University for the generous financial support offered throughout my course of study there.

I would like also to acknowledge Josie Maljar for her patience and skill in the typing of this manuscript.

Finally, my deepest thanks go to my husband for his constant encouragement and moral support.

To my parents

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

INTRODUCTION

Pairs of linear subspaces of a real n-dimensional inner product space of equal dimension have been studied since 1875 [20]. Since then, it has been known that such a pair of subspaces has a number of angles equal to their dimension as a set of unitary invariants. A treatment of the subject in somewhat more modern style is given in [14]. The subject was developed by S.N. Afriat [1] and others. The extension to the case of Hilbert space was completely analyzed by C. Davis [6]. He showed in this work that if a pair of subspaces of a Hilbert space, or equivalently a pair of orthoprojectors, are close to each other in some sense, then there is a unitary operator U which maps one of the subspaces onto the other, while being as close as possible to the identity. This unitary operator is called the direct rotation.

In 1970 C. Davis and W. Kahan [9] unified the study of a pair of subspaces of a Hilbert space by introducing the angle operator Θ between the two subspaces. The direct rotation mentioned above is related to the angle operator Θ by $U=\exp(J\Theta)$ where J is a partial isometry and plays the role of the imaginary unit.

For a pair of oblique projectors on a Hilbert space, or in general for a pair of projectors on a Banach space X a similar study was initiated by T. Kato [22, I, §4.6]. A similarity R between the two projectors with an expression similar to that of the direct rotation, namely $R=\exp K$ was developed by Z. Kovarik [25]. Here K resembles the angle operator Θ, and is called the oriented angle, and R is called the direct rotation between the two projectors. When the set of all projectors on X is equipped with a differential structure compatible with its embedding in $B(X)$, then the direct rotation R turns out to be of particular significance. This is because it gives rise to an interpolating path between the two projectors, within the set of all projectors which is in fact a geodesic arc between them [26].

The study of a pair of projectors is of interest to operator theory as well as to perturbation theory. From the perturbation theory point of view, the study of a pair of projectors helps to find by how much the invariant subspaces of an operator will change when the operator is slightly perturbed. Results of this sort can be found in [7], [8], [9] and [38]. There are also results in same spirit, which differ in nature from those mentioned above. The main concern of these results is an estimate of a single eigenvector (See for example [13], [21], [40]). The previous results were mainly concerned with selfadjoint operator, and selfadjoint perturbation. The same problem was investigated from a different point of view for a closed operator on a Hilbert space (possibly non selfadjoint), as in [37].

The problem of perturbation of invariant subspaces in a different notation means that you have a representation of an operator with respect to a certain decomposition of the underlying space which is close enough to be block diagonal, and we want to perform block diagonalization. A result of this type for partitioned matrices was given in [34].

In this thesis the study mainly goes through two broad lines. The first is concerned with the study of a pair of n-frames, where by a frame we mean a decomposition of the identity into n commuting projectors on a Banach space. The case of 2-frame is treated separately where we can get more global results. Naturally some of our results about a pair of projectors on a Banach space will generalize those results for a pair of orthoprojectors on a Hilbert space. Along the second line, we investigate the application of this study of a pair of n-frames in perturbation theory. Namely, in analogy with the previous results mentioned about perturbations of invariant subspaces where the framework was a Hilbert space and considering only orthogonal 2-frames, we study the same problem in a Banach space setting with n-frames; $n \geq 2$. Our results allow us to give an error bound on the invariant subspaces under perturbations.

The results we obtain have applications relevant to numerical analysis. The application of these results is natural when we are looking for the eigenvalues and eigenvectors of an operator A which is close to a sufficiently familiar operator A_0. However, in many cases, we are given only the operator A, and we try to determine its eigenvalues and eigenvectors by some approximate method, then for the investigation of the results we obtained, we artificially introduce an operator A_0, which is close to A and about which we have some information. This method is called artificial perturbation (See [12, §62] and [16]).

Chapter I is mainly devoted to known results which will be needed later. We also introduce the notation of n-frame on a Banach space, and the block matrix representation of an operator with respect to a particular frame.

In chapter II we discuss the direct rotation between a pair of projectors on a Banach space. We give a characterization of the direct rotation, which resolves a problem left open in [25]. Connectibility with geodesic arcs between two nonsymmetric involutions is investigated. We describe these geodesic arcs as the solutions of a system of differential equations which is arranged to reflect the algebraic structure of the underlying manifold.

Chapter III deals with a pair of n-frames E and F. We discuss two particular similarities between E and F. One is the balanced transformation, known in the context of finite dimensional Hilbert spaces to be that unitary which realizes the equivalence while deviating minimally from the identity in the Hilbert-Schmidt norm. The other similarity is the direct rotation which comes naturally in connection with the geodesic arc connecting the two frames. We give a local characterization of the balanced transformation. The two similarities coincide in the case of 2-frames as well as in some cases of n-frames; $n>2$. We give examples to illustrate these possibilities. We conclude the chapter by showing that the two similarities are still close to each other; at a distance of order $\|E-F\|^3$.

Finally, in Chapter IV we discuss the case when the two frames arise from an operator A and the perturbed operator $A+H$. We answer the natural question about how far the two frames are in terms of the perturbation H and the separation of parts of the spectrum of the operator A. This result depends on how to measure the difference between the two frames, and how to measure the separation of two operators. These two measures are defined, and we justify their usage. We treat the case of a selfadjoint operator separately and obtain results which generalize those known for orthogonal 2-frames.

CHAPTER I

PRELIMINARIES

In this chapter we will assemble definitions, notations and some of the fundamental results which will be needed in the sequel. The chapter is divided into four sections. The first section is concerned with basic notations and definitions. Section 2 deals with differential calculus on Banach spaces, operational calculus as well as some basic results from spectral theory. We state the major theorems of these subjects in the form in which we will need them later. In Section 3 we define frames, discuss source spaces and block matrix representations. In the last section we discuss the transformation functions for smooth paths of frames.

1.1. Notations and Definitions.

Certain notational conventions will be observed throughout this work and these are given here for the sake of continuity. Throughout $\mathcal{X}$, $\mathcal{Y}$ will denote Banach spaces. We denote by $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ the space of bounded linear operators from a real or a complex Banach space to another. $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ becomes a Banach space under the bound norm

$$\|A\| = \sup\{\|Ax\| : \|x\| \le 1\} .$$

$\mathcal{B}(\mathcal{X})$ is $\mathcal{B}(\mathcal{X}, \mathcal{X})$ with the additional structure of a Banach algebra (under composition). The identity operator will always be denoted by I and, when dealing with several Banach spaces, its domain will be obvious from the context. For a positive integer n, $\mathcal{B}^n(X)$ will be the n-fold direct sum of $\mathcal{B}(X)$. It will be an algebra when multiplication is defined coordinatewise. We make $\mathcal{B}^n(X)$ into a Banach algebra by setting

$$\|B\| = \|(B_1, B_2, \ldots, B_n)\| = \max_{1 \le i \le n} \|B_i\| .$$

$G(X)$ denotes the multiplicative group of invertible elements of $\mathcal{B}(X)$. It is an open subset of $\mathcal{B}(X)$, and we will often use the fact that the map $A \longmapsto A^{-1}$ of $G(X)$ into itself is continuous and differentiable.

The *spectrum* of A in $\mathcal{B}(X)$ is the set

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ is not invertible in } X\} .$$

The complement of $\sigma(A)$ is the *resolvent set* of A, it consists of all $\lambda \in \mathbb{C}$ for which $(\lambda I - A)^{-1}$ exists. If $A \in \mathcal{B}(X)$ then $\sigma(A)$ is a non-empty compact subset of $\mathbb{C}$ contained in the closed ball $|\lambda| \le \|A\|$. In fact, if

$$r(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\} ,$$

($r(A)$ is called the *spectral radius* of A), then

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{n\to\infty} [P(A^n)]^{1/n} = \inf_{n\geq 1} [P(A^n)]^{1/n} ,$$

where P is an algebra norm equivalent to the original norm with $P(I) = 1$.

In case the underlying space is a Hilbert space $\mathcal{H}$ then $A\varepsilon\mathcal{B}(\mathcal{H})$ is said to be a Hilbert-Schmidt operator if the quantity $\|A\|_{HS}$, defined by the equation

$$\|A\|_{HS} = \left(\sum_{\alpha\varepsilon A} \|Ax_\alpha\|^2\right)^{1/2} ,$$

is finite. Here $\{x_\alpha, \alpha\varepsilon A\}$ is a complete orthonormal set. $\|A\|_{HS}$ will be called the Hilbert-Schmidt norm. Indeed, the Hilbert-Schmidt norm is independent of the orthonormal basis used in its definition. Also the Hilbert-Schmidt norm is unitarily invariant. In fact, the set $HS(\mathcal{H})$ of all Hilbert-Schmidt operators is a two-sided ideal in the Banach algebra of all bounded linear operators in $\mathcal{H}$. Moreover, if $T\varepsilon HS(\mathcal{H})$ and $B\varepsilon\mathcal{B}(\mathcal{H})$ then

$$\|TB\|_{HS} \leq \|T\|_{HS} \|B\| \quad \text{and} \quad \|BT\|_{HS} \leq \|B\| \|T\|_{HS} .$$

$HS(\mathcal{H})$ with the Hilbert-Schmidt norm is a Banach algebra without identity. In addition $HS(\mathcal{H})$ is a Hilbert space with inner product defined by

$$\begin{aligned}(S,T) &= \sum_\alpha (Sx_\alpha, Tx_\alpha) \\ &= \sum_{\alpha,\beta} (Sx_\alpha, x_\beta)(x_\beta, Tx_\alpha)\end{aligned}$$

where $\{x_\alpha, \alpha \varepsilon A\}$ is a complete orthonormal system. We remark here that every $A \varepsilon HS(H)$ is compact and in fact the Hilbert-Schmidt norm can be expressed equivalently by

$$||A||_{HS} = \left(\sum_k \alpha_k^2\right)^{1/2}$$

where α_k are the repeated singular values of A, that is the eigenvalues of $(A^*A)^{1/2}$, which means that

$$||A||_{HS} = (\text{trace } A^*A)^{1/2} .$$

The n-fold direct sum of $HS(H)$ is $HS^n(H)$ with inner product defined by

$$(A,B)_n = \frac{1}{2} \sum_{j=1}^{n} \text{trace } (B_j^* A_j) \quad \text{for} \quad n \geq 2 .$$

For details on Hilbert-Schmidt operators see [11, ChapXI].

1.2. Differential Calculus on Banach Spaces.

Suppose X, Y are Banach spaces, U is an open subset of X, f maps U into Y. If there exists $T \varepsilon B(X,Y)$ such that

$$\lim_{h \to 0} \frac{||f(x+h) - f(x) - Th||}{||h||} = 0 ,$$

then f is said to be differentiable at x. In view of the uniqueness of the map T, we call it the Fréchet derivative of f at x and denote it by $f'(x)$ or $Df(x)$. The concept of partial derivative will be often used. For that, let f be a continuous mapping of an open subset U of $X_1 \times X_2$ into Y. For each point (a_1, a_2) of U we say that f is differentiable with respect to the first variable if the partial mapping

$$x_1 \longmapsto f(x_1, a_2)$$

is differentiable at a_1. The derivative of that mapping which is an element of $B(X_1, Y)$ is called the partial derivative with respect to the first variable and written $f_1'(a_1, a_2)$ or $D_1 f(a_1, a_2)$. For further discussion we refer to [10] and [28]. Next we give one consequence of the contraction mapping theorem, which is of local nature, namely the implicit function theorem.

Theorem 1.1. (The Implicit Function Theorem). *Let X, Y, Z be three Banach spaces and f a continuously differentiable mapping of an open subset A of $X \times Y$ into Z. Let (x_0, y_0) be a point of A such that $f(x_0, y_0) = 0$ and that the partial derivative $D_2 f(x_0, y_0)$ is a linear homeomorphism of Y onto Z. Then there is an open neighborhood U_0 of x_0 in X such that, for every open connected neighborhood U of x_0 contained in U_0, there is a unique continuous mapping u of U into Y such that $u(x_0) = y_0$, $(x, u(x)) \in A$ and $f(x, u(x)) = 0$ for any $x \in U$. Furthermore u is continuously differentiable in U, and its derivative is given by*

$$u'(x) = -(D_2 f(x,u(x)))^{-1} \circ (D_1 f(x,u(x))) .$$

Proof: See [10, 10.2]. ///

One of the most useful theorems in analysis is the mean value theorem. We note that the formulation of that theorem, in the case of Banach spaces, differs from the classical mean value theorem (for real valued functions).

Theorem 1.2. (Mean Value Theorem). *Let* X, Y *be Banach spaces,* f *a continuous mapping into* Y *of a neighborhood of a segment* S *joining two points* x_0, x_0+h *of* X. *If* f *is differentiable at every point of* S, *then*

$$\|f(x_0+h) - f(x_0)\| \leq \|h\| \sup_{0\leq\xi\leq 1} \|f'(x_0+\xi h)\| .$$

Proof: See [10, 8.5]. ///

It was realized recently that a starting point of the spectral theory is the operational calculus (See [15], [31]). We start by defining what we mean by a general operational calculus as given in [24].

Definition 1.3. *A general operational calculus is an ordered pair* (C,F) *which consists of*

(i) a topological algebra C *of complex-valued functions on a subset* Δ *of the complex plane, with ordinary pointwise operations, which contains the restrictions to* Δ *of polynomials, and of*

(ii) a continuous representation $F(.)$ *of* C *in* $B(X)$ *such that* $F(1)=I$.

Definition 1.4. *An operator* $A \in B(X)$ *is of class* C *if there exists a general operational calculus* $(C, F_A(.))$ *such that* $F_A(z)=A$, *any* $F_A(.)$ *with this property will be called an* C *-operational calculus for* A.

If we take $C=H(\Delta)$, the algebra of all complex valued functions which are holomorphic in the open set Δ, with the topology of uniform convergence on every compact subset of Δ then an operator A is of class $H(\Delta)$ if and only if its spectrum $\sigma(A)$ is contained in Δ. In this case the representation $F_A(.)$ is given uniquely by

$$F_A(f) = \frac{1}{2\pi i}\int_{\Gamma} f(z)\,(zI-A)^{-1}\,dz\ , \quad f \in H(\Delta)$$

where Γ is an oriented envelope of $\sigma(A)$ with respect to f. For the proof of this fact, see [19], Theorem 5.2.5. The above operational calculus is always called the *analytic operational calculus*. Throughout we will use analytic operational calculus unless otherwise stated, and we write $f(A)$ for $F_A(f)$. Now we determine how the spectrum of $f(A)$ is related to the spectrum of A. The following result is essentially due to I. Gelfand [17].

Theorem 1.6. (The Spectral Mapping Theorem). *If* $f(\lambda) \in H(\Delta)$ *and* $A \in B(X)$ *such that* $\sigma(A) \subset \Delta$ *then* $\sigma(f(A)) = f(\sigma(A))$.

Proof: See [19] Theorem 5.3.1. ///

We note here that most of the above concepts and results are valid, in general, for any element in a Banach algebra.

1.3. Frames and Block Matrix Representations.

By a *frame* on a Banach space X we mean $E\varepsilon B^n(X)$, where $E=(E_1,E_2,\ldots,E_n)$ and the E_j's satisfy

$$(1.1) \quad E_j\neq 0 \ , \ \sum_{j=1}^{n} E_j=I \ \text{ and } \ E_jE_k=\delta_{jk}E_k \quad 1\leq j,\ k\leq n \ .$$

that is, $E_1,E_2,\ldots,E_n$ is a finite resolution of the identity. We denote the set of all n-frames on X by $E^n(X)$. Given a frame E, it gives rise to a decomposition of X into the direct sum

$$X = X_1\oplus X_2\oplus\ldots\oplus X_n \ , \quad \text{where} \quad X_j=E_jX \ , \ 1\leq j\leq n \ .$$

Also, given a frame E in $B(X)$, any operator $A\varepsilon B(X)$ will have a partition with respect to this frame in the following way:

In the above decomposition of X into the direct sum let $i_j : X_j \longrightarrow X$ be the inclusion map and let $s_j^! : X \longrightarrow X_j$ be defined by $s_j^!x = E_jx \ \varepsilon \ X_j$ so that $i_j \ \varepsilon \ B(X_j, X)$ and $s_j^!\varepsilon B(X,X_j)$ and

$$(1.2) \quad s_j^!i_j=I\varepsilon B(X_j),\ s_j^!i_k=0\varepsilon B(X_k,\ X_j) \quad j\neq k \ \text{ and } \ i_js_j^!=E_j \ .$$

Define

$$(1.3) \qquad A_{jk} = s_j' A i_k \in \mathcal{B}(X_k, X_j) \quad 1 \leq j,\ k \leq n ,$$

then the *block matrix* of A with respect to E is

$$(1.4) \qquad [A]_E = \begin{bmatrix} A_{11} & \cdots\cdots\cdots & A_{1n} \\ A_{21} & \cdots\cdots\cdots & A_{2n} \\ \vdots & & \vdots \\ A_{n1} & \cdots\cdots\cdots & A_{nn} \end{bmatrix}$$

Let $\mathcal{BM}_E(X)$ denote the space of all matrices of the form (1.4) with entries $A_{jk} \in \mathcal{B}(X_k, X_j)$. There is one to one correspondence between $\mathcal{B}(X)$ and $\mathcal{BM}_E(X)$. This is because every $A \in \mathcal{B}(X)$ defines $[A]_E \in \mathcal{BM}_E(X)$ as shown above, while if we have n^2 operators $A_{jk} \in \mathcal{B}(X_k, X_j)$, they will define an operator $A \in \mathcal{B}(X)$ as follows

$$(1.5) \qquad A = \sum_{j,k=1}^{n} i_j A_{jk} s_k' \in \mathcal{B}(X) .$$

Equation (1.4) can be related to equation (1.5) as follows

$$A = [i_1, i_2, \ldots, i_n] \qquad \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} s_1' \\ \vdots \\ s_n' \end{bmatrix}$$

while

$$[A]_E = \begin{bmatrix} s_1' \\ \cdot \\ \cdot \\ \cdot \\ s_n' \end{bmatrix} A \, [i_1 \ldots i_n]$$

Note that BM_E with the usual operations of addition and multiplication (as partitioned matrices), becomes an algebra and equations (1.4) and (1.5) define an algebra isomorphism between $B(X)$ and $BM_E(X)$. Further if we norm $BM_E(X)$ by

$$(1.6) \qquad \|[A]_E\| = \max_{1\leq j\leq n} \sum_k \|A_{jk}\| ,$$

then the above isomorphism becomes an algebra homeomorphism. The idea of using a block matrix representation with respect to a particular frame E helps to use the familiar matrix ideas. Moreover it helps to avoid ambiguity in some situations, for example when we discuss E_1AE_1 , it has null space $(I-E_1)X$, but what really matters of its spectrum is the spectrum of its restriction to E_1X which will be the same as the spectrum of A_{11} .

Remark. In the setting given above i_j, s_j' may be more general, namely, if X_j's are Banach spaces linearly homoeomorphic to E_jX $1\leq j\leq n$, then any injections $i_j \varepsilon B(X_j,X)$ and surjections $s_j' \varepsilon B(X,X_j)$ which satisfy (1.2) will do. In case one of the X_j's is finite dimensional, a convenient choice of $X_j, i_j, s_j^!$ was indicated in [27, chap3], using a basis of X_j .

In the following we use some notation from the theory of partitioned matrices. Consequently, an operator $A \varepsilon \mathcal{B}(X)$ is said to be *block diagonal* if $[A]_E$ is a block diagonal matrix, and similarly A is *block off diagonal* if $[A]_E$ is block off diagonal. We can express $[A]_E$ as $[A]_E = D+S$

where

$$D = \begin{bmatrix} A_{11} & & 0 \\ & \ddots & \\ 0 & & A_{nn} \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 0 & A_{12} & \cdots & A_{1n} \\ A_{21} & 0 & \cdots & A_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ A_{n1} & A_{n2} & \cdots & 0 \end{bmatrix} .$$

D is called the diagonal part of $[A]_E$ and S is the off diagonal part of $[A]_E$. In terms of the operator A and the E_j's , the block diagonal part of A will be $\sum_{j=1}^{n} E_j A E_j$ (since $[\sum_{j=1}^{n} E_j A E_j]_E = D$) , and the off diagonal part will be $A - \sum_{j=1} E_j A E_j$. Hence it follows that $A \varepsilon \mathcal{B}(X)$ is block off diagonal if and only if

$$\sum_{j=1}^{n} E_j A E_j = 0 .$$

and A will be block diagonal if and only if $A = \sum_j E_j A E_j$, as well as $AE_j = E_j A$ $1 \leq j \leq n$. In that case A will be invertible if and only if A_{jj} is invertible $1 \leq j \leq n$ and consequently $A_{jj}^{-1} = s_j' A^{-1} i_j$, and A^{-1} will also be block diagonal, namely

$$A^{-1} = \sum_{j=1}^{n} i_j A_{jj}^{-1} s_j'$$

1.4. Transformation Functions

Differentiable paths of frames were analyzed in [22, II,§4-5]. It was shown there, that if we are given a continuously differentiable path

$$t \longmapsto F(t) \quad t \in [0,1]$$

of frames, then we can construct an operator valued function $U(t)$ (called the transformation function or Kato's transformation for $F(t)$) with the following properties:

(1) the inverse $U(t)^{-1}$ exists and both $U(t)$ and $U(t)^{-1}$ are continuously differentiable,

(2) $F(t) = U(t)\, F(0)\, U(t)^{-1}$.

For that purpose we define

$$L(t) = \frac{1}{2} \sum_{k=1}^{n} (F_k'(t)\, F_k(t) - F_k(t)\, F_k'(t))$$

and consider the differential equation

$$U' = L(t)\ U$$

for the unknown U. Since this is a linear differential equation which has a unique solution when the initial value $U(0)$ is specified, let $U(t)$ be the solution of the above differential equation with $U(0)=I$. Similarly the differential equation

$$V' = -VL(t)$$

has a unique solution for a given initial value $V(0)$. Let $V(t)$ be the solution for $V(0)=I$. It can be shown that $U(t)$, $V(t)$ are inverses to each other as follows:

$$(UV)' = U'V + UV' = L(t)UV - UVL(t) .$$

But the above equation is a linear differential equation in UV. Clearly $Z(t)=I$ is a solution of the above equation and satisfies the initial condition. By the uniqueness, we have

$$U(t)V(t) = I \qquad 0 \le t \le 1 .$$

On the other hand

$$(VU)' = -VLU + VLU = 0 .$$

Hence

$$V(t)\ U(t) = I \qquad 0 \leq t \leq 1$$

so that

$$V(t) = U(t)^{-1} \qquad 0 \leq t \leq 1$$

and, and by direct calculation it can be shown that

$$F(t) = U(t)\ F(0)\ U^{-1}(t) \qquad 0 \leq t \leq 1$$

as claimed.

It is worth remarking that the transformation function is not unique. For sufficiently small t, another $U(t)$ can be defined [22,II,§4.6], but the above transformation function has the advantage that it can be defined for any continuously differentiable path of frames, though the actual procedure for finding U is to solve a differential equation.

As we will be considering Banach manifolds, some concepts and results from differential geometry will be needed. The standard textbook for that is [29]. Some of the concepts carry over from finite dimensional manifolds, these notations are adopted mainly from [3], [4] and [18].

CHAPTER II

DISCUSSION OF DIRECT ROTATIONS AND GEODESIC PATHS

2.1. Introduction

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 maps one of the subspaces onto the other. A particular unitary was singled out by C. Davis [6] and independently by T. Kato [22]. In fact this goes back to Sz. Nagy [39], though in a different expression. This particular unitary is called the direct rotation. It maps one of the subspaces onto the other while being as close as possible to the identity.

We give first the definition of the direct rotation as introduced in [9]. To do so, let E_0, E_1 be the ortho-projectors on the two subspaces, then the frame $[E_0, I-E_0]$ will be an orthogonal frame, and as it was mentioned before, every operator on H will have a block matrix representation with respect to $[E_0, I-E_0]$.

Definition 2.1. *A unitary solution* $[U]_{[E_0, I-E_0]} = \begin{bmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{bmatrix}$ *of* $UE_0 = E_1U$ *is called a direct rotation from* E_0H *to*

E_1H if it satisfies the following additional conditions

(i) $C_0 \geq 0$, $C_1 \geq 0$

(ii) $S_1 = S_0^*$.

This unitary may be expressed as $U = \exp J\Theta$ where Θ is the angle operator between E_0H and E_1H and J is normal with spectrum $\subseteq \{0, \pm i\}$. The operator Θ will give a measure of how far E_0H is from E_1H . For a pair of orthoprojectors E_0, E_1 , one can find in [9] the assumptions under which the direct rotation exists and is unique. If E_0, E_1 are given then the square of the direct rotation is easy to compute. Namely

$$U^2 = (2E_1 - I)(2E_0 - I) .$$

By a principal square root of $(2E_1 - I)(2E_0 - I)$ we mean a unitary U such that $U^* + U$ is positive semidefinite. Davis and Kahan gave a characterization of the direct rotation which will be given in the next theorem. The proof can be found in [9].

<u>Theorem 2.2</u>. *Any direct rotation of E_0H to E_1H is a principal square root of $(2E_1 - I)(2E_0 - I)$. Also any principal square root of $(2E_1 - I)(2E_0 - I)$ is a direct rotation provided it takes $E_0H \cap (I - E_1)H$ onto $(I - E_0)H \cap E_1H$.*

Observation 2.3. In the finite dimensional case, if $\det(I+(2E_1-I)(2E_0-I)) \neq 0$ then the direct rotation is unique. For if we suppose that U,V are two principal square roots of $(2E_1-I)(2E_0-I)$, then $U^2 = (2E_1-I)(2E_0-I) = V^2$.

But then

$$(U^*+U)U = I + (2E_1-I)(2E_0-I) = (V^*+V)V ,$$

so that U^*+U and V^*+V are positive definite. Also we have

$$(U+U^*)^2 = 2I + U^2 + (U^{-1})^2 = 2I + V^2 + (V^{-1})^2 = (V+V^*)^2 .$$

Since the positive definite square root of a positive definite operator is unique, we have $U+U^* = V+V^*$ and so $(U^*+U)U = (V^*+V)V$ implies $U=V$. Also in the infinite dimensional case, we get uniqueness of the direct rotation if $-1 \notin \sigma((2E_1-I)(2E_0-I))$.

Finally, the extremal properties of the direct rotation were investigated in great detail in [9]. It was also shown there that the direct rotation takes an element of E_0H to E_1H by the most economical route. An extensive bibliography on the history of the subject was given in the above mentioned work [9].

2.2. Similarity between projectors.

Let E_0, E_1 be two bounded linear projectors on a Banach space X. In [22] Kato showed that if $\|E_1 - E_0\| < 1$, then E_0 and E_1 are similar. In fact, he gave an algebraic expression for that similarity in terms of E_0, E_1. We have already presented the case of a pair of orthoprojectors in Section 1.

For a pair of projectors E_0, E_1 and under more relaxed conditions than those in [22], Kovarik [25] constructed a similarity between E_0 and E_1 which resembles the direct rotation presented previously; cf. Def 2.1. We shall show how to construct this similarity. For this purpose, as well as future use, we need the following trigonometry of projectors.

Lemma 2.4. *For any two projectors E_0, E_1 consider the closeness operator*

$$(2.1) \qquad C_1 = C_1(E_0, E_1) = (E_0 + E_1 - I)^2 ,$$

and the separation operator

$$(2.2) \qquad S_1 = S_1(E_0, E_1) = (E_1 - E_0)^2 .$$

Let $T_i = 2E_i - I$ $(i = 0, 1)$ be the associated involution with E_0 and E_1 and let $V_1 = T_1 T_0$. Then the following identities hold:

$$\begin{aligned}
&(i) \quad C_1 = C_1(E_0,E_1) = I - E_0 - E_1 + E_0E_1 + E_1E_0 \\
&\qquad\qquad = \tfrac{1}{4}(T_0+T_1)^2 = \tfrac{1}{4}(2I+V_1+V_1^{-1}) \\
&(ii) \quad C_1(E_0,E_1) = C_1(I-E_0,I-E_1) = C_1(E_1,E_0) \\
&(iii) \quad C_1(E_0,E_1) \text{ commutes with both } E_0 \text{ and } E_1 \\
(2.3)\quad &(iv) \quad C_1E_0 = E_0E_1E_0 \text{ and } C_1E_1 = E_1E_0E_1 \\
&(v) \quad S_1 + C_1 = I \\
&(vi) \quad S_1(E_0,I-E_1) = C_1(E_0,E_1) \\
&(vii) \quad S_1(E_0,E_1) = \tfrac{1}{4}(T_1-T_0)^2 = \tfrac{1}{4}(2I-V_1-V_1^{-1})
\end{aligned}$$

The proof is clear when done in the given order.

We can think of S_1 as an operator analogue of $\sin^2\theta$ and C_1 as $\cos^2\theta$ where θ is the non obtuse angle between the ranges of E_0 and E_1. If the E_i's are one-dimensional orthoprojectors on the Euclidean plane then indeed $C_1 = \cos^2\theta$ and $S_1 = \sin^2\theta$ where θ is the angle between them.

In the Hilbert space setting where $\{E_0,E_1\}$ is a pair of orthoprojectors, the separation operator has the property $0 \leq S_1 \leq I$. This property makes it possible to define the operator $\Theta = \arcsin\ (S_1^{\frac{1}{2}})$, the operator angle, as a function of S_1, it commutes with E_0 and E_1. However, in general Banach space this loses its meaning (since any nonempty compact set in the plane can be $\sigma(S_1)$; see Example 1[25]). In Theorem 2.6 below an oblique operator angle in a general Banach space will be constructed. This operator need not commute with E_0 and E_1 . The following theorem exhibits a similarity between E_0 and E_1 under certain assumptions on $\sigma(S_1)$. We shall outline the proof, and especially mention those parts which will be needed in the sequel. For the details of the proof we refer to [25]. We shall need the following.

Definition 2.5. *For a set $N \subset B(X)$, we denote by $A\{N\}$ the norm closed subalgebra of $B(X)$ generated by N and the identity and closed under inversion if defined.*

Theorem 2.6. *Let E_0, E_1 be projectors in $B(X)$; $E_0 \neq E_1$. If the number 1 lies in the unbounded component of the complement of $\sigma(S_1)$ then there exists an involution $T_{\frac{1}{2}}$ in $A\{E_0, E_1\}$ such that $E_1 = T_{\frac{1}{2}} E_0 T_{\frac{1}{2}}$ and there exists $K \varepsilon A\{T_1 T_0\}$ such that the projector valued path*

$$(2.4) \qquad t \longmapsto E_t = (\exp tK)\, E_0 \exp(-tK),\ 0 \le t \le 1$$

connects E_0 with E_1 . Moreover $T_1 T_0 = \exp(2K)$ and $T_t K = -K T_t$ where $T_t = 2E_t - I$.

The basic ideas of the proof are as follows: Let

Γ be a simple smooth curve in the complement, in the complex plane, of $\sigma(S_1)\cup\{0\}$ connecting 1 with ∞. Note that Γ exists by the assumptions on $\sigma(S_1)$. Let $\Delta=\mathbb{C}\setminus\mathrm{trace}(\Gamma)$, then the function

(2.5) $\quad f_t(z) = \sin^2(t \arcsin\sqrt{z})$

is defined for t complex, and $z\varepsilon(0,1)$. This function can be continued to an analytic function on Δ. The function $f_t(z)$ has the properties:

$$f_{\frac{1}{2}}(z)=\tfrac{1}{2}\,(1-\sqrt{1-z})\ ,\ f_2(z)=4z(1-z),\ f_2(f_t(z))=f_{2t}(z)$$

$$f_t(z)\neq 1\ ,\ t=2^{-j}\ ,\ j\geq 0 \text{ an integer.}$$

$$f_t(z)\neq\tfrac{1}{2}\ ,\ t=2^{-j}\ ,\ j\geq 1 \text{ an integer.}$$

We define

(2.6) $\quad S_t = f_t(S_1)$ for all t complex.

It follows that $S_{t/2} = \frac{1}{2}\,(I-\sqrt{I-S_t})$, so that

(2.7) $\quad (I-2S_{t/2})^2 = I-S_t$.

We construct, by induction, involutions T_t for $t=2^{-j}$, $j\geq 0$ integer such that

(2.8) $\quad T_t^2 = I\ ,\ \dfrac{(T_0+T_t)^2}{4} = I - S_t$ and T_t commutes with S_t..

By Lemma 2.4 T_1 satisfies (2.8). Assuming T_t is constructed satisfying (2.8), we construct $T_{t/2}$ as follows:

$$(2.9) \quad T_{t/2} = \left(\frac{T_0+T_t}{2}\right) \ (I-2\ S_{t/2})^{-1} .$$

Note that $(I-2\ S_{t/2})^{-1}$ is well defined since $f_t(z) \neq 1/2$ $t=2^{-j}$ $j \geq 1$. It can be shown that $T_{t/2}$ satisfies (2.8) and furthermore $T_{t/2}\ T_0 = T_t\ T_{t/2}$ for $t=2^{-j}$, $j \geq 0$. In particular for $t=1$ this implies that $T_{1/2}\ T_0 = T_1\ T_{1/2}$ so that $T_{1/2}\ E_0 = E_1\ T_{1/2}$ and $T_{1/2} \varepsilon A\{E_0,E_1\}$, as it is clear from (2.9). With each T_t, associate V_t; $V_t = T_t T_0$. Then we have

$$(2.10) \quad V_{t/2} = \tfrac{1}{2}(I+V_t)\ (I-2S_{t/2})^{-1} = \tfrac{1}{2}(I-2S_{t/2})^{-1}(I+V_t)$$

$$= T_{t/2}T_0 = T_t T_{t/2} .$$

From (2.10) it follows that $V^2_{t/2} = V_t$ and hence $V^{2^j}_{2^{-j}} = V_1$. In addition it can be shown by induction and by using equation (2.10), that $V_t\ \varepsilon\ A(V_1), t=2^{-j}$ $j \geq 0$ an integer. Also as one would expect, $\lim_{j\to\infty} V_{2^{-j}} = I$. The construction of K goes along this line: there exists j_0 such that for any $j \geq j_0$, $\|I-V_{2^{-j}}\| < 1$. Define $2K_j = 2^j \log V_{2^{-j}}$, where the right hand side is defined by the Taylor series. Since $V^2_{t/2} = V_t$, it follows that $2^j \log V_{2^{-j}}$ is independent of j as long as it is defined. We call this value $2K$.

Now $(2K)\,2^{-j} = \log V_{2^{-j}}$ $j \geq j_0$, so that $(V_{2^{-j}})^{2^j} = \exp 2K$. Hence $V_1 = \exp 2K$ and $K \in A(V_1)$. Similarly, for $t = 2^{-j}$ $j \geq 0$, $V_{t/2} = \exp tK$. In particular $V_{1/2} = \exp K$ so that $V_{1/2} T_0 = T_1 V_{1/2}$. Thus $V_{1/2}$ gives a similarity between E_0, E_1 . Now for $t = 2^{-j}$ $j \geq 0$,

$$(2.11) \quad T_t = e^{tK}\, T_0\, e^{-tK} .$$

For $0 \leq t \leq 1$ we use (2.11) to define T_t and $V_t (= T_t T_0)$, so that T_t are involutions and $E_t = e^{tK}\, E_0\, e^{-tK}$ is a projector valued path connecting E_0 and E_1 as claimed in the theorem. ///

Remark 2.7. We note that both $T_{1/2}$ and $V_{1/2}$ provide a similarity between E_0 and E_1 , but one would use $V_{1/2}$ to measure how far E_1 is from E_0 since $V_{1/2} \to I$ as $E_1 \to E_0$. Thus if E_1 is a perturbation of E_0 , $V_{1/2}$ will be a perturbation of the identity, while the spectrum of any nontrivial involution is always $\{-1, 1\}$.

Now given two projectors E_0 , E_1 , $E_0 \neq E_1$, we ask under what conditions they can be connected by a projector valued path which is a straight line segment. The answer is given in the next proposition with a description of such pairs.

Proposition 2.8. *Let E_0, E_1 be different projectors in $B(X)$. Then the following statements are equivalent*

(*i*) $E_{t_0} = E_0 + t_0(E_1 - E_0)$ *is a projector for some* t_0, $t_0 \neq 0$, $t_0 \neq 1$.

(*ii*) $S_1(E_0, E_1) = 0$.

(*iii*) E_0 *and* E_1 *can be connected by the projector*

$$t \longmapsto E_t = E_0 + t(E_1 - E_0) \ , \quad 0 \leq t \leq 1 \ .$$

In each case the pair E_0 , E_1 *has the following block matrix representation with respect to the frame* $\{E_0, (I-E_0)\}$;

$$[E_0] = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad [E_1] = \begin{bmatrix} I & A \\ B & 0 \end{bmatrix} \ ,$$

where $AB=0$ *and* $BA=0$.

Proof: (ii) $\Rightarrow$ (iii). If $S_1(E_0,E_1) = 0$ then it follows from Theorem 2.6 that E_0 , E_1 can be connected by a projector valued path $E_t = \exp(tK)\, E_0 \exp(-tK)$, $0 \leq t \leq 1$. From (2.3) (vii) it follows that $(I-V_1)^2 = 0$, so that

$$K = -\frac{1}{2} \sum_{j=1} \frac{(I-V_1)^j}{j} = -\frac{1}{2}(I-V_1) = (E_1E_0 - E_0E_1) \ .$$

But

$$(E_1E_0 - E_0E_1)^2 = (E_1-E_0)^4 - (E_1-E_0)^2 ,$$

(cf. [22] Eq. (4.35)); hence $K^2=0$. The above path becomes

$$E_t = E_0 + t(E_1-E_0) \qquad 0\le t\le 1 .$$

(i) => (ii), suppose that there is a projector which is colinear with E_0, E_1, that is

$$E_{t_0} = E_0 + t_0(E_1 - E_0) \qquad t_0 \neq 0 , \qquad t_0 \neq 1$$

is a projector. Since $E_{t_0}^2 = E_{t_0}$, it follows that

$$0 = E_{t_0}^2 - E_{t_0} = (t_0^2 - t_0)\ S_1 = 0$$

and $S_1 = 0$.

(iii) =>(i) is clear. For the characterization of the pair E_0, E_1 in that case see [25] Remark (c). ///

Remarks 2.9.

(a) In the case of a Hilbert space and a pair of orthoprojectors, the condition $S_1=0$ would imply that E_0, E_1 have the forms given in Proposition 2.8 with $B=A^*$, so that $A^*A=0$, hence $A=0$ and $E_0=E_1$. This means that the separation operator in this case is a much stronger measure, in that no two distinct orthoprojectors can be connected by an orthoprojector valued path which is a straight line segment.

(b) If E_0, E_1 is a pair of projectors in a Banach space then $\frac{1}{2}(E_0+E_1)$ is not in general a projector. However, it would be a projector if and only if $S_1(E_0,E_1)=0$, and $\frac{1}{2}(E_0+E_1)$ will correspond to $E_{1/2}$ in the projector valued path $t \longmapsto E_0+t(E_1-E_0)$.

As we have mentioned in the above remark, $\frac{1}{2}(E_0+E_1)$ need not be a projector but the bisector of E_0 and E_1 will be obtained by "straightening" $\frac{1}{2}(E_0+E_1)$ to be a projector, namely $E_{1/2} = \frac{1}{2}(T_{1/2}+I)$ will be the bisector of E_0, E_1 where $T_{1/2}$ is given by equation (2.9) with $t=1$. $E_{1/2}$ will be the angle bisector of the acute angle between E_0, E_1. If we put $iW=K$ in Theorem 2.6, then $V_1=e^{2iW}$ and $S_1 = \frac{1}{4}(2-e^{2iW} - e^{-2iW}) = \sin^2 W$. Hence W resembles the oriented angle operator between E_0 and E_1. In the next lemma we generalize some trigonometric relations. They will show that the particular similarity $V_{1/2}$ is related to the particular angle bisector $E_{1/2}$ in a way which generalizes the known 2-space facts. This is similar to the results in [6], Section 4 in the Hilbert space setting with a pair of orthoprojectors.

Lemma 2.10.

(i) $\sin^2 W(E_0, E_{1/2}) = \sin^2 W(E_{1/2}, E_1)$

(ii) $V_{\frac{1}{2}} + V_{\frac{1}{2}}^{-1} = 2\ C_1(E_0,E_1)^{1/2}$

(iii) $C_1(E_0,E_{\frac{1}{2}}) = \frac{1}{2}[I+C_1(E_0,E_1)^{1/2}] = C_1(E_{\frac{1}{2}},E_1)$

(iv) $V_{\frac{1}{2}}(E_0,E_{\frac{1}{2}}) = V_{\frac{1}{2}}(E_{\frac{1}{2}},E_1)$

(v) $V_{\frac{1}{2}}(E_0,E_{\frac{1}{2}})^2 = V_{\frac{1}{2}}(E_0,E_1)$.

Proof: Since $\sin^2 W(E_0,E_{\frac{1}{2}}) = S_1(E_0,E_{\frac{1}{2}}) = \frac{1}{4}(T_{\frac{1}{2}}-T_0)^2 =$
$= \frac{1}{4}(I-T_{\frac{1}{2}}T_0-T_0T_{\frac{1}{2}})$, and $T_{\frac{1}{2}}T_0 = T_1T_{\frac{1}{2}}$ then

$$\sin^2 W(E_0,E_{\frac{1}{2}}) = \frac{1}{4}(2I-T_1T_{\frac{1}{2}}-T_{\frac{1}{2}}T_1) = S_1(E_{\frac{1}{2}},E_1) = \sin^2 W(E_{\frac{1}{2}},E_1) .$$

From (2.10) it follows that

$$\frac{1}{2}(V_{\frac{1}{2}}+V_{\frac{1}{2}}^{-1}) = \frac{1}{4}(I+V_1)\ (I-2\ S_{\frac{1}{2}})^{-1} + \frac{1}{4}(I+V_1^{-1})\ (I-2\ S_{\frac{1}{2}})^{-1}$$

$$= \frac{2I+V_1+V_1^{-1}}{4}\ (I-2\ S_{\frac{1}{2}})^{-1} = (I-S_1)\ (I-2\ S_{\frac{1}{2}})^{-1}$$

$$= C_1(E_o,E_1)^{\frac{1}{2}} .$$

Part (iii) can be proved similarly. From (2.10), (iv) and (v) follow directly. ///

2.3. Characterization of the Direct Rotation

In the previous section, the construction of $V_{1/2}$ was dependent on the function $f_t(z)$ defined by equation (2.5), where the domain of $f_t(z)$ is $\mathbb{C}\backslash\Gamma$ with Γ being a smooth simple curve connecting 1 , ∞ and not passing through zero. Hence $V_{1/2}$ may not be unique. It was an open problem in [25], to find conditions on the pair $\{T_0, T_1\}$ which would allow us to characterize $T_{1/2}$ or $V_{1/2}$ before it is constructed. Results of this type for the direct rotation between a pair of orthoprojectors were given in [6, Theorem 4.1] and [9, Proposition 3.3]. In this section we will give a characterization of the direct rotation between a pair of projectors in Banach spaces. First we need the following definition.

Definition 2.11. *Let T_0 and T_1 be two involutions. A direct rotation between T_0 and T_1 is $R \in G(X)$, which satisfies the following conditions*

(*i*) $T_1 = RT_0R^{-1}$

(*ii*) $R = \exp K$ *where* $K = \frac{1}{2}\log T_1T_0$ (*i.e.* $R^2 = T_1T_0$) .

We call R the *principal direct rotation* if we use the principal branch of the logarithm defining K . Also we note that the definition of K implies that $KT_0 + T_0K = 0$.

Here we use a holomorphic branch of the natural logarithm whose value at 1 is zero. Before stating the main theorem in this section we need some lemmas which also introduce some notation.

Lemma 2.13. *The function* $f : f(z) = \frac{z-1}{z+1}$ *maps conformally*

$$\Omega = \{z: \operatorname{Re} z>0\} \text{ onto } D = \{z: |z|<1\} .$$

Proof: The boundary of Ω, parametrized by $t \longmapsto it$, $-\infty<t\leq\infty$, is mapped onto the unit circle $t \longmapsto \frac{it-1}{it+1}$, and the interior point $1\varepsilon\Omega$ is mapped onto $0 \varepsilon D$. Hence the result follows by the Orientation Principle [5, III, 3.21]. ///

Lemma 2.14. *The function* $h(z) = \frac{2z}{1-z^2}$ *maps the set* $D = \{z: |z|<1\}$ *conformally onto the set* $G = \mathbb{C}\setminus\{z: z=x+iy$, $x=0$, $|y|\geq 1\}$ *while* $p(z) = \frac{z}{1+\sqrt{1+z^2}}$ *maps* G *conformally onto* D.

Proof: A standard result in conformal mapping [33] is that $z \longmapsto \frac{1}{2}(z+\frac{1}{z})$ maps D onto the complement of $[-1,1]$ and one composes this with rotation to get the map $w=f(z)=$ $=\frac{1}{2}(\frac{1}{z}-z)$ which maps D conformally onto the whole w-plane cut along the imaginary axis from $-i$ to $+i$. For the inverse, the quadratic equation $z^2+2wz-1=0$ has a solution

$$z = -w+\sqrt{1+w^2}$$

which maps the w-plane cut along the imaginary axis from $-i$ to $+i$ onto D, since the other solution $z=-w-\sqrt{w^2+1}$ maps ∞ to ∞. Composing the map f with inversion we get

$$h(z) = \frac{1}{\frac{1}{2}(\frac{1}{z}-z)} = \frac{2z}{1-z^2} \quad .$$

and $h(z)$ has the required mapping properties. To find the inverse of h, we choose that branch of f^{-1} which maps the complement of $[-i,i]$ onto D so that

$$p(z) = \frac{-1}{z} + \sqrt{(\frac{1}{z})^2+1} = \frac{z}{1+\sqrt{1+z^2}}$$

will map G onto D. ///

We will use the previous lemmas to show that an operator $R\varepsilon\mathcal{B}(X)$ with the property $\mathrm{Re}\, z>0$ for all $z\varepsilon\sigma(R)$ can be characterized in the following way.

Lemma 2.14. *If* $R\varepsilon\mathcal{B}(X)$, *then* $\mathrm{Re}\,\sigma(R)>0$ *if and only if* $R=g(H)$, *where* $H\varepsilon\mathcal{B}(X)$, $r(H)<1$ *and*

$$g(z) = \frac{1+z}{1-z} \quad .$$

Proof: Suppose $\mathrm{Re}\,\sigma(R)>0$, then Lemma 2.12 shows that the function $f(z)=\frac{z-1}{z+1}$ maps Ω onto D . Defining

$$H = f(R) \ ,$$

the spectral mapping theorem [Chapter I, Theorem 1.6] shows that $\sigma(H)\subset D$, hence $r(H)<1$. Now, since $g(z) = \frac{1+z}{1-z}$ is the inverse of f , we have $R=g(H)$ as claimed. On the other hand, if $R=g(H)$ with g and H satisfying the above conditions, then Lemma 2.12 together with spectral mapping theorem [Chapter I, Theorem 1.6] imply that $\mathrm{Re}\,\sigma(R)>0$.

If we set $L=R-R^{-1}$, then L can be expressed as a function of H as the following lemma shows.

Lemma 2.15. *Let* $R\varepsilon B(X)$, *then* $\mathrm{Re}\,\sigma(R)>0$ *if and only if* $L=R-R^{-1}=h(H)$, *with* $h(z)=\frac{2z}{1-z^2}$. *In this case* $H = p(L)$ *where* $p(z) = \frac{z}{1+\sqrt{1+z^2}}$.

Proof: The proof depends mainly on mapping properties of the functions h and p which were already discussed in Lemma 2.13, using the spectral mapping theorem [Chapter I, Theorem 1.6] the results follow. ///

Let $\Omega^2 = \mathbb{C}\backslash\{z: z\leq 0\}$, then if we define $\log z$, $z\in\Omega^2$ as the principal branch of the logarithm and write $\log z=\log r+i\theta$, then this function is holomorphic and satisfies

$$\exp(\log z) = z \quad , \qquad z\in\Omega^2$$

and

$$\log(\exp z)=z \qquad (-\pi<\mathrm{Im} z<\pi) .$$

If $A,B\in\mathcal{B}(X)$ and $\sigma(A)\subset\Omega^2$, $\sigma(B)\subset\{z:-\pi<\mathrm{Im} z<\pi\}$, then we have $\exp\log A=A$, and if we assume that $\sigma(\exp(B))\subset\Omega^2$, then $\log(\exp B)=B$. For the proof of these statements we refer to [2, I, 8.3].

<u>Theorem 2.16</u>. *Let* T_0 , $T_1\in B(X)$ *and suppose that* $\sigma(T_1T_0)$ *lies in* Ω^2 , $\Omega^2=\mathbb{C}\backslash\{z:z\leq 0\}$. *Then* $R\in G(X)$ *is a principal direct rotation if and only if*

(1) $T_1 = RT_0R^{-1}$

(2) *block diagonal part of* R = *block diagonal part of* R^{-1}

(3) $\mathrm{Re}\,\sigma(R)>0$

Proof: If R satisfies the conditions (1)-(3), then Lemma 2.14 implies that

$$R = (I-H)^{-1}\,(I+H)$$

where $r(H) < 1$. This restricts those R for which condition (3) is satisfied. Since the block diagonal part of R is

$$E_0RE_0+(I-E_0)R(I-E_0) = \tfrac{1}{2}(R+T_0RT_0)$$

condition (2) leads to

$$L+T_0LT_0 = 0\ , \quad L = R-R^{-1}\ .$$

In our block matrix notation this means that L is block off diagonal. This in turn will imply that

$$T_0L^2T_0 = -T_0LT_0L = L^2\ ,$$

and consequently

$$\tfrac{1}{2}(L^2+T_0L^2T_0) = L^2\ ,$$

so that L^2 is block diagonal. From lemma 2.14, 2.15, it follows that $\sigma(L)\subset G$ which in turn implies that $\sigma(I+L^2)$ does not intersect the negative real axis. Since $I+L^2$ is a block diagonal, the square root of $I+L^2$ can be calculated block by block. Hence

$$H = L\,(I+\sqrt{I+L^2})^{-1}$$

will be block off diagonal, since multiplication by a diagonal matrix does not change the structure. Define

$$(2.13)\quad K = \xi(H) ,$$

where

$$\xi(z) = 2 \text{ arc } \tanh z = \log\frac{1+z}{1-z} .$$

Here $\log\frac{1+z}{1-z} = 2 \sum_{n=1}^{\infty} \frac{1}{2n-1} z^{2n-1} \qquad |z|<1 .$

It then follows that

$$K = 2 \sum_{n=1}^{\infty} \frac{1}{2n-1} H^{2n-1} .$$

It can be shown that odd powers of H will be block off diagonal since H itself is block off diagonal. Hence indeed

$$KT_0 + T_0K = 0 .$$

Furthermore

$$K = \log(I-H)^{-1} (I+H) = \log R .$$

Thus

$$R = \exp K = \exp(-T_0KT_0) = T_0(\exp -K)T_0 ,$$

hence

(2.14) $R^2 = (\exp K)\, T_0\, (\exp -K)\, T_0 = T_1 T_0$.

This follows from condition (1). Equation (2.13) implies

$$H = (e^K - I)\,(e^K + I)^{-1} \quad .$$

But the function $\dfrac{e^z - 1}{e^z + 1}$ maps the infinite strip $\{z : |\operatorname{Im} z| < \pi/2\}$ onto the open unit disk D . This implies that $\sigma(K) \subset \{z : |\operatorname{Im} z| < \pi/2\}$ so that $\sigma(2K) \subset \{z : |\operatorname{Im} z| < \pi\}$. By (2.14) $R^2 = T_1 T_0$ and $R = \exp K$. Hence $\exp 2K = R^2 = T_1 T_0$ so that $K \equiv \frac{1}{2} \log T_1 T_0$ and R is a principal direct rotation.

On the other hand if $\sigma(T_1 T_0) \subset \Omega^2$, then by (2.3) (i) we have, $S_1 = \frac{1}{4}(2I - T_1 T_0 - T_0 T_1)$. So

$$I - S_1 = \phi(T_1 T_0)$$

where $\phi(z) = \frac{1}{4}(2 + z + \frac{1}{z})$. Under the assumption on $\sigma(T_1 T_0)$ it follows also that $\sigma(I - S_1)$ does not intersect the negative real axis so that $\sigma(S_1) \cap [1, \infty) = \emptyset$. Thus $\sigma(S_1)$ satisfies the assumption of Theorem 2.6. Hence if we take $\Gamma = \Gamma_0 = [1, \infty)$ and the principal branch of the logarithm, then we have from (2.10) that

$$V_{\frac{1}{2}} = \exp K \ , \quad K = \frac{1}{2} \log T_1 T_0 \ , \quad \text{and} \quad V_{\frac{1}{2}}^2 = T_1 T_0 \ .$$

Further $L = R - R^{-1} = \sinh K$ and

$$T_0 L\, T_0 = T_0 \sinh K \;\; T_0^{-1} = \sinh(T_0 K T_0) = -L$$

since $K = \frac{1}{2}\log T_1 T_0$ implies that $T_0 K T_0 = -K$. Now $R = \exp K$, $K = \frac{1}{2}\log T_1 T_0$ implies that $\mathrm{Re}\,\sigma(R) > 0$. This completes the proof. ///

We remark here that $T_{1/2}$ satisfies the conditions (1), (2) of Theorem 2.16 but not condition (3) (compare this with Remark 2.7).

2.4. Riemannian Geodesics Between Nonsymmetric Involutions.

The Riemannian geodesics between symmetric involutions were investigated in [26]. The geodesics are identified as minimal arcs between a pair of symmetric involutions whose straight line distance is less than 2. In this section we study the problem of Riemannian geodesics between nonsymmetric involutions. First we formulate the problem. Let T_0 be a (non-trivial) nonsymmetric involution on a Hilbert space H . Then the similarity orbit of T_0 is

$$(2.15) \quad \mathcal{Y}(H, T_0) = \{V T_0 V^{-1} \; , \; V \in G(H)\} \; .$$

we define

(2.16) $\mathcal{Y}_{HS}(H,T_0) = \{T \varepsilon \mathcal{B}(H)\ ,\ T-T_0 \varepsilon HS(H)$ and T is similar to $T_0\}$

$$= \mathcal{Y}(H,T_0) \cap (T_0 + HS(H))$$

where $HS(H)$ denotes the class of Hilbert-Schmidt operators. We also define

(2.17) $M_{HS}(T_0) = M(T_0) \cap HS(H)$,

where

$$M(T_0) = \{H \ \varepsilon\ \mathcal{B}(H) : T_0H + HT_0 = 0\} .$$

Now if the topology of $\mathcal{Y}_{HS}(H,T_0)$ is strengthened to be compatible with the Hilbert-Schmidt metric, we have the following theorem.

<u>Theorem 2.17</u>. *$\mathcal{Y}_{HS}(H,T_0)$ is a manifold modeled on the Hilbert space $M_{HS}(T_0)$.*

<u>Proof</u>: Since it is known that a C^∞ structure is determined by any C^∞ atlas on $\mathcal{Y}_{HS}(H,T_0)$ (cf. [4], Proposition 2.21), then in order to show that $\mathcal{Y}_{HS}(H_1T_0)$ is a manifold, we need not specify a complete atlas. We first describe a chart at T_0 . We define an open convex $M_{HS}(T_0)$ neighbourhood V_0 of 0 by

$$V_0 = \{H \in \mathcal{M}_{HS}(T_0) : \|T_0 H\|_{HS} < \log 2\} ,$$

and define the mapping

$$X^{-1} : H \longmapsto T_0 \exp(T_0 H) .$$

We claim that the range of X^{-1} is in $\mathcal{V}_{HS}(\mathcal{H},T_0)$. Now

$$X^{-1}(H) = T_0 \exp(T_0 H) = T_0 \exp(\tfrac{1}{2}T_0 H) \exp(\tfrac{1}{2}T_0 H)$$

$$= T_0 \exp(\tfrac{1}{2}T_0 H) T_0^2 \exp(\tfrac{1}{2}T_0 H) = \exp(-\tfrac{1}{2}T_0 H) T_0 \exp(\tfrac{1}{2}T_0 H)$$

(because $T_0 H T_0 = -H$) . Hence $X^{-1}(H)$ is similar to T_0 . Let $K = \tfrac{1}{2}T_0 H$, thus

$$X^{-1}(H) - T_0 = \exp(-K) T_0 \exp K - T_0$$

$$= (\exp(-K)-I) T_0 \exp K + T_0 (\exp K-I) .$$

Since $H \in HS(\mathcal{H})$, $\exp(\pm K)-I$ is in $HS(\mathcal{H})$ and so is $X^{-1}(H)-T_0$. X^{-1} has a formal inverse

$$X : T \longmapsto T_0 \log(T_0 T) .$$

To justify X , let T be in the range V of X^{-1} . That is $T = T_0 \exp(T_0 H)$, for some H in V_0 . Thus

$$\|T_0 T-I\|_{HS} = \|\exp T_0 H-I\|_{HS} \le \exp\|T_0 H\|_{HS}-1 < \exp\log 2-1 = 1.$$

Hence X is well defined on V and (V,X) is a chart at T_0. Since the mapping $(W,T) \longmapsto WTW^{-1}$ is a smooth action of $G(H)$ on $B(H)$, and the restricted action to $V_{HS}(H,T_0)$ will be transitive from the definition (2.16) of $V_{HS}(H,T_0)$, then we can describe a chart at any other point. It follows that the change of coordinates is smooth. Namely, if (U,X_U) is a chart at T_0, (V,X_V) is a chart at T_1 and $U \cap V \neq \emptyset$, then $X_V X_U^{-1} : X_U(U \cap V) \longrightarrow X_V(U \cap V)$ is a smooth map of open subsets of $M_{HS}(T_0)$. Now since X described above is a homeomorphism, the manifold topology will coincide with the topology given to $V_{HS}(H,T_0)$. This proves the theorem. ///

Let T be any point in $V_{HS}(H,T_0)$. As the space of tangency classes of smooth curves through T (See [29, IV.2]), the tangent space can be identified with the subspace $M_{HS}(T)$ of $HS(H)$ in the following manner. If $\alpha : (-\varepsilon,\varepsilon) \to B(H)$ has values in $V_{HS}(H)$ and $\alpha(0)=T$, then $(\alpha^2)'(0) = \alpha'(0)T + T\alpha'(0) = 0$ and $\alpha'(0) \in M(T)$. But $\alpha(t)-T_0 \in HS(H)$ implies that $\alpha'(0) \in HS(H)$. Hence $\alpha'(0) \in M_{HS}(T)$. On the other hand if $H \in M_{HS}(T)$, then $\alpha(t) = T\exp(tH)$ satisfies the following relations

$$\alpha^2(t) = I \text{ for all } t\ , \alpha(0)=T\ , \alpha'(0)=H \quad \alpha(t)-T_0 \in HS(H)\ .$$

Now $V_{HS}(H,T_0)$ with $M_{HS}(T)$ equipped with the Hilbert-

Schmidt inner product at each $T \in \mathcal{V}_{HS}(\mathcal{H},T_0)$ will be a Riemannian manifold.

A *tangent vector field* H on $\mathcal{V}_{HS}(\mathcal{H},T_0)$ - neighbourhood U of T_0 is a smooth function such that for each $T \in U$ we have $H(T) \in \mathcal{M}_{HS}(T)$. By an *affine connection* we mean a function of two vector fields whose value is again a vector field, that is, $(H,K) \longmapsto D_K H$. Moreover it has the linearity properties: For all f,g smooth functions on U and K,L,M vector fields we have

(1) $D_{fK+gL} H = f \cdot D_K H + g \cdot D_L H$

(2) $D_K(fH+gL) = f \cdot D_K H + g \cdot D_K L + (f'K) H + (g'K) L$.

A connection is said to be symmetric if it satisfies

(3) $[K,H] = D_K H - D_H K$,

where the Lie bracket $[K,H] = H'K - K'H$. Now if we have a Riemannian manifold and a symmetric connection which also satisfies

(4) $(D_K H,L) + (H,D_K L) = (H,L)'K$,

then the connection is called the *Riemannian connection*. If P_T^R denotes the orthoprojector from $HS(\mathcal{H})$ onto $\mathcal{M}_{HS}(T_K)$, then

(2.18) $(D_K^R H)(T) = P_T^R(H'(T)K(T))$, defines a Riemannian connection on $V_{HS}(H,T_0)$. The proof of the previous statement goes similar to that of the symmetric case given in [26]. In principle there always exists an orthoprojector of $HS(H)$ onto $M_{HS}(T)$. The following lemma shows how one would construct such a projector.

<u>Lemma 2.18</u>. *If $P\varepsilon B(H)$ is an oblique projector, then the orthogonal projector onto PH is given by*

$$Q = P(P+P^*-I)^{-2} \quad P^* \quad .$$

<u>Proof</u>: We will construct Q by means of the operational calculus. First we show that (P^*+P-I) is invertible. Using the identities in Lemma 2.4, we have

$$C = (P^*+P-I)^2 = \tfrac{1}{4}(T^*+T)^2 = I - \tfrac{1}{4}(T^*-T)^2 \geq I \ ;$$

this is because T^*-T is skew symmetric. Hence $0\notin\sigma(C)$. If $Q=PP^*$ is a projector, then by its definition it will be an orthogonal projector, in which case $||P||=1$ and consequently P itself will be an orthogonal projector. Thus $Q=P$ will be the required projector. Since $N(PP^*)\supset N(P^*)$, this implies $0\varepsilon\sigma(PP^*)$. Further for $\lambda\neq 0$, $\lambda\notin\sigma(C)$ we have the following partial fractions identity:

$$(\lambda I-PP^*)[\lambda^{-1}I+PC^{-1}((\lambda I-C)^{-1}-\lambda^{-1}I)P^*] =$$

$$= [\lambda^{-1}I+PC^{-1}((\lambda I-C)^{-1}-\lambda^{-1}I)P^*](\lambda I-PP^*) = I \ ,$$

so that

$$(\lambda I-PP^*)^{-1} = \lambda^{-1}I+PC^{-1}((\lambda I-C)^{-1}-\lambda^{-1}I)P^* \ .$$

The above equation shows that 0 is an isolated point of $\sigma(PP^*)$. Hence we can define the spectral projector

$$I-Q = \frac{1}{2\pi i}\int_{|\lambda|=\varepsilon} (\lambda I-PP^*)^{-1}\, d\lambda$$

$$= \frac{1}{2\pi i}\int_{|\lambda|=\varepsilon} [\lambda^{-1}I+PC^{-1}\ ((\lambda I-C)^{-1}-\lambda^{-1}I)P^*]d\lambda$$

$$= \frac{1}{2\pi i}\int_{|\lambda|=\varepsilon} \frac{d\lambda}{\lambda} + \frac{1}{2\pi i}\int_{|\lambda|=\varepsilon} PC^{-1}(\lambda I-C)^{-1}P^* d\lambda + \frac{1}{2\pi i}\int_{|\lambda|=\varepsilon} PC^{-1}\lambda^{-1}P^* d\lambda$$

$$= I-PC^{-1}P^* \ ,$$

so that $Q=PC^{-1}P^*$. One can easily check that $PQ=Q$ and $QP=P$ so that $R(Q)=R(P)$, also $P^*Q=P^*$, $QP^*=Q$ imply $N(Q)=N(P^*)$ and hence Q is the orthogonal projector on $R(P)$. The lemma is proved. ///

It was shown in [26, Lemma 1], that $M(T)$ can be complemented and $P_T : P_T H = \frac{1}{2}(H-THT)$ is a projector with range $M(T)$. According to Lemma 2.18, it follows that the orthoprojector P_T^R on $M_{HS}(T)$ is

$$P_T^R = P_T(P_T+P_T^*-I)^{-2}\ P_T^* \quad .$$

A curve through T_0 is called a *geodesic* of D^R if $D^R_{\alpha'}\alpha'=0$ on the domain of α . For D^R defined by (2.18) the resulting second order differential equation is

$$(2.19) \qquad P_T(P_T+P_T^*-I)^{-2}\ P_T^*(\alpha'') = 0$$

In principle the differential equation (2.19) has locally a unique solution with initial conditions $\alpha(0)=T_0$ and $\alpha'(0)=H_0$ where $H_0 \in M_{HS}(T_0)$. In the following proposition we will simplify the differential equation and then reduce it to a system of first order differential equations using Kato's transformation, introduced in Chapter 1.

Proposition 2.19. *If $T_0 \in B(H)$ is a nonsymmetric involution and if $H_0 \in M_{HS}(T_0)$, $H_0 \neq 0$, then the Riemannian geodesic through T_0 in the direction of H_0 is the solution of the differential equation*

$$(2.20) \qquad \alpha''+(\alpha^*\alpha)\alpha''(\alpha\alpha^*) + 2(\alpha^*\alpha)\alpha(\alpha)^2(\alpha\alpha^*) = 0 \ ,$$

with the initial conditions $\alpha(0)=T_0$ and $\alpha'(0)=H_0$.

Further, the above equation can be reduced to a system of first order differential equations (2.23).

Proof: The Riemannian geodesic through T_0 in the direction of H_0 is the solution of (2.19) with initial conditions $\alpha(0)=T_0$, $\alpha'(0)=H_0$. First we show that $P_T^*=P_{T^*}$. Recall that $P_T(H)=\frac{1}{2}(H-THT)$, and the scalar product in $HS(\mathcal{H})$ is given by $(A,B)=\operatorname{tr} AB^* = \operatorname{tr} B^*A$. Hence

$$(P_T^*H,K) = (H,P_TK) = \operatorname{tr} H\,(P_TK)^*$$

$$= \operatorname{tr} \frac{1}{2} H(K-TKT)^* = \operatorname{tr} \frac{1}{2} H(K^*-T^*K^*T^*)$$

$$= \frac{1}{2} \operatorname{tr} HK^* - \frac{1}{2} \operatorname{tr} HT^*K^*T^*$$

$$= \frac{1}{2} \operatorname{tr} HK^* - \frac{1}{2} \operatorname{tr} T^*HT^*K^*$$

$$= \frac{1}{2} \operatorname{tr} (H-T^*HT^*)K^* = (P_{T^*}H,K)$$

hence $P_T^*=P_{T^*}$, and since $N(P_T^R)=N(P_T^*)$, equation (2.19) reduces to $P_{T^*}(\alpha'')=0$. Since

$$(2.21)\qquad P_{T^*}(H'(T)K(T)) = \tfrac{1}{2}(H'(T)K(T) - T^*(H'(T)K(T))T^*)$$

with H a vector field, implies $H(T)\in M_{HS}(T)$, that is $TH(T)T+H(T)=0$, differentiating along K will yield

$$(2.22) \quad K(T)H(T)T + T(H'(T)K(T))T + TH(T)K(T) + H'(T)K(T)=0 \ .$$

Substituting from (2.22) into (2.21), we get

$$P_{T*}(H'(T)K(T)) = \frac{1}{2}[H'(T)K(T) - (T^*T)T(H'(T)K(T))T(TT^*)]$$

$$= \frac{1}{2}(H'(T)K(T) + (T^*T)(H'(T)K(T))(TT^*) + \frac{1}{2}(T^*T)T(KH+HK)(TT^*)=0$$

$$P_{T*}\ \alpha'' = \frac{1}{2}[\alpha'' + \alpha^*\alpha\ \alpha''\ \alpha\alpha^* + 2(\alpha^*\alpha)\ (\alpha'^2)\ (\alpha\alpha^*)] \ .$$

Hence (2.20) is the differential equation for the Riemannian geodesics as claimed. In reducing it to a system of first order differential equations, the basic idea is Kato's transformation [Chapter I, Section 4]. If α is a smooth path of involutions in $\mathcal{Y}_{HS}(\mathcal{H},T_0)$ and passing through T_0 then there exists a transformation function $U(t)$ such that

$$\alpha(t) = U(t)\alpha(0)U^{-1}(t)$$

where

$$U'(t) = L(t)\ U(t), U(0)=I \ , \quad L(t)=\frac{1}{2}\ \alpha'(t)\alpha(t).$$

Now α being a geodesic, it satisfies

$$\alpha'' + \alpha^*\ \alpha''\ \alpha^* = 0 \ .$$

But $L=\frac{1}{2}\ \alpha'\alpha$, so

$$\alpha'' = 2L'\alpha + 4L^2\alpha = 2L'\, U\alpha(0)U^{-1} + 4L^2\, U\alpha(0)U^{-1} \ .$$

Hence we obtain the system,

$$(2.23) \qquad [2L'\, U\alpha(0)U^{-1} + 4L^2\, U\alpha(0)U^{-1}](U^{-1})^* \,\alpha(0)^*\, U^* +$$

$$+ (U^{-1})^* \,\alpha^*(0)\, U^*[2L'\, U\alpha(0)\, U^{-1} + 4L^2\, U\alpha(0)U^{-1}] = 0$$

and $U'(t) = L(t)\, U(t)$

with initial conditions

$$U(0)=I \ , \ L(0)=\tfrac{1}{2}\, \alpha(0)\, \alpha'(0) \ . \qquad ///$$

CHAPTER III

CLOSENESS TO GEODESICS

In this chapter we study the similarity between a pair of n-frames E and F. Special attention will be paid to two particular similarities, namely the balanced transformation and the direct rotation which arises from the geodesic connecting E,F. We give a local characterization of the balanced transformation. We also show that, in general, the balanced transformation is different from the direct rotation. Nevertheless, the balanced transformation defines a frame-valued path, connecting E and F, which is close to the geodesic connecting E and F. We also show that the difference between the direct rotation and the balanced transformation is of order $\|E-F\|^3$. Finally, we show that this order cannot be improved in general. Throughout this chapter and the next chapter we make the assumption that n is fixed, and we say simply "frame" to mean "n-frame".

3.1. Closeness and Similarity Between Frames

We begin by recalling that E being a frame means that $E \varepsilon \mathcal{B}^n(X)$ and that the components E_j of E satisfy

$$(3.1) \qquad E_j \neq 0 \ , \quad \sum_{j=1}^{n} E_j = I \ , \ E_j E_k = \delta_{jk} E_k \quad 1 \leq j,k \leq n \ .$$

Recall also that the set of all $E \varepsilon B^n(X)$ and satisfying (3.1) is denoted by $E^n(X)$. Two frames E and F are said to be similar if there exists an invertible operator $V \varepsilon B(X)$ such that $VE_j = F_j V$, $j=1,\ldots,n$, or shortly $VE=FV$. In Chapter II, we have already discussed similarity between 2-frames. For a given pair of frames E and F one can construct an operator A which intertwines between E,F. By that we mean $A \varepsilon B(X)$ such that $AE=FA$. The following lemma provides means for such a construction.

<u>Lemma 3.1</u>. *Let* $A,B \varepsilon B^n(X)$ *and* $C \varepsilon B(X)$. *Let* $P_{B,A}(C) : B^n(X) \times B^n(X) \times B(X) \longrightarrow B(X)$ *be given by*

$$(B,A,C) \longmapsto \sum_{j=1}^{n} B_j C A_j \quad .$$

Then P *is a continuous trilinear map. Moreover, if* E,F *are fixed elements in* $E^n(X)$, *then* $P_{F,E} : B(X) \longrightarrow B(X)$ *is a projector. Also* $S \varepsilon B(X)$ *is a similarity between* E,F *if and only if* $S \varepsilon G(X) \cap \text{Range } (P_{F,E})$.

<u>Proof</u>: The statement that $P_{B,A}(C)$ is continuous and trilinear is clear from the definition. Using equation (3.1) it can be proved that $P_{F,E}$ is a projector, and

for any $A \in \mathcal{B}(X)$, the operator $S = \sum_{j=1}^{n} F_j A E_j$ intertwines between E and F . (Just write $FS=SE$ in components.) If S is invertible then it will be a similarity. On the other hand, if S is a similarity between E and F , then certainly $S = \sum_{j=1}^{n} F_j S E_j$ and the lemma is proved. ///

Let

$$(3.2) \qquad C_j = (E_j + F_j - I)^2 \; , \qquad j=1,\dots,n \; ,$$

and suppose that each C_j is invertible. Define

$$(3.3) \qquad S = S(F,E) = P_{F,E}(I) = \sum_{j=1}^{n} F_j \, E_j \; .$$

S intertwines between E and F and the above condition implies that S is invertible, where $S^{-1} = \Sigma E_j C_j^{-1} F_j$. This follows from $F_j E_j F_j = C_j F_j = F_j C_j$ and $E_j F_j E_j =$ $= C_j E_j = E_j C_j$. The similarity given by (3.3) is not balanced since $S(F,E)^{-1} \neq S(E,F)$ unless all the C_j's are the identity operator. However, a balanced transformation can be constructed, under more restrictive conditions on the spectrum of the C_j's , as given by the following proposition.

<u>Proposition 3.2</u>. *Let* E,F *be two frames;* $E \neq F$. *Assume that the spectrum of each* C_j , $1 \leq j \leq n$, *does not separate* 0 *from* ∞ . *Then the operator*

$$(3.4)\quad U = U(F,E) = \sum_{j=1}^{n} F_j C_j^{-\frac{1}{2}} E_j = \left(\sum_{j=1}^{n} F_j E_j\right)\left(\sum_{j=1}^{n} E_j F_j E_j\right)^{-1/2}$$

is a similarity between F *and* E *which is also balanced. Further* $U(F,E)$ *and* $U(E,F)$ *have the block matrix representation* $[U(F,E)] = (U_{jk})_{j,k=1}^{n}$, $[U^{-1}(F,E)] = (U_{jk}^{-1})_{j,k=1}^{n}$ *with*

$$(3.5)\qquad U_{jk} = F_{k,jk}(F_{k,kk})^{-1/2}\ ,\quad U_{jk}^{-1} = (F_{j,jj})^{-1/2}\, F_{j,jk}$$

where

$$[F_i] = (F_{i,jk})_{j,k=1}^{n} \qquad i = 1,\dots,n\ .$$

Proof: Since 0 lies in the unbounded component of the complement of the spectrum of each C_j , every $C_j^{-\frac{1}{2}}$ is well defined (cf. [36, Chap.10]). Clearly $\sum_{j=1}^{n} F_j C_j^{-\frac{1}{2}} E_j$ intertwines between E and F . Since C_j commutes with F_j and E_j and $E_j F_j E_j = E_j C_j E_j$ by definition (3.2) of C_j , it follows that $\sum_{j=1}^{n} F_j C_j^{-\frac{1}{2}} E_j$ is invertible and its inverse is

$\sum_{j=1}^{n} E_j\, C_j^{-\frac{1}{2}} F_j$. Hence $U(F,E)$ is a balanced similarity.

To show the last equality in (3.4) we note that $T = \sum_{j=1}^{n} E_j F_j E_j$ is block diagonal (with respect to the frame E). The block diagonal elements of T are (see Chapter I),

$$T_{jj} = s_j^! \left(\sum_{j=1}^{n} E_j F_j E_j \right) i_j = s_j^! C_j i_j = s_j^! F_j i_j = F_{j,jj}$$

$$j = 1, \ldots, n$$

so that

$$T_{jj}^{-\frac{1}{2}} = s_j^! C_j^{-\frac{1}{2}} \, i_j = F_{j,jj}{}^{-\frac{1}{2}} \quad .$$

Hence

$$T^{-\frac{1}{2}} = \sum_{j=1}^{n} i_j \; T_{jj}^{-\frac{1}{2}} \; s_j^! = \sum_{j=1}^{n} E_j \; C_j^{-\frac{1}{2}} E_j \quad ,$$

which proves the last equality in (3.4). It remains to find the entries of the block matrix of $U(F,E)$ and $U(E,F)$. By (3.4)

$$U(F,E) = \left(\sum_{j=1}^{n} F_j E_j \right) T^{-\frac{1}{2}} \quad .$$

Thus we need only to find the entries of the block matrix of $S = \sum_{j=1}^{n} F_j E_j$. But $[S] = [S_{jk}]_{j,k=1}^{n}$ where $S_{jk} = s_j^! \, S \, i_k = F_{k,jk}$. Similarly since we can express $S^{-1} = T^{-1}\left(\sum_{j=1}^{n} E_j F_j \right)$ we can find the entries of the block matrix of S^{-1}, namely $S_{jk}^{-1} = (F_{j,jj})^{-1} \; F_{j,jk}$. Now (3.5) follows since $U(F,E) = ST^{-\frac{1}{2}}$ and $U(E,F) = T^{\frac{1}{2}} S^{-1}$. ///

Remarks 3.3.

1. The similarity given by (3.4) will be called the *balanced transformation*. If E and F are frames of orthoprojectors on a Hilbert space then $U(F,E)$ will be unitary and is the same as the one defined by C. Davis [7], where the underlying Hilbert space was finite dimensional. In the general Banach space setting, it was introduced by Z. Kovarik [27, Proposition 1].

2. If we assume that $\|E-F\|<1$, then the assumption of Proposition 3.2 will be satisfied since the definition of the norm in $B^n(X)$ implies that $\|E_j-F_j\|<1$ and hence $C_j^{-\frac{1}{2}}$ will be defined by the binomial series.

3. The assumption that $E \neq F$ implies that E does not commute with F, since if it does, then the identities (see [25] Remark c)

$$(E_j+F_j-I)(E_j-F_j) = (F_jE_j-E_jF_j) = (F_j-E_j)(E_j+F_j-I)$$

will imply that $E_j=F_j$, $j=1,\dots,n$.

Now we turn to the study of the direct rotation between two nearby frames E,F. This direct rotation comes up in connection with the study of $E^n(X)$ as a Banach manifold. For a given frame $E \varepsilon E^n(X)$, the similarity orbit of E is

$$(3.6) \qquad \mathcal{E}^n(X,E) = \{F \varepsilon E^n(X) : F \text{ is similar to } E\}$$

$$= \{VEV^{-1} : V \varepsilon G(X)\} .$$

The similarity orbits were studied in [27] and it was shown there that they have a differential structure as given in the next theorem.

Theorem 3.4. *Let* $E \varepsilon E^n(X)$ *. Then* $E^n(X,E)$ *defined in* (3.6) *is a regularly embedded split submanifold of* $B^n(X)$ *which can be modelled on the Banach space*

$$(3.7) \qquad M(E) = \{L \varepsilon B(X) : \sum_{j=1}^{n} E_j L E_j = 0\} .$$

Further, the tangent space can be identified with

$$(3.8) \qquad F(E) = \left\{ H \varepsilon B^n(X) : H_j = E_j H_j + H_j E_j \quad 1 \leq j \leq n, \ \sum_{j=1}^{n} H_j = 0 \right\} .$$

Proof: See [27]. ///

As a Banach manifold, $E^n(X,E)$ is equipped with a suitably defined connection. The geodesic lines through E in the direction of $H_0 \varepsilon F(E)$ are given explicitly by the arc

$$(3.9) \qquad t \longmapsto \exp(tL)\, E \, \exp(-tL)$$

where

$$(3.10)\quad L = \frac{1}{2} \sum_{j=1}^{n} (H_{0j}E_j - E_jH_{0j}) \in M(E) .$$

In principle, it is known that for two nearby frames there exists a geodesic arc connecting them [29, IV, §2]. This means that, given E and F sufficiently close, there exists a geodesic $t \longmapsto \alpha(t)$, $t\in[0,1]$ such that $\alpha(0)=E$, $\alpha(1)=F$ and

$$(3.11)\quad \alpha(t) = e^{tL}\, E\, e^{-tL} , \quad 0\le t\le 1 , \quad L\in M(E) .$$

In particular,

$$(3.12)\quad F=(\exp L)\ E \exp(-L) , \quad \sum_{j=1}^{n} E_j L E_j = 0 .$$

We call $R=\exp L$ the *direct rotation* between E and F. The name is justified as follows. In the case of Hilbert space, if we have frames of orthoprojectors then, under more restrictions on the similarity orbits, the similarity orbits will be Riemannian manifolds [27, Chapter 8]. Consequently the geodesics will be locally minimal and the arc length of (3.11) between $E=\alpha(0)$ and $F=\alpha(1)$ will be $\|L\|_{HS}$. Hence L resembles the oriented angle between the two frames, and indeed R is a direct rotation between E and F. We call L the generator of the geodesic between E and F. It is clear that the existence of L is guaranteed by the closeness of E and F. Equation (3.11) shows that L can be defined as the solution of the system of equations

$$F_j = (\exp L)\, E_j \exp(-L) \quad 1\le j\le n\,, \quad \sum_{j=1}^{n} E_j L E_j = 0\,.$$

In fact the above system is equivalent to a single operator equation

$$(3.13)\quad \exp L - P_{F,E}(\exp L) + P_{E,E}(L) = 0 \quad ;$$

for the equivalence we refer to [27, Chap. 9]. Adding up the above results for two frames E and F which are sufficiently close, the direct rotation R between E and F is $R=\exp L$ where L is the solution of the operator equation (3.13); moreover L can be obtained as the limit of the sequence

$$(3.14)\quad L_0=0\,,\ L_{m+1}=L_m-P_{E,E}(L_m)-\exp L_m+P_{F,E}(\exp L_m)\,,$$
$$m=0,1,\ldots$$

the transformation $L_m \to L_{m+1}$ being contractive.

The next proposition answers the question about the possibility that two frames can be connected by a frame valued path which is also a straight line segment.

Proposition 3.5. *Let* E,F *be two frames, then the line segment connecting* E *and* F *will be in* $E^n(X)$ *if and only if the* C_j*'s defined by* (3.2) *satisfy* $C_j=I$, $j=1,\ldots,n$.

Proof: From the definition of a frame it follows that E and F as elements in the algebra $B^n(X)$ are idempotents. If we define $C=(C_1,\ldots,C_n)$ then the above condition amounts to $C=I$ in $B^n(X)$. Since $(F+E-I)^2 = C$, we can apply Proposition 2.8 to show that the line segment connecting E and F lies in the set of idempotents in $B^n(X)$. This means that

$$F(t) = E+t(F-E)$$

is an idempotent for all $0\leq t\leq 1$. Component-wise this means That $F_j^2(t) = F_j(t)$ $0\leq t\leq 1$, $j=1,\ldots,n$. But since $F_j(t) = E_j+t(F_j-E_j)$, and E,F being frames implies that $\sum_{j=1}^{n} F_j(t) = I$ $0\leq t\leq 1$, for $F(t)$ to be a frame we need only to show that $F_j(t)F_k(t) = 0$ $j\neq k$. We do this by showing that $F(t)$ is similar to E . Note that

$$\begin{aligned} C_j(E_j,F_j(t)) &= I - (F_j(t)-E_j)^2 \\ &= I - t^2(F_j-E_j)^2 = I , \quad j=1,\ldots,n . \end{aligned}$$

Hence $S(F(t),E) = \sum_{j=1}^{n} F_j(t)E_j$ is invertible and its inverse is $\sum_{j=1}^{n} E_j\ F_j(t)$. Thus

$$SE_jS^{-1} = \left(\sum_{k=1}^{n} F_k(t)E_k\right) E_j \left(\sum_{k=1}^{n} E_kF_k(t)\right) = F_j(t),$$

$$j = 1,\dots,n\ .$$

But E being a frame implies $F_j(t)\ F_k(t)=0\quad j\neq k$. That is $t \longmapsto E+t(F-E)$ lies in $\mathcal{E}^n(X)$. The converse is easy to show, for if the line segment $t \longmapsto E(t) =$ $= E+t(F-E)$ lies in $\mathcal{E}^n(X)$, then $E_j(t) = E_j+t(F_j-E_j)$ $j=1,\dots,n$ is a projector. Thus

$$0 = E_j^2(t)-E_j(t) = (t^2-t)(F_j-E_j)^2\ ,$$

which implies that $C_j=I\quad j=1,\dots,n$. This completes the proof. ///

We conclude this section by pointing out that the above proposition shows that the similarity orbits given by (3.6) are ruled manifolds.

3.2. Local Characterization of the Balanced Transformation

In this section, we continue to study the balanced transformation introduced in Section 1. Recall that the balanced transformation $U(F,E)$ is

$$U(F,E) = \sum_{j=1}^{n} F_j \, C_j^{-1/2} E_j \, .$$

Note that as $F \longrightarrow E$, $U(F,E) \longrightarrow I$. Hence if F is sufficiently close to E , certainly U will be close to the identity. Throughout we will assume that E and F are sufficiently close frames. Define

(3.15) $\quad K = \log U(F,E)$,

so that $U(F,E) = \exp K$. The following lemma records some properties of the block matrix of $\exp K$.

Lemma 3.6. *For K defined by (3.15) we have*

(1) $\quad \exp K = \sum_{j=1}^{n} F_j (\exp K) E_j$,

(2) *$\sinh K$ is block off diagonal. That is*

$$\sum_{j=1}^{n} E_j \, (\sinh K) E_j = 0$$

(3) *the block diagonal part of $\exp K$ is equal to the block diagonal part of $\cosh K$;*

Proof: Lemma 3.1 shows that $\exp K \in \mathrm{Range}(P_{F,E})$ whence (1) follows. Since $\sinh K = \frac{1}{2}(\exp K - \exp(-K)) = \frac{1}{2}(U(F,E) - U^{-1}(F,E))$, we can apply equations (3.5) with $j=k$

to find that the diagonal blocks of $U(F,E)$ and $U^{-1}(F,E)$ are equal to $F_{j,jj}^{1/2}$. Hence $\sinh K$ is block off diagonal. As $\exp K = \sinh K + \cosh K$, the operator $\sinh K$ being block off diagonal implies (3). The lemma is proved. ///

The next theorem shows that K as defined by equation (3.15) may be equivalently defined as the solution of the operator equation (3.16) given below.

<u>Theorem 3.7</u>. *If* $\|E-F\|$ *is sufficiently small then* $K=\log U(F,E)$ *is the unique solution of*

$$(3.16) \quad \exp K - P_{F,E}(\exp K) + P_{E,E}(\sinh K) = 0 ,$$

with the property that $K \longrightarrow 0$ *as* $F \longrightarrow E$.

<u>Proof</u>: First we show that $K=\log U(F,E)$ is a solution of (3.16). This follows directly from Lemma 3.6. But since $U(F,E) = \left(\sum_{j=1}^{n} F_j E_j\right)\left(\sum_{j=1}^{n} E_j F_j E_j\right)^{-1/2}$ implies that $U \longrightarrow I$ as $F \longrightarrow E$, we have $K \longrightarrow 0$. To show the local uniqueness of K , we appeal to the implicit function theorem. For that, let E be a fixed frame and define

$$\psi : B^n(X) \times B(X) \longrightarrow B(X)$$

by

$$\psi(F,K) = \exp K - P_{F,E}(\exp K) + P_{E,E}(\sinh K) .$$

It follows from Lemma 3.1 that the map $\mathcal{P}_{\cdot,E}(\cdot)$ is a smooth map. Hence ψ will be a smooth map. Also

$$\psi(E,0) = 0 .$$

The partial derivative of ψ with respect to the second variable at $(E,0)$ is

$$D_2\psi(E,0) = I \varepsilon \mathcal{B}(\mathcal{B}(X)) .$$

This can be justified as follows

$$\psi(E,H) - \psi(E,0) = \exp H - \mathcal{P}_{E,E}(\exp H) + \mathcal{P}_{E,E}(\sinh H) .$$

Since $\exp H = \sum_{n=0}^{\infty} \frac{H^n}{n!}$ and $\sinh H = \sum_{n=0}^{\infty} \frac{H^{2n+1}}{(2n+1)!}$, one has $\exp H = I+H+O(\|H\|^2)$ and $\sinh H = H+O(\|H\|^3)$, so that

$$\psi(E,H)-\psi(E,0) = \exp H-\mathcal{P}_{E,E}(I+H+O(\|H\|^2))+\mathcal{P}_{E,E}(H+O(\|H\|^3))$$

$$= H+O(\|H\|^2) .$$

Thus $D_2\psi(E,0)\,H = H$ and $D_2\psi(E,0) = I \varepsilon \mathcal{B}(\mathcal{B}(X))$. Then the conditions of the implicit function theorem are satisfied. Hence (Theorem 1.1) there exists a neighbourhood W_0 of E in $\mathcal{B}^n(X)$ and a smooth function $g_0 : W_0 \longrightarrow \mathcal{B}(X)$ such that $\psi(F,K)=0$ if and only if $K=g_0(F)$ for $F \varepsilon W_0$. So if we take $W = W_0 \cap \mathcal{E}^n(X)$ and $g = g_0|W : W \longrightarrow \mathcal{B}(X)$, for such a neighbourhood, $K=g(F)$ is the unique solution of

(3.16) with the property that $K \longrightarrow 0$ as $F \longrightarrow E$. The conclusion of the theorem follows. ///

The previous theorem helps to characterize the balanced transformation locally as follows.

Theorem 3.8. *Let E,F be two frames and suppose that*

(1) V is a similarity between F,E,

(2) $\|E-F\|$ is sufficiently small,

(3) V is close to the identity.

Then V will be the balanced transformation between E and F if and only if it satisfies

$$(3.17) \quad \sum_{j=1}^{n} E_j \, (V-V^{-1}) \, E_j = 0 \, .$$

Proof: It follows from Lemma 3.6 that the balanced transformation satisfies (3.17). On the other hand, suppose V satisfies (3.17). Then condition (3) implies that $K=\log V$ is well defined, $V=\exp K$. V being a similarity implies $\exp K = \sum_{j=1}^{n} F_j (\exp K) E_j$. Now equation (3.17) implies that K satisfies equation (3.16). Hence by Theorem 3.7 $K=\log U(F,E)$ which reads that $V=U(F,E)$. This proves the theorem. ///

In case of a 2-frame the above theorem takes the following form.

Corollary 3.9. *Let* $E=(E_1,E_2)$, $F=(F_1,F_2)$ *be 2-frames. Then under the same assumptions as in Theorem 3.8,*

$$V^2 = \sum_{j=1}^{2} F_j E_j \qquad V \in A(E,F)$$

if and only if $\sum_{j=1}^{2} E_j (V-V^{-1}) E_j = 0$.

This corollary can be also deduced from the semi-global characterization of the direct rotation given in Theorem 2.16. This is because in case of 2-frames the direct rotation and the balanced transformation coincide.

As we mentioned in the introduction, the balanced transformation $U(F,E)$, for sufficiently close frames F and E, defines a path connecting F and E in $E^n(X)$ as follows.

(3.18) $t \longmapsto (\exp tK)\, E \exp(-tK)$, $0 \leq t \leq 1$

where $K = \log U(F,E)$. A glance at the geodesic between F and E , given by equation (3.11), shows that the path (3.18) will be a geodesic provided $K = \log U(F,E) \in M(E)$, where $M(E)$ is defined in (3.7). This always happens in case of a 2-frame (See Theorem 2.6). Still, it might happen in case of n-frame $n>2$ as will be shown in example 3.10 below. At the end of the next section, we give an example for the other case when the two paths do not coincide.

<u>Example 3.10</u>. Let $E_j = e_j e_j^T$ be the coordinate projectors on $\mathbb{C}^3$ $(j=1,2,3)$, and let $Q=e^{\varepsilon H}$ where

$$H = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} .$$

Clearly $H \varepsilon M(E)$. H has eigenvalues $0,1,-1$. Let $f(\lambda)=e^{\varepsilon\lambda}$ then $\exp(\varepsilon H) = \sum_{k=1}^{3} f(\lambda_k)\, L_k(H)$, where

$$L_k(H) = \prod_{\substack{i=1 \\ i\neq k}}^{3} \frac{(H-\lambda_i I)}{\lambda_k-\lambda_i} .$$

Then

$$L_1(H) = I-H^2 \ , \ L_2(H) = \tfrac{1}{2}(H^2+H) \ , \ L_3(H) = \tfrac{1}{2}(H^2-H) \ ,$$

$$\exp(\varepsilon H) = \begin{bmatrix} 1 & 1 & -1 \\ -1 & -1 & 1 \\ -1 & -1 & 1 \end{bmatrix} + e^{\varepsilon} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} + e^{-\varepsilon} \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} .$$

$$\exp(\varepsilon H) = \begin{bmatrix} 1 & 1-e^{-\varepsilon} & -1+e^{-\varepsilon} \\ -1+e^{\varepsilon} & -1+e^{\varepsilon}+e^{-\varepsilon} & 1-e^{-\varepsilon} \\ -1+e^{\varepsilon} & -1+e^{\varepsilon} & 1 \end{bmatrix},$$

$$\exp(-\varepsilon H) = \begin{bmatrix} 1 & 1-e^{\varepsilon} & -1+e^{\varepsilon} \\ -1+e^{-\varepsilon} & -1+e^{-\varepsilon}+e^{\varepsilon} & 1-e^{\varepsilon} \\ -1+e^{-\varepsilon} & -1+e^{-\varepsilon} & 1 \end{bmatrix}.$$

If we let $F=QEQ^{-1}$ with $Q=\exp(\varepsilon H)$, then F will be a frame and Q will be the direct rotation between E and F. Now we compute $U(F,E) = \left(\sum_{j=1}^{n} F_jE_j\right)\left(\sum_{j=1}^{n} E_jF_jE_j\right)^{-1/2}$. Direct calculation gives

$$\sum_{j=1}^{n} F_jE_j = Q \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1+e^{\varepsilon}+e^{-\varepsilon} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \sum_{j=1}^{n} E_jF_jE_j = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (-1+e^{\varepsilon}+e^{-\varepsilon})^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$U(F,E) = Q = \exp(\varepsilon H)$.

In fact for the above particular example, we can even show more. Namely, for any $H \in M(E)$ with eigenvalues $\{-\lambda,0,\lambda\}$ the balanced transformation between E and F, where $F = \exp(\varepsilon H)\, E \exp(-\varepsilon H)$ will be $\exp(\varepsilon H)$. This is because in this case we have

$$\exp(\varepsilon H) = P_1 + e^{\lambda\varepsilon} P_2 + e^{-\lambda\varepsilon} P_3$$

where

$$P_1 = \frac{1}{\lambda^2}(\lambda^2 - H^2), \quad P_2 = \frac{1}{2\lambda^2}(\lambda H + H^2) \text{ and } P_3 = \frac{1}{2\lambda^2}(-\lambda H + H^2) .$$

Now

$$\exp(\varepsilon H) = \frac{1}{\lambda^2}(\lambda^2 - H^2) + \frac{e^{\lambda\varepsilon}}{2\lambda^2}(\lambda H + H^2) + \frac{e^{-\lambda\varepsilon}}{2\lambda^2}(-\lambda H + H^2) .$$

From the characterization of the balanced transformation (Theorem 3.8), in order that $\exp(\varepsilon H) = U(F,E)$ we have to have the diagonal elements of $\exp(\varepsilon H)$ the same as $\exp(-\varepsilon H)$. For that let $[H^2]_E = (H_{ij})$, then

$$(\exp(\varepsilon H))_{jj} = 1 - \frac{H_{jj}}{\lambda^2} + \frac{e^{\lambda\varepsilon}}{2\lambda^2} H_{jj} + \frac{e^{-\lambda\varepsilon}}{2\lambda^2} H_{jj}$$

$$j = 1,2,3$$

$$(\exp(-\varepsilon H))_{jj} = 1 - \frac{H_{jj}}{2\lambda^2} + \frac{e^{-\lambda\varepsilon}}{2\lambda^2} H_{jj} + \frac{e^{\lambda\varepsilon}}{2\lambda^2} H_{jj}$$

$$j = 1,2,3$$

(because $H \varepsilon M(E)$ implies the diagonal elements are zeros). We remark that this example does not fall into the list of instances given in [27, Chapter 10, (i) - (iii)].

3.3. Closeness Between Direct Rotation and Balanced Transformation

In this section we investigate how close the balanced transformation is to the direct rotation between two given frames E and F. Both similarities offer interpolating paths between E and F in $E^n(X)$. The direct rotation $R=\exp L$ gives the path (3.11) where the generator of the path satisfies the operator equation (3.13), i.e.

$$\exp L - P_{F,E}(\exp L) + P_{E,E}(L) = 0 \ ,$$

and the balanced transformation $U(F,E) = \left(\sum_{j=1}^{n} F_j E_j\right) \left(\sum_{j=1}^{n} E_j F_j E_j\right)^{-1/2}$ gives rise to the path (3.18) where K is the solution of the operator equation (3.16), i.e.

$$\exp K - P_{F,E}(\exp K) + P_{E,E}(\sinh K) = 0 \ .$$

Note that two similarities between F and E are related by an invertible E-diagonal operator, that is S,T are similarities between F,E if and only if $S=TD$ where D is an invertible E-diagonal operator. We use this fact to transform the equations (3.16) and (3.13) into equivalent forms which are more suitable, in our opinion, for finding an estimate for $\|R-U(F,E)\|$. Now since F can be represented as $F=QEQ^{-1}$, we have $\exp L=QD$ and $\exp K=QD_0$, where D,D_0 are invetible E-diagonal operators.

Lemma 3.11. *Under the transformation $F=QEQ^{-1}$, equations (3.13) and (3.16) have the equivalent forms*

(3.13)' $\quad D - P_{E,E}(D) + P_{E,E}(\log(QD)) = 0$

(3.16)' $\quad D_0 - P_{E,E}(D_0) + P_{E,E}(\sinh \log(QD_0)) = 0$.

<u>Proof</u>: To show that (3.13)' is equivalent to (3.13), we note first that L being a solution of (3.13) implies

$$\exp L = P_{F,E}(\exp L) = \sum_{j=1}^{n} F_j(\exp L)E_j = Q \sum_{j=1}^{n} E_j Q^{-1}(\exp L)E_j .$$

Let $D = Q^{-1}\exp L$; then we have $D = P_{E,E}(D)$. Next $P_{E,E}(L)=0$ implies $P_{E,E}(\log(QD))=0$, hence D satisfies (3.13)'. On the other hand, if D satisfies (3.13)', then $P_{E,E}$ being a projector implies $P_{E,E}(\log(QD))=0$, and $P_{E,E}(D)=D$, so if we put $L = \log(QD)$, then L satisfies (3.13). A similar argument shows the equivalence between (3.16) and (3.16)'. ///

Next we use a method of solving operator equations of the form $G(A)=0$ where G is a function defined on an open subset of $\mathcal{B}(X)$. This method is known, in case of real equations, as Newton's method. Its generalizations and modifications were investigated in detail in [23, XVIII]. We will be interested in the special case when the operator equation has the form

$$G(A) = \pi(A) + R(A) = 0 ,$$

where the solution A_0 of the approximate equation $\pi(A)=0$ is known. Then on the basis of certain conditions we can draw a conclusion about $\|A^*-A_0\|$ where A^* is the solution of $G(A)=0$. Our intention is to use this idea to get an estimate on $\|\exp L - U(F,E)\|$. This will be given in the next theorem.

Theorem 3.12. *Consider the operator equations (with a fixed $Q \varepsilon G(X)$)*

$$(3.19) \quad G(Q,D) = D - P_{E,E}(D) + P_{E,E}(\log(QD)) = 0$$

$$(3.20) \quad \pi(Q,D) = D - P_{E,E}(D) + P_{E,E}(\sinh \log(QD)) = 0 .$$

Then (Q,D^*) is a solution of (3.19) and (Q,D_0) is a solution of (3.20), where $D^* = Q^{-1}(\exp L)$, $L \varepsilon M(E)$ is given by (3.13), and $D_0 = Q^{-1}(\exp K)$ with $K = \log U(F,E)$. Furthermore

$$G(Q,D) = \pi(Q,D) + R(Q,D)$$

where

$$(3.21) \quad R(Q,D) = P_{E,E}(\log(QD) - \sinh \log(QD))$$

and

$$(3.22) \quad \|\exp L - U(F,E)\| = O(\|E-F\|^3) .$$

Proof: The statements that (Q,D^*) and (Q,D_0) are solutions of (3.19) and (3.20) follow directly from Lemma 3.11. Also, (3.21) follows from (3.19) and (3.20). In order to prove (3.22) we need to find estimates on $R(Q,D)$, $\pi(Q,D)$

and their derivatives. We start by observing that from our assumption that E and F are sufficiently close, we know that QD is close to the identity if Q is. From (3.21) we have

$$R(Q,D) = P_{E,E}(\log(QD) - \sinh \log(QD)) .$$

Let

$$\tilde{R}(Q,D) = \log(QD) - \sinh \log(QD)$$

We use the functional calculus approach to bound $\tilde{R}(Q,D)$, that is

$$\tilde{R}(Q,D) = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda)\, (\lambda - QD)^{-1}\, d\lambda ,$$

where $\Gamma = \{\lambda : |\lambda - 1| = \rho < 1\}$ and $f(\lambda) = \log\lambda - \sinh(\log\lambda)$. And as we observed above we can assume that $\|I-QD\| < \delta$, $\delta = \frac{1}{2}\rho$. By this we can bound $\|(\lambda-QD)^{-1}\|$ for $\lambda \varepsilon \Gamma$ as follows. For $\lambda \varepsilon \Gamma$, $\lambda = 1+\rho e^{i\theta}$ and

$$(\lambda - QD)^{-1} = [\rho e^{i\theta} - (QD-I)]^{-1} =$$

$$= \rho^{-1} e^{-i\theta} \sum_{k=0}^{\infty} \rho^{-k}\, e^{-ik\theta} (QD-I)^{k} .$$

So $\|(\lambda-QD)^{-1}\| \leq M = 2/\rho$. Since $f(\lambda) = \log\lambda - \sinh(\log\lambda)$ and $\lambda \varepsilon \Gamma$, we have

$$f(\lambda) = \log\lambda - \sum_{k=0}^{\infty} \frac{(\log\lambda)^{2k+1}}{(2k+1)!} ,$$

and from

$$\log\lambda = - \sum_{j=1}^{\infty} \frac{(1-\lambda)^j}{j}$$

we obtain

$$f(\lambda) = - \frac{(1-\lambda)^3}{3!} + O((1-\lambda)^5) ,$$

so that for $\lambda\varepsilon\Gamma$ $|f(\lambda)|<\rho^3/5$. The above integral expression of $\tilde{R}(Q,D)$ gives

$$\|\tilde{R}(Q,D)\| \le \frac{1}{2\pi} \max_{\lambda\varepsilon\rho} |f(\lambda)| \; M\,(2\pi\rho)$$
$$\le \frac{2}{5}\,\rho^3 \; ;$$

Hence

(3.23) $\|R(Q,D_0)\| \le \frac{2}{5}\,\rho^3 \;\; \|P_{E,E}\| = \zeta.$

We will denote by $\tilde{R}'(Q,D)$ the derivative of $\tilde{R}(Q,D)$ with respect to D . The derivative of $\tilde{R}(Q,D)$ is given explicitly by (See [36, 10.35])

$$\tilde{R}'(Q,D)H = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda)\; (\lambda-QD)^{-1}\; QH(\lambda-QD)^{-1}\, d\lambda .$$

So

$$\|\tilde{R}'(Q,D)H\| \le \frac{4}{5}\; \rho^2\|Q\|\;\|H\| .$$

Thus

$$(3.24)\quad \|\tilde{R}'(Q,D_0)\| \le \frac{4}{5}\rho^2 \|Q\| \|P_{E,E}\| = \alpha .$$

We will find a bound on the second derivative, using its explicit expression. From

$$\tilde{R}'(Q,D)H = \frac{1}{2\pi i}\int_\Gamma f(\lambda)\,(\lambda-QD)^{-1}\,QH\,(\lambda-QD)^{-1}\,d\lambda ,$$

we obtain

$$\tilde{R}'(Q,D+K)H-\tilde{R}'(Q,D)H = \frac{1}{2\pi i}\int_\Gamma f(\lambda)\left\{(\lambda-QD)^{-1}[QK(\lambda-QD)^{-1}\,QH\, + \right.$$

$$\left. + \, QH(\lambda-QD)^{-1}QK](\lambda-QD)^{-1} + O\left(\|K\|^2\right)\right\} d\lambda .$$

It follows that

$$\|\tilde{R}''(Q,D)HK\| \le \frac{2}{5}\rho^4 M^3 \|Q\|^2 \|K\| \|H\|$$

and

$$(3.25)\quad \|\tilde{R}''(Q,D)\| \le \frac{16}{5}\rho\|P_{E,E}\| \|Q\|^2 = p .$$

Next we bound $\pi''(Q,D)$. Since

$$\pi(Q,D) = D - P_{E,E}(D) + P_{E,E}\sinh(\log(QD)) ,$$

a similar argument as above gives

$$\|\tilde{\pi}''(Q,D)HK\| \le \|Q\|^2 \|H\| \|K\| M^3 \max_{\lambda\varepsilon\Gamma} |g(\lambda)|$$

where

$$g(z) = \frac{1}{2}(z-\frac{1}{z}) = -[(1-z)+O(1-z)^2] .$$

Hence

$$(3.26) \quad \|\pi''(Q,D)\| \le \frac{8}{\rho} \|Q\|^2 \|P_{E,E}\| = k .$$

Clearly α in equation (3.24) can be made less than 1 by taking ρ small enough. Let

$$h = \frac{\zeta(k+p)}{(1-\alpha)^2} .$$

Here ζ,k,p are defined in (3.23), (3.25), (3.26) respectively. Now if $\alpha<1$, $h<1/2$ we can carry out the successive approximations on

$$S(D) = D - G(Q,D)$$

with the majorizing function

$$\phi(t) = (\frac{k+p}{2})\, t^2+\alpha t+\zeta ,$$

which indeed majorizes S (see [23], XVIII). Also $h<1/2$ implies that $\phi(t)$ has real fixed points and

$$r_0 = \frac{1 - \sqrt{1-2h}}{h} \; \frac{\zeta}{1-\alpha}$$

will be the unique root of $t=\phi(t)$ in $[0,\delta]$.

Now appealing to Theorem 1 [23, XVIII, § 2], we see that the successive approximations starting at D_0 will converge to D^* and $\|D^*-D_0\| \leq r_0 = O(\rho^3)$. Hence

$$\|\exp L - U(F,E)\| \leq O(\|F-E\|^3) .$$

This completes the proof of the theorem. ///

Remark 3.13. The estimate (3.22) justifies in the finite dimensional problem the use of $U(F,E)$ instead of $\exp L$. The reason is that $U(F,E)$ has an algebraic expression while L is obtained by an iterative process. If an iterative process is set up for computing L , a good initial approximation would be K .

We now give an example to show that the order in (3.22) cannot be improved in general.

Example 3.14.

Let $E_j = e_j e_j^T$ be the coordinate projectors in C^3. Let

$$H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 7 & 0 \end{bmatrix} .$$

The eigenvalues are $1, 2, -3$ and

$$P_1 = \frac{1}{4}\begin{bmatrix} 6 & -1 & -1 \\ 6 & -1 & -1 \\ 6 & -1 & -1 \end{bmatrix}, \quad P_2 = \frac{1}{5}\begin{bmatrix} -3 & 2 & 1 \\ -6 & 4 & 2 \\ -12 & 8 & 4 \end{bmatrix},$$

$$P_3 = \frac{1}{20}\begin{bmatrix} 2 & -3 & 1 \\ -6 & 9 & -3 \\ 18 & -27 & 9 \end{bmatrix},$$

$$\exp\varepsilon H = \frac{1}{4}e^{\varepsilon}\begin{bmatrix} 6 & -1 & -1 \\ 6 & -1 & -1 \\ 6 & -1 & -1 \end{bmatrix} + \frac{1}{5}e^{2\varepsilon}\begin{bmatrix} -3 & 2 & 1 \\ -6 & 4 & 2 \\ -12 & 8 & 4 \end{bmatrix} + $$

$$+ \frac{1}{20}e^{-3\varepsilon}\begin{bmatrix} 2 & -3 & 1 \\ -6 & 9 & -3 \\ 18 & -27 & 9 \end{bmatrix}$$

$F = \exp(\varepsilon H)E\exp(-\varepsilon H)$

$Q = \exp(\varepsilon H)$

$$= \frac{1}{20}\begin{bmatrix} 30e^{\varepsilon}-12e^{2\varepsilon}+2e^{-3\varepsilon} & -5e^{\varepsilon}+8e^{2\varepsilon}-3e^{-3\varepsilon} & -5e^{\varepsilon}+4e^{2\varepsilon}+e^{-3\varepsilon} \\ 30e^{\varepsilon}-24e^{2\varepsilon}-6e^{-3\varepsilon} & -5e^{\varepsilon}+16e^{2\varepsilon}+9e^{-3\varepsilon} & -5e^{\varepsilon}+8e^{2\varepsilon}-12e^{-3\varepsilon} \\ 30e^{\varepsilon}-48e^{2\varepsilon}+18e^{-3\varepsilon} & -5e^{\varepsilon}+32e^{2\varepsilon}-27e^{-3\varepsilon} & -5e^{\varepsilon}+16e^{2\varepsilon}+9e^{-3\varepsilon} \end{bmatrix}.$$

Similarly

$$Q^{-1} = \exp(-\varepsilon H)$$

$$= \frac{1}{20}\begin{bmatrix} 30e^{-\varepsilon}-12e^{-2\varepsilon}+2e^{3\varepsilon} & -5e^{-\varepsilon}+8e^{-2\varepsilon}-3e^{3\varepsilon} & -5e^{-\varepsilon}+4e^{-2\varepsilon}+e^{3\varepsilon} \\ 30e^{-\varepsilon}-24e^{-2\varepsilon}-6e^{3\varepsilon} & -5e^{-\varepsilon}+16e^{-2\varepsilon}+9e^{3\varepsilon} & -5e^{-\varepsilon}+8e^{-2\varepsilon}-3e^{3\varepsilon} \\ 30e^{-\varepsilon}-48e^{-2\varepsilon}+18e^{3\varepsilon} & -5e^{-\varepsilon}+32e^{-2\varepsilon}-27e^{3\varepsilon} & -5e^{-\varepsilon}+16e^{-2\varepsilon}+9e^{3\varepsilon} \end{bmatrix}.$$

Certainly $Q \neq U(F,E)$ since the diagonal elements of Q and Q^{-1} are not equal.

$$U(F,E) = \sum_{j=1}^{3} F_jE_j\left(\sum_{j=1}^{3} E_jF_jE_j\right)^{-1/2}$$

$$F_1E_1 = Q \text{ diag } \left\{\frac{3}{2}e^{-\varepsilon} - \frac{3}{5}e^{-2\varepsilon} + \frac{1}{10}e^{3\varepsilon}, 0, 0\right\}$$

$$F_2E_2 = Q \text{ diag } \left\{0, -\frac{1}{4}e^{-\varepsilon} + \frac{4}{5}e^{-2\varepsilon} + \frac{9}{20}e^{3\varepsilon}, 0\right\},$$

$$F_3E_3 = Q \text{ diag } \left\{0, 0, -\frac{1}{4}e^{-\varepsilon} + \frac{4}{5}e^{-2\varepsilon} + \frac{9}{20}e^{3\varepsilon}\right\},$$

$$\sum_{j=1}^{3} F_jE_j = Q \text{ diag } Q^{-1}$$

$$= Q \text{ diag } \left\{\frac{3}{2}e^{-\varepsilon}-\frac{3}{5}e^{-2\varepsilon}+\frac{1}{10}e^{3\varepsilon}, -\frac{1}{4}e^{-\varepsilon}+\frac{4}{5}e^{-2\varepsilon}+\frac{9}{20}e^{3\varepsilon}, -\frac{1}{4}e^{-\varepsilon}+\frac{4}{5}e^{-2\varepsilon}+\frac{9}{20}e^{3\varepsilon}\right\},$$

and

$$\sum_{j=1}^{3} E_jF_jE_j = \text{diag } Q \text{ diag } Q^{-1}$$

$$U(F,E) = Q \text{ diag } Q^{-1} (\text{diag } Q \text{ diag } Q^{-1})^{-1/2}$$

$$= Q\sqrt{(\text{diag } Q^{-1})(\text{diag } Q)^{-1}} \quad .$$

So that explicitly

$$U(F,E) = Q \text{ diag } \{A,B,B\} \; .$$

The first diagonal element A is

$$\sqrt{\frac{10+60\varepsilon^3+O(\varepsilon^4)}{10-60\varepsilon^3+O(\varepsilon^4)}} = 1 + 6\varepsilon^3 + O(\varepsilon^4) \; .$$

The second and third diagonal elements B are

$$\sqrt{\frac{20+70\varepsilon^2+20\varepsilon^3+O(\varepsilon^4)}{20+70\varepsilon^2-20\varepsilon^3+O(\varepsilon^4)}} = 1 + \varepsilon^3 + O(\varepsilon^4) \; .$$

Hence

$$U(F,E) = Q\,(1+O(\varepsilon^3)) \; ,$$

and

$$\|U(F,E)-Q\| \le \|Q\|\, O(\varepsilon^3) \; .$$

But $\|F-E\| = O(\varepsilon)$, hence

$$\|U(F,E)-Q\| = O(\|F-E\|^3) \; ,$$

and the order cannot be improved to $o(\|F-E\|^3)$.

CHAPTER IV

PERTURBATION OF INVARIANT SUBSPACES

Let A be a selfadjoint operator on a Hilbert space $\mathcal{H}$ and H be a "small" selfadjoint perturbation of A. Then a question arises concerning how much an invariant subspace of A will change under the perturbation H. This question is related to the geometry of a pair of subspaces and an answer to this question must be based on such geometry. This problem was investigated in [7], [8] and [9]. Taking the angle operator Θ mentioned in Section 1, Chapter II as a measure of the difference between the subspaces, it was shown in the above mentioned works that, bounds on trigonometric functions of the operator angle Θ can be obtained from the gap between parts of the spectrum of A or A+H, and the perturbation H. In the present chapter, we study the same problem in the Banach space setting.

In Section 1 we give the formulation of the problem and we show how to associate with A and A+H spectral projector frames E and F respectively. The result depends on two main ingredients. The first is a measure of the separation of the spectra of two operators. This will be developed in Section 2. The second is a measure of how far the two spectral projector frames E, F are. As a measure of that we take the generator L of the geodesic between E and F which we called the oriented angle. Its use as a measure is already justified in Chapter III. In Section 3 we get a bound for L, obtained from the separation of parts of the spectrum of A and from the perturbation H. Finally in Section 4 we consider the same problem, except that the frames E and F arise from selfadjoint operators A, $A+H$ in a Hilbert space $\mathcal{H}$. In this case E and F will be orthoprojector frames. As it was mentioned before; a detailed study is already known in the case of 2-frames. We study the case of n-frames where $n \geq 2$ and we give a result of a different nature than the one given in Section 3. In fact our result generalizes some of those known results for 2-frames [9].

4.1. Spectral Projector Frame of a Linear Operator

Let $A \in \mathcal{B}(X)$ be a bounded linear operator on a Banach space such that the spectrum of A, $\sigma(A)$, is separated into several non-empty parts $\sigma_1(A), \ldots, \sigma_n(A)$. Each $\sigma_j(A)$, $1 \leq j \leq n$, is a spectral set (i.e. open and closed in $\sigma(A)$) and is enclosed by a simple Jordan curve Γ_j (or a finite collection of simple closed Jordan curves). It is known [11, VII, 3.20] that the spectral sets give rise to spectral projectors E_j $1 \leq j \leq n$, where E_j is defined by

$$(4.1) \quad E_j = \frac{1}{2\pi i} \int_{\Gamma_j} (\lambda I - A)^{-1}\, d\lambda \ , \quad 1 \le j \le n \ ,$$

each associated with the spectral set σ_j , and they satisfy

$$(4.2) \quad E_j \neq 0 \ , \ \sum_{j=1}^{n} E_j = I \ , \ E_j E_k = \delta_{jk} E_k \quad (j,k=1,\ldots,n) \ .$$

A commutes with each E_j , so that A is decomposed according to the decomposition

$$X = X_1 \oplus X_2 \oplus \ \ldots \ \oplus X_n \ ,$$

where $X_j = E_j X$. The part $A|_{X_j}$ of A in X_j has spectrum $\sigma_j(A)$ and $A|_{X_j} \varepsilon \mathcal{B}(X_j)$. Equation (4.2) indicates that E is a frame. We call it the spectral projector frame associated with A . The commutativity of A with all the E_j's mean that A has a block diagonal matrix with respect to the frame E.

Now we discuss how the spectrum of A changes when A undergoes a small perturbation. In the finite dimensional case, the spectrum consists only of eigenvalues, and it is known that the eigenvalues of an operator A depend on A continuously [22, II, §5-8]. Even in general Banach space, the spectrum $\sigma(A)$ changes continuously with A if the perturbation commutes with A [11, VII. 6.10]. There are other kinds of restricted continuity of the spectrum. For example, $\sigma(A)$ changes continuously with A if A

varies over the set of selfadjoint operators on a Hilbert space [22, VIII, §1.2]. However, this is not true for more general perturbations, nevertheless the spectrum is still an upper semicontinuous function from $B(X)$ into the compact subsets of the complex plane [2, I, §6]. By this we mean for each $A \in B(X)$ and each neighbourhood V of $\sigma(A)$, there exists a neighbourhood U of A in $B(X)$ such that

$$\sigma(B) \subset V \qquad (B \in U) .$$

Roughly speaking, the upper semicontinuity says that $\sigma(A)$ does not expand suddenly when A is changed continuously, but it may very well "shrink" suddenly [22, IV, §3, example 3.8].

We showed above that $\sigma(A)$ is upper semicontinuous. In fact a finer result can be shown. Namely, each separated part of $\sigma(A)$ is upper semicontinuous. This will be clear from the next theorem, whose proof can be found in [22, IV, §3.4].

Theorem 4.1. *Let $A \in B(X)$ have spectrum separated as above. Let*

$$X = X_1(A) \oplus \ldots \oplus X_n(A)$$

be the associated decomposition of X. Then there exists a $\delta > 0$ depending on A and the Γ_j's with the following *properties. For any* $H \in B(X)$ *with* $\|H\| < \delta$, $A+H$ *has spectrum likewise separated into* n *non-empty parts* $\sigma_1(A+H),\ldots,\sigma_n(A+H)$ *and each* $\sigma_j(A+H)$ *is included in* Γ_j. *In the associated decomposition*

$$X = X_1(A+H) \oplus \ldots \oplus X_n(A+H) \ ,$$

$X_1(A+H),\ldots,X_n(A+H)$ *are isomorphic respectively with* $X_1(A),\ldots,X_n(A)$.

Under the assumption that the perturbation H of A is sufficiently small, the spectrum of $A+H$ will be separated into $\sigma_1(A+H),\ldots,\sigma_n(A+H)$. In a similar way to what we did with A we can define the spectral projectors F_j associated with the spectral sets $\sigma_j(A+H)$ $j=1,\ldots,n$. Namely

$$F_j = \frac{1}{2\pi i} \int_{\Gamma_j} (\lambda I-A-H)^{-1} \, d\lambda \ , \ 1 \le j \le n \ .$$

By the same argument as before F_j will satisfy (4.2). Let $F=(F_1,\ldots,F_n)$, then F will be the spectral projector frame associated with $A+H$. We pose the problem as follows: given A, let H be a small perturbation of A. Associate with A and $A+H$ the spectral projector frames E and F respectively. It is known that to draw a conclusion about how far F is from E, we have to put restrictions on the separation of the parts of the spectrum of A. To summarize, we want to bound a measure of the difference between E and F by a function of the perturbation H and the gap between the parts of the spectrum of A.

4.2. The Separation of Two Operators

Let X, Y be Banach spaces. Let $B \in \mathcal{B}(X)$ and $C \in \mathcal{B}(Y)$. Then B and C define an operator $T \in \mathcal{B}(\mathcal{B}(X,Y))$ by

$$(4.3) \qquad T(P) = PB - CP, \quad P \in \mathcal{B}(X,Y).$$

The above operator equation was studied by Lumer and Rosenblum [7] (also in [32] and [35]) when P, B and C belong to a Banach algebra. Their techniques, which apply here, show that

$$(4.4) \qquad \sigma(T) = \sigma(B) - \sigma(C) = \{\beta - \gamma : \beta \in \sigma(B), \gamma \in \sigma(C)\}$$

We use the above result to define a measure for the separation of the spectra of B and C. Since the distance between the spectra may change violently with small perturbations, it cannot be taken as a reasonable measure of the separation of the spectra. However, for self-adjoint operators on Hilbert spaces, the distance between the spectra can be used as a measure due to the continuity of the spectra as mentioned in Section 4.1.

<u>Definition 4.2</u>. *Let $B \in \mathcal{B}(X)$ and $C \in \mathcal{B}(Y)$. Then the separa-*

tion between B and C , denoted Sep(B,C) is

$$Sep(B,C) = \begin{cases} \|T^{-1}\|^{-1} & 0 \notin \sigma(T) \\ 0 & 0 \in \sigma(T) \end{cases},$$

where T is defined by (4.3).

Theorem 4.3. *The separation of B and C satisfies the inequality*

$$Sep(B,C) \leq \inf|\sigma(B)-\sigma(C)| .$$

If $Sep(B,C) \neq 0$ then

$$Sep(B,C) = \inf_{\|P\|=1} \|T(P)\| .$$

Proof: Equation (4.4) implies that $0 \notin \sigma(T)$ if and only if $\sigma(B) \cap \sigma(C) = \emptyset$. To show the first statement, we note that if $0 \in \sigma(T)$, then it holds automatically. If $0 \notin \sigma(T)$, then $r(T^{-1}) \leq \|T^{-1}\|$. This, together with (4.4), implies the inequality. The second statement is clear, since if T^{-1} exists, then

$$\|T^{-1}\|^{-1} = \inf_{\|P\|=1} \|T(P)\| .$$

We remark that the function Sep is not symmetric. However when X and Y are Hilbert spaces and P in (4.3) is restricted to $HS(X,Y)$ so that $T \in B(HS(X,Y))$, then in this case Sep is symmetric and is the same as the distance between the spectra of B and C if they are selfadjoint (see [37], Theorem 2.6).

Now we show that taking $\|T^{-1}\|^{-1}$ as a measure of the separation is justified. It will be shown that Sep is stable under small perturbations.

<u>Theorem 4.4.</u> *The function*

$$Sep : \mathcal{B}(X) \times \mathcal{B}(Y) \longrightarrow \mathbb{R}^+$$

is Lipschitz continuous with constant 1.

<u>Proof</u>: First we show that

$$\mathrm{Sep}(B+H,\ C+Q) \geq \mathrm{Sep}(B,C) - \|H\| - \|Q\| .$$

If $\mathrm{Sep}(B,C) - \|H\| - \|Q\| \leq 0$ then the above inequality is satisfied. So we consider the case when

$$\mathrm{Sep}(B,C) - \|H\| - \|Q\| > 0 \quad .$$

Let $K \varepsilon \mathcal{B}(\mathcal{B}(X,Y))$ be defined by

$$K(P) = PH - QP .$$

If $S=T+K$ is invertible, then $\mathrm{Sep}(B+H,\ C+Q) = \|S^{-1}\|^{-1}$.

Now

$$\|K\| \leq \|H\| + \|Q\| < \mathrm{Sep}(B,C) = \|T^{-1}\|^{-1} .$$

Since

$$\|(S-T)T^{-1}\| \leq \frac{\|H\|+\|Q\|}{\mathrm{Sep}(B,C)} < 1$$

then from $S = [I+(S-T)T^{-1}]T$, it follows that

$$S^{-1} = T^{-1} \sum_{n=0}^{\infty} (-)^n [(S-T)T^{-1}]^n .$$

Thus

$$\|S^{-1}\| \leq \|T^{-1}\| \frac{1}{1-\|S-T\| \|T^{-1}\|} ,$$

and

$$\|S^{-1}\|^{-1} \geq \|T^{-1}\|^{-1} (1-\|K\| \|T^{-1}\|) \geq \|T^{-1}\|^{-1} -$$

$$- (\|H\| + \|Q\|) .$$

So

$$\mathrm{Sep}(B+H,\ C+Q) \geq \mathrm{Sep}(B,C) - \|H\| - \|Q\| .$$

From the above inequality, it also follows that

$$\mathrm{Sep}(B,C) \geq \mathrm{Sep}(B+H,\ C+Q) - \|H\| - \|Q\| .$$

Adding up, we have

$$|\mathrm{Sep}(B+H,\ C+Q) - \mathrm{Sep}(B,C)| \leq \|H\| + \|Q\| .$$

4.3. Error Bound

In the setting of Section 1, the smooth path

$$A(t) = A+t\,H \qquad 0\le t\le 1$$

connects $A=A(0)$ with $A+H=A(1)$. According to Theorem 4.1, $\sigma(A(t))$ will be separated into n parts $\sigma_j(A(t))$ $1\le j\le n$, for all $t\varepsilon[0,1]$. Moreover, each $\sigma_j(A(t))$ will be enclosed by Γ_j. The spectral projectors $\tilde{F}_j(t)$ associated with the spectral sets $\sigma_j(A(t))$ are then given by

$$(4.5) \qquad \tilde{F}_j(t) = \frac{1}{2\pi i}\int_{\Gamma_j} (\lambda I-A-tH)^{-1}\,d\lambda\ , \quad 1\le j\le n,\ t\varepsilon[0,1]\ .$$

For each t, $0\le t\le 1$, $\tilde{F}(t)=(\tilde{F}_j(t))_{j=1}^{n}$ will be the spectral projector frame corresponding to $A(t)$. Since $t \longmapsto A(t)$, $0\le t\le 1$ is a smooth path in $\mathcal{B}(X)$, it follows that

$$(4.6) \qquad t \longmapsto \tilde{F}(t) = (\tilde{F}_j(t))_{j=1}^{n}$$

is a frame valued continuously differentiable path, with

$$\tilde{F}_j'(t) = \frac{1}{2\pi i}\int_{\Gamma_j} (\lambda I-A-tH)^{-1}\,H(\lambda I-A-tH)^{-1}\,d\lambda\ , \quad 1\le j\le n\ .$$

In particular,

$$(4.7) \qquad \tilde{F}_j'(0) = \frac{1}{2\pi i}\int_{\Gamma_j} (\lambda I-A)^{-1}\,H(\lambda I-A)^{-1}\,d\lambda\ , \quad 1\le j\le n\ .$$

Furthermore, it follows from $\tilde{F}_j^2(t)=\tilde{F}_j(t)$ that

$$(4.8) \qquad \tilde{F}_j'(t)\,\tilde{F}_j(t) + \tilde{F}_j(t)\,\tilde{F}_j'(t) = \tilde{F}_j'(t)\ , \quad j=1,\dots,n\ .$$

Also

$$\sum_{j=1}^{n} \tilde{F}_j(t)=I \quad \text{implies that} \quad \sum_{j=1}^{n} \tilde{F}_j'(t) = 0\ .$$

Equation (4.6) gives an interpolating continuously differentiable path connecting $E=\tilde{F}(0)$ and $F=\tilde{F}(1)$ within $E^n(X)$. Recall that $E^n(X)$ denotes the set of all frames in $B^n(X)$. Theorem 3.4 showed that $E^n(X)$ is a Banach manifold and the tangent space $T(E)$ at $E\varepsilon E^n(X)$ was identified with

$$T(E) = \{H\varepsilon B^n(X) : H_j = E_jH_j + H_jE_j\ ,$$

$$\Sigma H_j = 0\ ,\ 1\leq j\leq n\}\ .$$

It follows from equation (4.8) that $\tilde{F}'(0)$ is indeed in the tangent space at E. Let

$$\tilde{H}_0 = \tilde{F}'(0) = (\tilde{F}_j'(0))_{j=1}^{n}\ .$$

Define $\tilde{L}$ as follows

$$(4.9) \qquad \tilde{L} = \frac{1}{2}\sum_j (\tilde{H}_{0j}\,E_j - E_j\,\tilde{H}_{0j})\ .$$

We note that

$$(4.10) \quad \sum_{j=1}^{n} E_j \tilde{L} E_j = \frac{1}{2}\left[\sum_{j=1}^{n} E_j \tilde{H}_{0j} E_j - \sum_{j=1}^{n} E_j \tilde{H}_{0j} E_j\right] = 0 ,$$

so that $\tilde{L} \varepsilon M(E)$, where $M(E)$ is the model space defined by equation (3.7).

We remark that $\tilde{L}$ as defined by equation (4.9) was used to define Kato's transformation which we have mentioned in Chapter I. Further $\tilde{L}$ satisfies

$$\tilde{L} = \frac{1}{2} \sum_{j=1}^{n} \left(\tilde{H}_{0j} E_j - E_j H_{0j}\right) = \sum_{j=1}^{n} \tilde{H}_{0j} E_j = - \sum_{j=1}^{n} E_j \tilde{H}_{0j} .$$

The equality of the three members of the above equation follows since

$$\sum_{j=1}^{n} \tilde{F}_j'(0) E_j + \sum_{j=1}^{n} E_j \tilde{F}_j'(0) = \sum_{j=1}^{n} (\tilde{F}_j'(0)^2)' =$$

$$= \sum_{j=1}^{n} (\tilde{F}_j(0))' = 0 .$$

In what follows, we try to bound $\tilde{L}$ defined by equation (4.9), though our goal is to find a bound on the generator L of the geodesic between E,F . Finding an estimate on $\tilde{L}$ will be the first step to bounding L .

Adopting the notation introduced in Chapter 1, and fixing the frame E , the spectral projector frame corresponding to A , we have

$$(4.11) \quad [A]_E = \begin{bmatrix} A_{11} & & & \\ & A_{22} & & 0 \\ & & \ddots & \\ 0 & & & A_{nn} \end{bmatrix} ,$$

$$(4.12) \quad [H]_E = \begin{bmatrix} H_{11} & \cdots & H_{1n} \\ H_{21} & \cdots & H_{2n} \\ H_{n1} & \cdots & H_{nn} \end{bmatrix} .$$

We use (4.11) and (4.12) to find the entries of the block matrix of $\tilde{F}_j'(0)$. This will be given by the next lemma.

<u>Lemma 4.5</u>. *Let* $[\tilde{F}_i'(0)]_E = (\tilde{E}'_{i,jk})_{j,k=1}^n$, *then*

$$[\tilde{F}_i'(0)]_E = \begin{bmatrix} & & \tilde{E}'_{i,1i} & & \\ 0 & & & & 0 \\ & & \tilde{E}'_{i,2i} & & \\ & & \vdots & & \\ \tilde{E}'_{i,i1} & \tilde{E}'_{i,i2} \cdots & 0 & \tilde{E}'_{i,i,i+1} & \cdots \tilde{E}'_{i,in} \\ & & \tilde{E}'_{i,i+1,i} & & \\ 0 & & \vdots & & 0 \\ & & \tilde{E}'_{i,ni} & & \end{bmatrix}$$

$i = 1, \ldots, n$.

<u>Proof</u>: Recall that $\tilde{F}_i'(0)$ as given by equation (4.8) is

$$\tilde{F}_i'(0) = \frac{1}{2\pi i} \int_{\Gamma_i} (\lambda I - A)^{-1} H (\lambda I - A)^{-1} d\lambda .$$

Let

$$[\tilde{F}_i'(0)]_E = (\tilde{E}_{i,jk}')_{j,k=1}^n .$$

Expressing A and H by their block matrices, as given by (4.11) and (4.12), in the right hand side of the above expression of $\tilde{F}_i'(0)$, we get

$$\tilde{E}_{i,jk}' = \frac{1}{2\pi i} \int_{\Gamma_i} (\lambda I - A_{jj})^{-1} H_{jk} (\lambda I - A_{kk})^{-1} d\lambda .$$

By our assumption on the spectrum of A , we know that Γ_i encloses only $\sigma(A_{ii})$, and therefore

$$\tilde{E}_{i,jk}' = 0 \quad j \neq i \text{ and } k \neq i .$$

Also since $E_i \tilde{F}_i'(0) E_i = 0$, as it follows from (4.8), we have $\tilde{E}_{i,ii}' = 0$. That is, in matrix representation, $\tilde{E}_{i,jk}'$ has all entries zero except for the i^{th} row and i^{th} column.

We now study $\tilde{F}_i'(0)$ more closely: the only values of $\tilde{E}_{i,jk}'$ to be considered are when $j=i$ or $k=i$. If we fix $k=i$, then with each j we associate the following map

$$S_{ji}: \; B(X_i, X_j) \longrightarrow B(X_i, X_j)$$

$$(4.13) \quad S_{ji}(Q) = \frac{1}{2\pi i} \int_{\Gamma_i} (\lambda I - A_{jj})^{-1} \, Q \, (\lambda I - A_{ii})^{-1} \, d\lambda \; .$$

The next theorem determines the inverse of S_{ji}, $j=1,\ldots,n$, $j \neq i$.

<u>Theorem 4.6</u>. *The map* S_{ji} :

$$S_{ji}(Q) = \frac{1}{2\pi i} \int_{\Gamma_i} (\lambda I - A_{jj})^{-1} \, Q \, (\lambda I - A_{ii})^{-1} \, d\lambda, \quad Q \in B(X_i, X_j)$$

is a bounded linear map and it is the inverse of the map T_{ji}, *where*

$$(4.14) \quad T_{ji}(P) = PA_{ii} - A_{jj}P \, , \qquad P \in B(X_i, X_j)$$

$j=1,\ldots,n \quad j \neq i$.

<u>Proof</u>: It is clear that $T_{ji} \in B(B(X_i, X_j))$. We will show that

$$T_{ji} \, S_{ji} = S_{ji} \, T_{ji} = I \, .$$

From (4.14) we have

$$T_{ji}(S_{ji}(Q)) = S_{ji}(Q) \, A_{ii} - A_{jj} \, S_{ji}(Q) \, ,$$

and by (4.13) this expression equals

$$\frac{1}{2\pi i}\int_{\Gamma_i} (\lambda I-A_{jj})^{-1}\, Q(\lambda I-A_{ii})^{-1}\, A_{ii}\,d\lambda - \frac{1}{2\pi i}\int_{\Gamma_i} A_{jj}(\lambda I-A_{jj})^{-1}\, Q(\lambda I-A_{ii})^{-1}\, d\lambda \quad .$$

But since

$$A_{jj}(\lambda I-A_{jj})^{-1} = -I + \lambda(\lambda I-A_{jj})^{-1}$$

and

$$(\lambda I-A_{ii})^{-1}\, A_{ii} = -I + \lambda(\lambda I-A_{ii})^{-1} \ ,$$

we have

$$T_{ji}(S_{ji}(Q)) = \frac{1}{2\pi i}\left[\int_{\Gamma_i} Q(\lambda I-A_{ii})^{-1}\, d\lambda - \int_{\Gamma_i} (\lambda I-A_{jj})^{-1} Q\,d\lambda\right] .$$

By the operational calculus

$$\frac{1}{2\pi i}\int_{\Gamma_i} (\lambda I-A_{ii})^{-1}\, d\lambda = I \quad ,$$

and for $j \neq i$,

$$\frac{1}{2\pi i}\int_{\Gamma_i} (\lambda I-A_{jj})^{-1}\, d\lambda = 0 \quad .$$

Hence $T_{ji}(S_{ji}(Q)) = Q$. On the other hand

$$S_{ji}(T_{ji}(P)) = \frac{1}{2\pi i} \int_{\Gamma_i} (\lambda I - A_{jj})^{-1} (PA_{ii} - A_{jj}P)(\lambda I - A_{ii})^{-1} \, d\lambda \ .$$

This reduces by similar computations to

$$S_{ji}(T_{ji}(P)) = P \ ,$$

and the theorem is proved. ///

The above theorem says that

$$\tilde{E}'_{i,ji} = S_{ji}(H_{ji}) \quad j=1,\ldots,n;\ j \neq i \ ,$$

where $S_{ji} = T_{ji}^{-1}$, and T_{ji} is given by (4.14). This gives only the entries of $[\tilde{F}'_i(0)]_E$ in the i^{th} column.

Similarly if we fix $j=i$ in $\tilde{E}'_{i,jk}$, then

$$\tilde{E}'_{i,ik} = \frac{1}{2\pi i} \int_{\Gamma_i} (\lambda I - A_{ii})^{-1} H_{ik} (\lambda I - A_{kk})^{-1} \, d\lambda$$

where $k=1,2,\ldots,n$, $k \neq i$. With each k we associate a bounded linear map S_{ik} as follows

$$S_{ik}(Q) = \frac{1}{2\pi i} \int_{\Gamma_i} (\lambda I - A_{ii})^{-1} Q (\lambda I - A_{kk})^{-1} \, d\lambda \ ,$$

with $Q \in \mathcal{B}(X_k, X_i)$. By the same argument as in Theorem 4.6, we have, S_{ik} is the inverse of T_{ik} defined by

$$T_{ik}(P) = A_{ii}P - PA_{kk} \quad k=1,\ldots,n \quad k \neq i \ .$$

We also note that $\tilde{E}'_{i,ik} = S_{ik}(H_{ik})$ $k=1,\ldots,n$ $k\neq i$ where $S_{ik} = T_{ik}^{-1}$.

In Section 2 we defined a measure for the separation of the spectra of two operators. Specializing this to the case where $B=A_{ii}$, $C=A_{jj}$, we obtain

$$\mathrm{Sep}(A_{ii}, A_{jj}) = \|T_{ji}^{-1}\|^{-1} .$$

By Theorem 4.6 we have

$$\mathrm{Sep}(A_{ii}, A_{jj}) = \|S_{ji}\|^{-1} \quad i,j=1,\ldots,n \quad j\neq i . \tag{4.15}$$

The following theorem will give a bound on $\tilde{L}$ as defined by equation (4.9) in terms of $\mathrm{Sep}(A_{ii}, A_{jj})$ and the perturbation H .

Theorem 4.7. *Let* $A\varepsilon B(X)$, *and suppose that the spectrum of* A , $\sigma(A)$, *is separated into* n *non-empty parts* $\sigma_j(A)$. *Let* E *be the spectral projector frame corresponding to* A, *and* $H\varepsilon B(X)$ *be a perturbation of* A *with* F *the spectral projector frame associated with* $A+H$. *Set*

$$\sigma = \min_{\substack{1\leq i,j\leq n \\ i\neq j}} \mathrm{Sep}(A_{ii}, A_{jj})$$

and

$$M = n^3\|E\|^3 ,$$

then

(4.16) $\|\tilde{L}\| \leq \frac{M}{\sigma}\|H\|$.

Proof: From equation (4.9) we have

$$\tilde{L} = \frac{1}{2}\sum_{i=1}^{n}(\tilde{F}_i'(0)E_i - E_i\tilde{F}_i'(0))$$

and

$$\|\tilde{L}\| \leq \sum_{i=1}^{n}\|\tilde{F}_i'(0)\| \; \|E_i\| \; .$$

By the way we normed $B^n(X)$, $\|E\| = \max_{1\leq i\leq n}\|E_i\|$, so that

$$\|\tilde{L}\| \leq \|E\| \sum_{i=1}^{n}\|\tilde{F}_i'(0)\| \; .$$

We will use the block matrix structure of $\tilde{F}_i'(0)$ developed in Lemma 4.5, and Theorem 4.6 to bound $\tilde{F}_i'(0)$. Recall from Chapter I that the norm defined in $BM_E(X)$ is

$$\|[\tilde{F}_i'(0)]_E\| = \max_j \sum_k \|\tilde{E}_{i,jk}'\| \; .$$

Direct calculation shows that

$$\|\tilde{F}_i'(0)\| \leq n\,\|E\| \quad \|[\tilde{F}_i'(0)]_E\| \; .$$

But for $i=1,\ldots,n$

$$[\tilde{F}_i'(0)]_E = (\tilde{E}_{i,jk}')_{j,k=1}^n \quad ,$$

where

$$\tilde{E}_{i,jk}' = 0 \quad j \neq i \ , \quad \text{or} \quad k \neq i \ , \quad \text{or} \quad i=j=k$$

$$\tilde{E}_{i,ji}' = S_{ji}(H_{ji}) \quad 1 \leq j \leq n \ , \ j \neq i$$

$$\tilde{E}_{i,ik}' = S_{ik}(H_{ik}) \quad 1 \leq k \leq n \ , \ k \neq i \ .$$

Equation (4.15) implies

$$\|\tilde{E}_{i,ji}'\| \leq \frac{1}{\mathrm{Sep}(A_{ii},A_{jj})} \ \|H_{ji}\| \quad 1 \leq j \leq n \ , \ j \neq i$$

$$\|\tilde{E}_{i,ik}'\| \leq \frac{1}{\mathrm{Sep}(A_{kk},A_{ii})} \ \|H_{ik}\| \quad 1 \leq k \leq n \ , \ k \neq i \ .$$

Hence

$$\|\tilde{E}_{i,jk}'\| \leq \frac{1}{\sigma} \|H_{jk}\| \quad 1 \leq j, \ k \leq n \ ,$$

and consequently

$$\|[\tilde{F}_i'(0)]_E\| \leq \frac{1}{\sigma} \ \|[H]_E\| \ .$$

But since

$$\|\tilde{L}\| \leq \|E\| \sum_{i=1}^n \|\tilde{F}_i'(0)\| \leq n\|E\|^2 \quad \sum_{i=1}^n \|[F_i'(0)]_E\|$$

we have

$$\|\tilde{L}\| \leq \frac{M}{\sigma} \|H\| \ . \quad ///$$

The rest of this section will be devoted to finding an estimate on L, the generator of the geodesic between E,F. In Chapter III we discussed how to obtain the generator L, and it turned out that it can be obtained as the solution of the operator equation

$$\exp L - P_{F,E}(\exp L) + P_{E,E}(L) = 0 \tag{4.17}$$

Theorem 4.8. *Under the assumptions of Theorem 4.7, let $L\varepsilon\mathcal{B}(X)$ be the generator of the geodesic between E and F. Then*

$$\|L-\tilde{L}\| \leq K\,\|\exp\tilde{L}\|$$

where K is a constant given by (4.18) below.

Proof: Consider the map

$$\phi(F,L) = \exp L - P_{F,E}(\exp L) + P_{E,E}(L) \quad F\varepsilon\mathcal{B}^n(X),\ L\varepsilon\mathcal{B}(X) .$$

Clearly ϕ is a smooth map $\mathcal{B}^n(X) \times \mathcal{B}(X) \longrightarrow \mathcal{B}(X)$. Further $\phi(E,0)=0$, $\phi_2'(E,0)=I\varepsilon\mathcal{B}(\mathcal{B}(X))$, where $\phi_2'(E,0)$ denotes the partial derivative with respect to the second argument. Hence the conditions of the implicit function theorem are satisfied. Therefore there exists a neighbourhood U_0 of E in $\mathcal{B}^n(X)$ and a neighbourhood V of $(E,0)$ in $\mathcal{B}^n(X) \times \mathcal{B}(X)$ and a unique smooth map $L_0\colon U_0 \longrightarrow \mathcal{B}(X)$ such that for each $F\varepsilon U_0$ we have $(F, L_0(F))\,\varepsilon\, V$ and $\phi(F,L_0(F))=0$.

So if $F \varepsilon U_0 \cap E^n(X)$ then $L = L_0(F)$ will be the generator of the geodesic between E and F. The actual procedure of finding L goes as follows: For a fixed $F \varepsilon U_0 \cap E^n(X) = U \subset U_0$ define a function

$$\psi^F(L) = L - \phi(F,L) \, , \; L \varepsilon L_0(U_0) \; .$$

We will show that ψ^F is a contraction map for any $F \varepsilon U$. To do so, we show first that for every $\varepsilon > 0$, a $\delta > 0$ can be found such that $\psi^F(L)$ maps the sphere $\|L\| \leq \varepsilon$ into itself for $\|F-E\| < \delta$. Differentiating ψ^F we have

$$\psi^F(L)' = I - \phi_2'(F,L) = \phi_2'(E,0) - \phi_2'(F,L)$$

$$\|\psi^F(L)'\| \leq \|\phi_2'(E,0) - \phi_2'(F,L)\| \; .$$

By the continuity of ϕ_2', the qunatity on the right hand side can be made arbitrarily small. So let

$$\|\psi^F(L)'\| \leq \alpha < 1 \quad (\|L\| \leq \varepsilon \; , \; \|F-E\| < \delta) \; .$$

Since $\psi^F(0) = -\phi(F,0)$, then

$$\|\psi^F(0)\| = \|\phi(F,0)\| \leq \|\phi(F,0) - \phi(E,0)\| \; .$$

By the continuity of ϕ, the right hand side in the above inequality can be made as small as we please by reducing δ, let δ be sufficiently small that

$$\|\psi^F(0)\| = \|\phi(F,0) - \phi(E,0)\| \le \varepsilon(1-\alpha) \;,\; (\|F-E\| < \delta) \;.$$

Now for $\|F-E\|<\delta$ and $\|L\| \le\varepsilon$, we have

$$\|\psi^F(L)\| \le \|\psi^F(0)\| + \|\psi^F(L) - \psi^F(0)\|$$

$$\le \varepsilon(1-\alpha) + \sup_{0<\theta<1} \|\psi^{F'}(\theta L)\| \, \|L\| \;.$$

(Here we have applied the mean value theorem for the second quantity). Thus

$$\|\psi^F(L)\| \le \varepsilon(1-\alpha) + \varepsilon\alpha = \varepsilon \;.$$

Hence ψ^F maps the sphere $\|L\|\le\varepsilon$ into itself, while the derivative is bounded by a constant $\alpha\varepsilon(0,1)$. That is ψ^F is a contraction with α as a contraction factor. Hence there exists in the sphere $\|L\|\le\varepsilon$ a unique fixed point L of ψ^F , which implies $\phi(F,L)=0$, where L will be the geodesic between E and F . Knowing that ψ^F is a contraction, we can obtain L by successive approximations. Starting the successive approximation at $\tilde{L}$, we obtain

$$L_0 = \tilde{L} \quad L_{n+1} = \psi^F(L_n) \;,\; n=1,2,\ldots \;.$$

Thus from the error formula for the successive approximations

$$\|L-L_n\| \le \frac{\alpha^n}{1-\alpha} \|L_1-L_0\| \quad n=0,1,\ldots \;.$$

Since $L_1=\tilde{L}-\phi(F,\tilde{L})$ and $P_{E,E}(\tilde{L})=0$ (as it follows from (4.10)) we obtain

$$\|L-\tilde{L}\| \le \frac{1}{1-\alpha} \|L_1-\tilde{L}\| = \frac{1}{1-\alpha}\|\phi(F,\tilde{L})\|$$

$$\le \frac{1}{1-\alpha} \|\exp\tilde{L} - P_{F,E} \exp\tilde{L}\|$$

$$\le \frac{1}{1-\alpha} \|1 - P_{F,E}\| \, \|\exp\tilde{L}\| .$$

This completes the proof of the theorem with

$$(4.18) \quad K = \frac{1}{1-\alpha} \|1 - P_{F,E}\| . \quad ///$$

We note here that the estimate

$$(4.19) \quad \|\exp L\| = \Big\|\sum_{n=0}^{\infty} \frac{L^n}{n!}\Big\| \le \sum_{n=0}^{\infty} \frac{\|L\|^n}{n!} = \exp\|L\|$$

is often overestimated, a better estimate can be given by

$$(4.19') \quad \|\exp L\| \le \inf_{t\varepsilon R} e^{-t+\|L+tI\|} .$$

This is because

$$\exp L = \exp(L+tI-tI) = \exp(-t)\exp(L+tI) , \quad t\varepsilon R ,$$

So that

$$\|\exp L\| \le e^{-t} \|\exp(L+tI)\| \le e^{-t+\|L+tI\|} ; \ t\varepsilon R .$$

We will give an example to show that the strict inequality sign may hold in (4.19) and equality in (4.19').

Example 4.9. Let

$$L = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{with eigenvalues } +i, -i \ .$$

Thus

$$\exp L = e^{i}P_1 + e^{-i}P_2$$

where

$$P_1 = \frac{L+iI}{2i}, \quad P_2 = \frac{L-iI}{2i} .$$

Hence

$$\exp L = \begin{bmatrix} \cos 1 & -\sin 1 \\ \sin 1 & \cos 1 \end{bmatrix} .$$

We use here the Euclidean norm, i.e. $\|L\| = [r(L^*L)]^{1/2}$. Note that L is skew symmetric $r(L^*L) = r(-L^2) = r(I) = 1$. Hence $\|L\| = 1$. Since L is skew symmetric, it follows that $\exp L$ is unitary, so $\|\exp L\| = 1$ and $\|\exp L\| < \exp\|L\| = e$. On the other hand $\|L+tI\| = \sqrt{1+t^2}$ and

$$\operatorname*{Inf}_{t\in R} \exp(-t+\|L+tI\|) = \exp \inf_{t\in R} (-t+\sqrt{1+t^2})$$

$$= \exp \inf_{t\in R} \frac{1}{t+\sqrt{1+t^2}} = 1$$

and indeed

$$\|\exp L\| = \inf_{t\in R} \exp(-t+\|L+tI\|) .$$

We summarize the results of this section in the following theorem which at the same time answers the open question [9, X, Question 3], though we have a more general setting.

Theorem 4.10. *Let* $A\in B(X)$ *, suppose the spectrum of* A *,* $\sigma(A) = \bigcup_{j=1}^{n} \sigma_j(A)$ *where the* $\sigma_j(A)$*'s are disjoint spectral sets. Let* E *be the spectral projector frame corresponding to* A *, and let* H *be a small perturbation of* A *, with* F *the spectral projector frame associated with* $A+H$ *. Then we have the following estimate on the oriented angle* L *between* E *and* F *:*

$$\|L\| \leq K\,\gamma(\tilde{L}) + \|\tilde{L}\|$$

where

$$\gamma(\tilde{L}) = \operatorname*{Inf}_{t\in R} e^{-t+\|\tilde{L}+tI\|} .$$

and

$$\|\tilde{L}\| \leq \frac{M}{\sigma} \|H\| .$$

Proof: Since $\|L\| \leq \|L-\tilde{L}\| + \|\tilde{L}\|$, then Theorem 4.8 together with inequality (4.19) give an estimate on the first term. For the second term we use theorem 4.7. ///

4.4. Error bound in the selfadjoint case.

In the previous section we considered an operator A with the spectral projector frame E associated with it and the spectral projector frame F associated with $A+H$. If A and H are selfadjoint operators on a Hilbert space $\mathcal{H}$, then the frames E and F will be orthoprojector frames. In this section we will be mainly considering this case.

Suppose that the spectrum of $\sigma(A)$ is confined to n intervals I_j, $j=1,\ldots,n$ and $E=(E_1,\ldots,E_n)$ is the spectral orthoprojector frame associated with A. E will give rise to the decomposition of $\mathcal{H}$ as follows:

$$\mathcal{H} = \mathcal{H}_1 \oplus \ldots \oplus \mathcal{H}_n$$

where $\mathcal{H}_j = E_j \mathcal{H}$.

Let

$$s_j : \mathcal{H}_j \longrightarrow \mathcal{H} \quad j=1,2,\ldots,n$$

be the injections of $\mathcal{H}_j$'s into $\mathcal{H}$, then

$$s_j s_j^* = E_j \quad , \quad s_j^* s_j = I \quad , \quad s_k^* s_j = 0 \quad j \neq k \quad j,k=1,\ldots,n .$$

Also let

$$\tilde{s}_j : \mathcal{H} \ominus \mathcal{H}_j \longrightarrow \mathcal{H}$$

such that

$$\tilde{s}_j\tilde{s}_j^* + s_js_j^* = I .$$

Hence $\tilde{s}_j\tilde{s}_j^* = I - E_j = \tilde{E}_j$ and $\tilde{s}_j^*s_j = 0 = s_j^*\tilde{s}_j$.

Clearly

$$[A]_E = \begin{bmatrix} A_{11} & & 0 \\ & \ddots & \\ 0 & & A_{nn} \end{bmatrix} .$$

For small selfadjoint perturbations H , the spectrum of A+H will be contained in the intervals J_j , j=1,...,n . Let F be the spectral orthoprojector frame of A+H . As above F gives rise to the decomposition of H into the direct sum

$$H = H_1' \oplus \dots \oplus H_n'$$

with $H_j' = F_j\, H$.

Similarly let t_j j=1,2,...,n be the injections of H_j' into H , hence as before

$$t_jt_j^* = F_j,\ t_j^*t_j = I \text{ and } t_k^*t_j = 0\ j\neq k\ (j,k=1,\dots,n) .$$

For each j let $\tilde{t}_j : H \ominus H_j' \longrightarrow H$ be such that

$$\tilde{t}_j\tilde{t}_j^* + t_jt_j^* = I,\ \tilde{t}_j^*t_j = 0 = t_j^*\tilde{t}_j .$$

F is the spectral projector frame of A+H , therefore

$$[A+H]_F = \begin{bmatrix} B_{11} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & B_{nn} \end{bmatrix}.$$

With respect to the frames $[E_j, \tilde{E}_j]$ and $[F_j, \tilde{F}_j]$, the operators A and $A+H$ will have the following block matrix representation

$$[A]_{[E_j, \tilde{E}_j]} = \begin{bmatrix} A_{jj} & 0 \\ 0 & \tilde{A}_{jj} \end{bmatrix},$$

$$[A+H]_{[F_j, \tilde{F}_j]} = \begin{bmatrix} B_{jj} & 0 \\ 0 & \tilde{B}_{jj} \end{bmatrix}.$$

As mentioned before, the distance between the spectra of two selfadjoint operators is a stable measure of the separation of the spectra. This will be the measure we use in the next theorem.

Theorem 4.11. *Let* A , $A+H$ *be as described above, and suppose*

(4.20) $\quad dist(\sigma(\tilde{B}_{jj}), \sigma(A_{jj})) \geq \delta_j , \quad \delta_j > 0 , \quad j=1,\ldots,n .$

Then

(4.21) $\|\Sigma\delta_j \tilde{F}_j E_j\| \leq n \|H\|$.

Also let $V(F,E) = \Sigma F_j E_j$, *and set* $\delta = \min_{1\leq j\leq n} \delta_j$, *then*

$$\delta\|I-V\| \leq n \|H\| .$$

Furthermore if we assume also that

$$dist(\sigma(B_{jj}) , \sigma(\tilde{A}_{jj})) \geq \delta_j \ ; \ j=1,\ldots,n ,$$

then

$$\delta\|F-E\| \leq \|H\| .$$

In particular for $n=2$ *we have*

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

Proof: Since

$$A = s_j A_{jj} s_j^* + \tilde{s}_j \tilde{A}_{jj} \tilde{s}_j^* \quad j=1,\ldots,n ,$$

and

$$A+H = t_j B_{jj} t_j^* + \tilde{t}_j \tilde{B}_{jj} \tilde{t}_j^* \quad j=1,\ldots,n ,$$

we have for $j=1,\ldots,n$

$$Hs_j = (A+H)\, s_j - As_j = (A+H)\, s_j - s_j A_{jj} .$$

Taking the adjoint of the above equation, we get

$$s_j^* H = s_j^*(A+H) - A_{jj} s_j^*$$

and multiplying from the right by $\tilde{t}_j$

$$s_j^* H \tilde{t}_j = s_j^*(A+H) \tilde{t}_j - A_{jj} s_j^* \tilde{t}_j = s_j \tilde{t}_j \tilde{B}_{jj}$$

$$- A_{jj} s_j^* \tilde{t}_j .$$

Set

$$X_j = s_j^* \tilde{t}_j \in \mathcal{B}(\tilde{F}_j H, E_j H) , \quad Q_j = s_j^* H \tilde{t}_j \in \mathcal{B}(\tilde{F}_j H, E_j H)$$

and the above equation becomes

$$Q_j = X_j \tilde{B}_{jj} - A_{jj} X_j .$$

This operator equation was considered in [9, V], and under our assumption (4.20), Theorem 5.1 there can be applied to our situation, to give

$$\delta_j \| X_j \| \leq \| Q_j \| .$$

This leads to

$$\delta_j \| s_j^* \tilde{t}_j \| = \delta_j \| E_j \tilde{F}_j \| = \delta_j \| \tilde{F}_j E_j \| \leq \| s_j^* H \tilde{t}_j \| = \| E_j H \tilde{F}_j \| ,$$

consequently

$$(4.22) \quad \delta_j \| \tilde{F}_j E_j \| \leq \| \tilde{F}_j H E_j \| \leq \| H \| \qquad j=1,\ldots,n .$$

Adding up we get inequality (4.21). Since

$$V = V(F,E) = \sum_{j=1}^{n} F_j E_j \quad , \delta = \min_{1 \le j \le n} \delta_j \ ,$$

we have

$$\delta \| I - V \| \le n \| H \|$$

by virtue of (4.21) and $I - V = \Sigma \tilde{F}_j E_j$.

If we further assume that dist$(\sigma(B_{jj})\ ,\ \sigma(\tilde{A}_{jj})) \ge \delta_j$ $j=1,2,\dots,n$, then an inequality similar to (4.22) can be obtained namely

$$\delta_j \| F_j \tilde{E}_j \| \le \| F_j H \tilde{E}_j \| \qquad j=1,\dots,n \ .$$

Lemmas (6.1), (6.2) in [9] together with the above inequality and inequality (4.21) imply

$$\delta_j \| \tilde{F}_j E_j + F_j \tilde{E}_j \| \le \| \tilde{F}_j H E_j + F_j H \tilde{E}_j \| \le \| H \| \quad j=1,\dots,n \ .$$

But since

$$\delta_j \| \tilde{F}_j E_j + F_j \tilde{E}_j \| = \delta_j \| (\tilde{F}_j E_j + F_j \tilde{E}_j)(2E_j - I) \| = \delta_j \| F_j - E_j \|$$

(the bound norm being unitary invariant), hence we have

$$\delta_j \| F_j - E_j \| \le \| H \| \qquad j=1,\dots,n \ .$$

This implies

$$\delta \| F - E \| \le \| H \| \ .$$

For 2-frames this is exactly the $\sin\Theta$ Theorem [9], since $\| F_1 - E_1 \| = \| \sin\Theta \| = \| F_2 - E_2 \|$. This completes the proof.

<u>Remark 4.12</u>. We mention here that, in a private communication Davis wrote, that under assumption (4.20), he got a stronger estimate than that in (4.21), namely

$$\| \Sigma \delta_j \tilde{F}_j E_j \| \le \| H \| .$$

BIBLIOGRAPHY

[1] Afriat, S.N. *Orthogonal and Oblique Projectors and the Characteristics of Pairs of Vector Spaces.* Proc. Cambridge Philos. Soc. 53 (1957), pp. 800-816.

[2] Bonsall, F.F. and J. Duncan. *Complete Normed Algebras.* Springer Verlag, New York, 1973.

[3] Boothby, W.M. *An Introduction to Differentiable Manifolds and Riemannian Geometry.* Academic Press, New York, 1975.

[4] Brickell, F. and R.S. Clark. *Differentiable Manifolds.* Van Nostrand Reinhold, New York, 1970.

[5] Conway, J. *Functions of One Complex Variable.* Springer-Verlag, New York, 1978.

[6] Davis, C. *Separation of Two Linear Subspaces.* Acta Sci. Math. 19 (1958), pp. 172-187.

[7] Davis, C. *The Rotation of Eigenvectors by a Perturbation.* J. Math. Anal. Appl. 6 (1963), pp. 159-173.

[8] Davis, C. *The Rotation of Eigenvectors by a Perturbation II.* J. Math. Anal. Appl. 11 (1965), pp. 20-27.

[9] Davis, C. and W. Kahan. *The Rotation of Eigenvectors by a Perturbation III.* SIAM J. Numer. Anal. 7 (1970), pp. 1-46.

[10] Dieudonné, J. *Foundations of Modern Analysis.* Academic Press, New York, 1960.

[11] Dunford, N. and J.T. Schwartz. *Linear Operators.* Parts I, II. Interscience, New York, 1963.

[12] Faddeev, D.K. and V.N. Faddeeva. *Computational Methods of Linear Algebra.* W.H. Freeman, San Francisco, 1963.

[13] Falk, S. *Einschleissungssätze für die Eigenvektoren normaler Matrizenpaare.* Zeitschrift für Angewandte Mathematik und Mechanik 45 (1965), pp. 47-56.

[14] 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. Friedricks, O.E. Neugebauer and J.J. Stocker, Interscience, New York, (1948).

[15] Foias, C. *Une Application des Distribution Vectorelles á la Théorie Spectrale.* Bull. Sci. Math. (2) 84 (1960), pp. 147-158.

[16] Gavurin, M.K. *Approximate Determination of Proper Values and Perturbation Theory.* Amer. Math. Soc. Transl. (2) 16 (1960), pp. 385-388.

[17] Gelfand, I. *Normierte Ringe*. Rec. Math (Math. Sbornik) N.S. 9 (1941), pp. 3-24.

[18] Goetz, A. *Introduction to Differential Geometry*. Addison Wesley, Don Mills, 1970.

[19] Hille, E. and R.S. Phillips. *Functional Analysis and Semigroups*. Amer. Math. Soc. Colloq. Publ. Vol. 31, Amer. Math. Soc., Providence, R.I., 1957.

[20] Jordan, C. *Essai sur la Géometrie à n Dimensions*. Bull. Soc. Math. France 3 (1875), pp. 103-174.

[21] Kato, T. *On the Upper and Lower Bounds of Eigenvalues*. Journal of the Physical Society of Japan 4 (1949), pp. 334-339.

[22] Kato, T. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 1966.

[23] Kantorovič, L.V. *Functional Analysis in Normed Spaces*. The MacMillan, New York, 1964.

[24] Kantorovitz, S. *Classification of Operators by Means of their Operational Calculus*. Trans. Amer. Math. Soc. 115 (1965), pp. 194-224.

[25] Kovarik, Z.V. *Similarity and Interpolation between Projectors*. Acta Sci. Math. (Szeged) 39 (1977), pp. 311-315.

[26] Kovarik, Z.V. *Manifolds of Linear Involutions*. Lin. Alg. Appl. 24 (1979), pp. 271-287.

[27] Kovarik, Z. *Manifolds of Projector Frames.* McMaster Univ. Math. Report No. 103, Hamilton, Canada (1978).

[28] Lang, S. *Analysis II.* Addison-Wesley, New York, 1969.

[29] Lang, S. *Differential Manifolds.* Addison-Wesley, New York, 1972.

[30] Lumer, G. and M. Rosenblum. *Linear Operator Equations.* Proc. Amer. Math. Soc., 10 (1969), pp. 32-41.

[31] Maeda, F. *Generalized Spectral Operators on Locally Convex Spaces.* Pacific J. Math., 11 (1963), pp. 177-192.

[32] Neudecker, H. *A Note on Kronecker Matrix Products and Matrix Equation System.* SIAM J. Appl. Math. 17 (1969), pp. 603-606.

[33] Phillips, E.G. *Function of a Complex Variable.* Longman, London, 1957.

[34] Powers, D. *A Block Gerschgorin Theorem.* Lin. Alg. Appl. 13 (1976), pp. 45-52.

[35] Rosenblum, M. *On the Operator Equation* $BX-XA=Q$. Duke Math. J. 23 (1956), pp. 263-269.

[36] Rudin, W. *Functional Analysis.* McGraw-Hill, New York, 1973.

[37] Stewart, G.W. *Error Bounds for Approximate Invariant Subspaces.* SIAM J. Numer. Anal. 8 (1971), pp. 796-808.

[38] Swanson, C.A. *An Inequality for Linear Transformation with Eigenvalues.* Bull. Amer. Math. Soc. 67 (1961), 607-608.

[39] Sz.-Nagy and F. Riesz. *Leçons d'Analyse Fonctionnelle.* Budapest: Académiai Kiadó, 1953.

[40] Weinberger, H.F. *Error Bounds in the Ràyleigh-Ritz Approximation of Eigenvectors.* J. Research of the National Bureau of Standards. Section B 64B (1960), pp. 217-225.