Deformations of operators

Citation for published version (APA):

Document status and date:
Published: 01/01/1978

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
https://www.tue.nl/index.php?id=71870

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 12. Feb. 2019
Deformations of Operators

by

H.A.M. Daniëls

T.H.-Report 78-WSK-02

June 1978
## A. Summary

## B. Basic Notions and Notations

### I. The theory of Arnold. Deformations of Matrices

#### § 0. Introduction

#### § 1. Holomorphic mappings in $\mathbb{C}^{n \times n}$. Instability of the Jordan normal form

#### § 2. Deformations of matrices

#### § 3. The orbit and the centralizer

#### § 4. Construction of versal deformations

#### § 5. Functions of versal families

### II. On Orbits and Centralizers of Operators

#### § 0. Introduction

#### § 1. The orbit and the centralizer in a Banach algebra

#### § 2. The centralizer of an operator in $\mathcal{L}(H)$

#### § 3. The centralizer of a normal compact operator

#### § 4. Commutators of operators

#### § 5. Hilbert-Schmidt operators

#### § 6. Examples

#### § 7. The embedding of HS in $\text{HS}^+$

#### § 8. Heuristics

### Appendix

### III. Deformations of Hilbert-Schmidt Operators

#### § 0. Introduction

#### § 1. Slices in $G^+$

#### § 2. Deformations in $\text{HS}^+$

#### § 3. An exponential map

#### § 4. Weakly-versal deformations. Weak-versality $\Rightarrow$ transversality

#### § 5. Deformations of Hilbert-Schmidt operators

## References
Abstract

In this paper we consider versal families of operators. The theory of versal families of various kinds of objects and applications to the corresponding fields have been studied extensively by V.I. Arnold in [ARN II].

A more specialized paper of V.I. Arnold, which inspired us to study families of operators, is [ARN I]. This paper deals with families of matrices. In chapter I we shall give a short description of the theory in [ARN I].

Chapter III deals with deformations of Hilbert-Schmidt operators and in this chapter we shall prove a generalization of one of the theorems in [ARN I]. As a preparation we investigate some properties of operators on Hilbert space in chapter II.

AMS Subject Classification: 47A55, 47B10.
B. Basic Notions and Notations

In this paper some fundamental theorems on Functional Analysis and Differential Geometry are used. Most of the basic concepts and theorems on Functional Analysis, used in this paper, can be found in every textbook on this subject. For example in [DUN I] and [DUN II] or in [HIL]. Some of them, which are more specific and deal with operators on Hilbert space, can be found as exercises in [HAL I].

The basic concepts of Differential Geometry such as Differential Calculus and the theory of Manifolds can be found essentially in [LAN] and [ABR]. The theory of finite dimensional Manifolds is essentially in [GOL].

To cause no ambiguity we want to give here some definitions and notations of rather fundamental concepts, which are defined and notated in many (slightly) different ways.

Functional Analysis

Throughout this paper $H$ will denote a separable infinite dimensional Hilbert space. The letters $E$ and $F$ will denote Banach spaces. A subset $V$ of $E$ is a subspace of $E$ iff $V$ is a linear space and $V$ is closed in $E$. Let $V$ be a subspace of $E$, $V$ splits in $E$ iff $V$ has a closed complement $V' \subset E$, i.e. a subspace $V'$ such that $E = V \oplus V'$.

$L(E \to F)$ denotes the Banach space of bounded linear operators from $E$ into $F$. The norm on $L(E \to F)$ is given by

$$\|A\| := \sup_{x \in E, \|x\|=1} \|Ax\|.$$ 

Furthermore, if $A \in L(E \to F)$, $\text{Ker}(A)$ is the subspace $A^{+}(0) \subset E$. The linear manifold $\text{Ran}(A) \subset F$ is the set $\{Ax \mid x \in E\}$.

$L(E)$ denotes the Banach space of bounded linear operators of $E$ into itself. If $A \in L(E)$, $\sigma(A)$ denotes the spectrum of $A$. If $\lambda \in \mathbb{C}\setminus \sigma(A)$ then $R(\lambda, A) := (\lambda I - A)^{-1}$ is the resolvent of $A$. (I stands for the identity operator.) The theory of spectral sets and operator functions, such as projections defined with contour integrals, can be found in [DUN I], Ch. VII.

If $A \in L(H)$ then $A^{*}$ will denote the adjoint of $A$.

In this paper we use two different topologies on $L(H)$, the uniform operator topology induced by the norm and sometimes the strong operator topology (see [HAL I], Ch. 11 and Ch. 12).
Differential Geometry

In this paper the derivative of a map should be thought of as a linear operator.

B1. Definition (differentiable mapping). If \( f \) is a continuous map from an open subset \( U \subseteq E \) into \( F \) and \( x \in U \) then \( f \) is differentiable at \( x \) iff there is a bounded linear map \( D_x f : \mathcal{L}(E \rightarrow F) \) such that

\[
\lim_{\|h\| \to 0} \frac{\|f(x + h) - f(x) - (D_x f)h\|}{\|h\|} = 0.
\]

The linear map \( D_x f \) is necessarily unique. \( f \) is of class \( C^1 \) in \( U \) (notation \( f \in C^1(U \rightarrow F) \)) iff \( f \) is differentiable at each point of \( U \) and the map \( x \to D_x f \) is continuous from \( U \) into \( \mathcal{L}(E \rightarrow F) \) (norm topology) (see also [LAN], Ch. I, § 3 and [ABR], Ch. I, § 1).

B2. Definition (submanifold). Suppose \( M \) is a \( C^r \)-manifold. A subset \( N \subseteq M \) is a \( C^r \)-submanifold iff at every point \( x \in N \) there is an admissible chart (i.e. compatible with the atlas of \( M \)) \((U_x, \phi)\) such that \( \phi(U_x) = V_1 \times V_2 \), where \( V_1 \) and \( V_2 \) are open neighbourhoods of the origins in the Banach spaces \( F_1 \) respectively \( F_2 \), such that \( \phi(x) = (0,0) \) and \( \phi(U_x \cap N) = V_1 \times \{0\} \) (see also [LAN], Ch. II, § 2 and [ABR], Ch. IV, § 17).

B3. Remark. Note that for \( x \in N \) the tangent space \( T_x N \) to \( N \) at \( x \) is a splitting subspace of the tangent space \( T_x M \) to \( M \) at \( x \) (see [ABR], Ch. II, § 17, p. 45).

B4. Definition (double splitting map). Suppose \( f \) is a map from the \( (C^p, p \geq 1) \) manifold \( M_1 \), into the \( (C^q, q \geq 1) \) manifold \( M_2 \), differentiable at \( x \in M_1 \). Then \( f \) is called double splitting at \( x \in M_1 \) iff

B4.1. \( \ker(D_x f) \) splits in \( T_x M_1 \).

B4.2. \( \text{Ran}(D_x f) \) is closed and splits in \( T_{f(x)} M_2 \).

B5. Definition (transversality of a map and a submanifold). Let \( N \) be a \( (C^p, p \geq 1) \) submanifold of the manifold \( M \). Suppose \( f \) is a map from the \( (C^q, q \geq 1) \) manifold \( \Lambda \) into \( M \), differentiable at \( \lambda \in \Lambda \). \( f \) is called transversal to \( N \) at \( \lambda \) iff
B5.1. \( f(\lambda) \in N \).

B5.2. \( f \) is double splitting at \( \lambda \).

B5.3. \( \text{Ran}(D_\lambda f) \) contains a closed complement of \( T_{f(\lambda)} N \) in \( T_{f(\lambda)} M \).

\( f \) is minimal transversal to \( N \) at \( \lambda \) iff \( f \) is transversal and

\[
T_{f(\lambda)} M = T_{f(\lambda)} N \oplus (D_\lambda f)T_{\lambda} A.
\]

B6. Definition (transversality of two submanifolds). Suppose \( N_1 \) and \( N_2 \) are both submanifolds of the \( C^p \)-manifold \( M \). \( N_1 \) is transversal to \( N_2 \) at \( x \) iff

B6.1. \( x \in N_1 \cap N_2 \).

B6.2. \( T_M x = T_{N_1} x + T_{N_2} x \).

\( N_1 \) is minimal transversal to \( N_2 \) at \( x \) if the sum in B6.2 is a direct sum \( (T_{N_1} x \cap T_{N_2} x = \{0\}) \).

B7. Remark. In many books transversality is defined as follows: \( f \) is transversal to the submanifold \( N \) at \( \lambda \in \Lambda \) iff \( f(\lambda) \not\in N \) or B5.1, B5.2, B5.3.
I. The theory of Arnold. Deformations of Matrices

§ 0. Introduction

In this chapter we shall give a short description of the theory of ver- sal deformations of matrices given by Arnold in [ARN I]. This description covers the sections § 2, § 3 and § 4. We give an additional result in § 5.

§ 1. Holomorphic mappings in $\mathbb{C}^{n \times n}$. Instability of the Jordan normal form

In this section we consider holomorphic mappings from an open subset of $\mathbb{C}^p$ into the matrixalgebra $\mathbb{C}^{n \times n}$ ($\mathbb{C}^{n \times n}$ is the algebra consisting of all $n \times n$ complex matrices, provided with the usual operations).

Holomorphy is defined in the usual way, that is

1.1. Definition. A map $A: U \rightarrow \mathbb{C}^{n \times n}$, where $U$ is open in $\mathbb{C}^p$, is holomorphic in $U$ iff each $\lambda_0 \in U$ has a neighbourhood where $A(\lambda)$ can be developed in a power series

$$A(\lambda) = \sum_{|\alpha|=0}^{\infty} A_\alpha (\lambda - \lambda_0)^\alpha$$

convergent in some matrix-norm; here $\alpha$ is the multi-index $(\alpha_1, \alpha_2, \ldots, \alpha_p)$; $|\alpha| := \alpha_1 + \ldots + \alpha_p$; $A_\alpha \in \mathbb{C}^{n \times n}$ and $(\lambda - \lambda_0)^\alpha := (\lambda_1 - \lambda_0_1)^{\alpha_1} \ldots (\lambda_p - \lambda_0_p)^{\alpha_p}$.

The same definition is used if $\mathbb{C}^{n \times n}$ is replaced by any Banach space (then the sum must be convergent in the Banach space norm).

1.2. Remark. It is well-known that $A$ is holomorphic iff for every bounded linear functional $L$ on the Banach space, the mapping $\lambda \rightarrow L(A(\lambda))$ is holomorphic from $\mathbb{C}^p$ into $\mathbb{C}$ (see [HIL], Ch. III, § 2). In our case ($\mathbb{C}^{n \times n}$) this implies that all entries of $A(\lambda)$, $a_{ij}(\lambda)$, are holomorphic functions of $\lambda$.

The converse is also true of course.

1.3. Instability of the Jordan normal form

Suppose $A(\lambda)$ is a holomorphic map (which will also be called a family) from $\mathbb{C}^p$ into $\mathbb{C}^{n \times n}$. If $A(\lambda)$ is reduced to its Jordan normal form $J(\lambda)$ (see [GAN], Ch. VII, § 7), then in general, $J(\lambda)$ is not a holomorphic function of $\lambda$; $J(\lambda)$ sometimes depends even discontinuously on $\lambda$. 
1.3.1. Example. Define

$$A(\lambda) = \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}; \lambda \in \mathbb{C}.$$ 

The Jordan normal form of $A(\lambda)$ is given by

$$J(\lambda) = \begin{cases} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} & \text{if } \lambda \neq 0 \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \text{if } \lambda = 0. \end{cases}$$

In this example the smoothness of the family is lost by reducing the family to its Jordan normal form. So, if a matrix is only known approximately it is unwise to reduce it to the Jordan normal form. In studying smooth families, we are therefore interested in normal forms to which a family can be reduced without losing the smoothness.

§ 2. Deformations of matrices

2.1. Definition. A deformation of a matrix $A_0$ is a map $A: \Lambda \to \mathbb{C}^{n \times n}$ with $A(0) = A_0$ and holomorphic in an open set containing the origin of $\Lambda$. The space $\Lambda (= \mathbb{C}^p$ for some $p \in \mathbb{N})$ is called the base of the deformation (or the base of the family).

2.2. Definition. Two deformations of $A_0$, $A(\lambda)$ and $B(\lambda)$, are said to be similar, if there is a deformation $C(\lambda)$ of the identity matrix such that

$$A(\lambda) = C(\lambda)B(\lambda)C^{-1}(\lambda)$$

for $\lambda$ in some open set in $\Lambda$ containing $0$ ($C^{-1}(\lambda)$ means $(C(\lambda))^{-1}$).

In other words: the germ of $A(\lambda)$ at $\lambda = 0$ is the germ of $C(\lambda)B(\lambda)C^{-1}(\lambda)$ at $\lambda = 0$.

2.3. Definition. If $A(\lambda)$ is a deformation of $A_0$, depending on $k$ complex parameters, and $\varphi: \mathbb{C}^k \to \mathbb{C}^k$ is a map which is holomorphic in a neighbourhood of $0 \in \mathbb{C}^k$ and satisfies $\varphi(0) = 0$, then we call $A(\varphi(\mu))$ the deformation of $A_0$ induced by $A$ under $\varphi$. Clearly $A(\varphi(\mu))$ depends on $l$ parameters.
2.4. Definition. A deformation \( A(\lambda) \) of \( A_0 \) is called versal iff every deformation \( B(\mu) \) of \( A_0 \) is similar to a deformation induced by \( A \) under a suitable change of the parameters i.e. if there exist \( \lambda \in \mathbb{N} \), a function \( \phi: \mathbb{T}^k \to \mathbb{C}^k \) with \( \phi(0) = 0 \) and a deformation \( C(\mu) \) of the identity matrix such that

\[
B(\mu) = C(\mu)A(\phi(\mu))C^{-1}(\mu).
\]

2.5. Example.

\[
\begin{bmatrix}
\lambda_1 & \lambda_2 \\
\lambda_3 & \lambda_4
\end{bmatrix}
\]
is a versal deformation of \( \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \)

which depends on 4 complex parameters.

2.6. Example.

\[
\begin{bmatrix}
1 + \lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix}
\]
is a 2-parameter versal deformation of \( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \).

2.7. Definition. Let \( N \subset \mathbb{C}^{n \times n} \) be a complex analytic submanifold of the complex analytic manifold \( M := \mathbb{C}^{n \times n} \) (analytic manifold means that the charts are open subsets of \( \mathbb{C}^{n \times n} \) and the chart-functions are holomorphic in the sense of definition 1.1). Let \( A: \Lambda \to M \) be holomorphic in a neighbourhood of \( \lambda \in \Lambda \). Then the map \( A \) is said to be transversal to \( N \) at \( \lambda \in \Lambda \) iff

2.7.1. \( T_A(\lambda)M = (D_A)T_{\Lambda} + T_A(\lambda)N. \)

\( T_A(\lambda)M \) is the tangent space of \( M \) at \( A(\lambda) \), \( T_{\Lambda} \) is the tangent space of \( \Lambda \) at \( \lambda \) and \( T_A(\lambda)N \) is the tangent space of \( N \) at \( A(\lambda) \) which is a subspace of \( T_A(\lambda)M \).

2.8. Remark. The reader should compare definition 2.7 with definition B5 and note that since \( \Lambda \) and \( M \) are both finite dimensional, \( D_{\lambda}A \) is automatically double splitting at any point \( \lambda \in \Lambda \). The sum in 2.7.1 is not necessarily direct. It is possible, although it is not very interesting, that \( \dim T_A(\lambda)M = \dim T_A(\lambda)N = \dim (D_{\lambda}A)T_{\lambda} = n^2. \)
§ 3. The orbit and the centralizer

Consider the space of all $n \times n$ matrices $\mathfrak{gl}_{n \times n}$ and the (Lie) group $G$ of all non-singular matrices. $G$ is an open set in $\mathbb{C}^{n \times n}$ containing the identity matrix $e$. It is well known that $G$ is a connected analytic submanifold of $\mathbb{C}^{n \times n}$.

3.1. Remark. Note that the group of non-singular matrices in $\mathbb{R}^{n \times n}$ is not connected.

3.2. Definition. If $A_0$ is fixed in $\mathfrak{gl}_{n \times n}$ we define the map $\alpha_{A_0}: G \to \mathfrak{gl}_{n \times n}$ by

$$\alpha_{A_0}(g) := gA_0g^{-1}; g \in G.$$ 

$\alpha_{A_0}(G)$ is an analytic submanifold of $\mathfrak{gl}_{n \times n}$ and it is called the orbit $N$ of $A_0$ under the action of the group $G$ (see [GIB]).

$\alpha_{A_0}$ is a holomorphic map, the derivative in $e$: $D_e \alpha_{A_0}$ is a map from the Lie-algebra $T_e G = \mathfrak{gl}_{n \times n}$ into $T_{A_0} \mathfrak{gl}_{n \times n} (= \mathfrak{gl}_{n \times n})$, and satisfies

$$(D_e \alpha_{A_0})C = [C, A_0] = CA_0 - A_0C.$$ 

The derivative $D_{g_0} \alpha_{A_0}$ at an arbitrary point $g_0 \in G$ is given by

$$D_{g_0} \alpha_{A_0}(h) = g_0^{-1}[g_0^{-1}h, A_0]g_0^{-1}.$$ 

The proof of this statement is left to the reader. In Chapter II, § 5 we shall prove an analogous result.

3.3. Remark. If $A$ and $B$ are linear operators $[A, B] := AB - BA$ is called the commutator of $A$ and $B$.

For the sake of brevity we shall write $\text{Ad}_{A_0}$ for $D_e \alpha_{A_0}$.

3.4. Definition. The kernel of the linear map $\text{Ad}_{A_0}$ is called the centralizer of $A_0$ and is denoted by $Z(A_0)$. It consists of all the matrices that commute with $A_0$.

The range of $\text{Ad}_{A_0}$ is the tangent space to the orbit $N$ of $A_0$ at $A_0$.

3.5. Lemma. The codimension of the orbit is the dimension of the centralizer.
Proof. Since $T_e G$ and $T^{n \times n}_{A_0}$ are both vector spaces with the same dimension, $n^2$, and $A_{0 \cdot A}$ is linear we have

$$\dim(\text{Ker } A_{0 \cdot A}) + \dim(\text{Ran } A_{0 \cdot A}) = n^2.$$ 

Hence

$$\text{codim(orbit)} = \dim(\text{centralizer}).$$

We are now able to prove the following fundamental theorem of Arnold. We reproduce the proof in some detail because it is the guideline for further investigations.

3.6. Theorem (Arnold). Equivalence of versality and transversality. A deformation $A(\lambda)$ of $A_0$ is versal iff the mapping $A(\lambda)$ is transversal to the orbit of $A_0$ at $\lambda = 0$.

Proof. Versality implies transversality. Let $A(\lambda)$ be a versal deformation of $A_0$. If $B(\mu)$ is any deformation of $A_0$, then by the versality of $A$, we have

$$B(\mu) = C(\mu)A(\psi(\mu))C^{-1}(\mu).$$

Taking the derivative at $\mu = 0$, of both sides, we get

3.6.1. $\forall_{\lambda \in T_{A_0}^A} (D_0B)\lambda = [(D_0C)\lambda, A_0] + (D_0A)(D_0C)^{-1}(\mu).$

Since 3.6.1 holds for every $B$, and each vector in $T^{n \times n}_{A_0}$ can be written as $(D_0B)\lambda$ for a suitable $B$; each vector is the sum of a vector in the tangent space to the orbit of $A_0$ and a vector in the image of $D_0A$; this is exactly the transversality of the map $A(\lambda)$ at $\lambda = 0$.

Transversality implies versality.

This is more complicated. Suppose $A$ is a transversal mapping. Let $N$ denote the orbit of $A_0$ and $A$ the base of the deformation $A(\lambda)$. By the transversality we have

3.6.2. $T^{n \times n}_{A_0} = T_{A_0} N + (D_0A)T_{A_0} A.$

Without loss of generality we may assume that the sum is a direct sum (i.e. $T_{A_0} N$ and $(D_0A)T_{A_0} A$ are linearly independent), for, if the dimension of $T_{A_0} A$ is greater than the codimension of $T_{A_0} N$ we replace $A$ by a submanifold $A_0 \subset A$ such that the restriction of $A$ to $A_0$ is still transversal to the orbit and
the new sum is a direct sum. If it is proved that a restriction of $A$ is versal, then $A$ itself is certainly versal.

Next we choose a submanifold $V$ in $G$ such that $e \in V$ and $V$ is minimal transversal (see Def. B6) to the centralizer of $A_0$, so

3.6.3. $T_e V \oplus T_e Z(A_0) = T_e G (= \mathbb{C}^{n \times n})$.

(The submanifold $V$ can be chosen of the form $e + (B \cap W)$, where $B$ is the open unit ball in $\mathbb{C}^{n \times n}$ and $W$ is a complement of $Z(A_0)$ in $\mathbb{C}^{n \times n}$.) Define the map $\beta: V \times \Lambda \to \mathbb{C}^{n \times n}$ by

$$\beta(v, \lambda) := vA(\lambda)v^{-1}.$$ 

Then $\beta$ is a holomorphic mapping in a neighbourhood of $(e, 0)$ (considered as a function of $\mathbb{C}^{\dim V + \dim \Lambda}$ into $\mathbb{C}^{n \times n}$) and the derivative at $(e, 0)$

$$\beta_* := \text{D}_{(e, 0)} \beta: T_e V \times T_0 \Lambda \to T_{A_0} \mathbb{C}^{n \times n}$$

is given by

$$\beta_*(w, \zeta) = [w, A_0] + (D_0 A) \zeta; \ w \in T_e V, \ \zeta \in T_0 \Lambda.$$ 

From 3.6.1 and 3.6.2 it follows that $\ker \beta_*$ is trivial and hence $\text{ran} \beta_* = T_{A_0} \mathbb{C}^{n \times n}$. Hence $\beta_*$ is an invertible linear operator. Applying the inverse function theorem we may conclude that $\beta$ is a holomorphic diffeomorphism from a neighbourhood of $(e, 0)$ of $V \times \Lambda$ onto an open set in $\mathbb{C}^{n \times n}$ containing $A_0$. Hence, if $B(\mu)$ is any deformation of $A_0$ and $\mu$ is sufficiently small we have

$$B(\mu) = \beta(v, \lambda)$$

for some $v \in V$ and $\lambda \in \Lambda$. Define

$$C(\mu) := \pi_1 \beta^{-1}(B(\mu))$$

$$\varphi(\mu) := \pi_2 \beta^{-1}(B(\mu))$$

(where $\pi_1$ and $\pi_2$ are the projections of $V \times \Lambda$ onto $V$ respectively $\Lambda$) then for small $\mu$

$$B(\mu) = C(\mu)A(\varphi(\mu))C^{-1}(\mu)$$

which proves the versality of the family $A$. □
3.7. Remark. Note that, if \( f: \mathbb{C}^p \rightarrow E_1 \times E_2 \) is holomorphic, where \( E_1 \) and \( E_2 \) are Banach spaces and \( \pi_1 \) is bounded projection of \( E_1 \times E_2 \) onto \( E_1 \), then \( \pi_1 f: \mathbb{C}^p \rightarrow E_1 \) is holomorphic.

§ 4. Construction of versal deformations

It follows from theorem 3.6 that constructing versal deformations is the same thing as constructing transversal deformations. To do this, the space \( \mathbb{C}^{n \times n} \) is equipped with the usual innerproduct.

4.1. Definition. If \( A \) and \( B \in \mathbb{C}^{n \times n} \) we define

\[
(A, B) := \text{trace}(AB^*) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \overline{b}_{ij}
\]

( , ) has three useful properties

4.1.1. \( (A, A) = \|A\|_F^2 \)

4.1.2. \( (A, B) = (B^*, A^*) \)

4.1.3. \( (XA, B) = (A, X^*B) \)

where \( X \in \mathbb{C}^{n \times n} \) and \( \|A\|_F \) is the Euclidean norm on \( \mathbb{C}^{n \times n} \).

4.2. Lemma. Let \( A_0 \in \mathbb{C}^{n \times n} \). The orthogonal complement (with respect to the innerproduct just defined) of the tangent space to the orbit of \( A_0 \) is the adjoint of the centralizer of \( A_0 \), which is equal to \( Z(A_0^*) \).

Proof. For the proof we refer to [ARNII]. It is a special case of theorem 5.7 in chapter II of this paper.

Note that this lemma constitutes a different proof for \( \text{codim}(\text{orbit}) = \text{dim}(\text{centralizer}). \) Since every versal deformation is transversal to the tangent space to the orbit, the minimum number of parameters equals the codimension of the orbit which is the dimension of the centralizer \( =: d \).

Hence every matrix has a versal deformation with minimum number of parameters equal to \( d \). It can be chosen in the following way

\[
A_0 + B(\lambda)
\]

where \( A_0 \) is the matrix and \( B(\lambda) \) is a family (orthogonal) transversal to the tangent space of the orbit (in the adjoint of the centralizer, \( Z(A_0^*) \)). For
an explicit computation of $Z(A_0^*)$ (where $A_0$ is a Jordan normal form) and explicit examples of versal deformations we refer to [ARNI]. A way to find new versal families from given versal families is described in the next section.

§ 5. Functions of versal families

Let $T \in \mathbb{C}^{n \times n}$ then $F(T)$ denotes the class of all functions of a complex variable which are locally holomorphic in some open set containing $\sigma(T)$. The open set need not to be connected and depends on $f \in F(T)$. If $f \in F(T)$ one can define $f(T)$ which is again an element of $\mathbb{C}^{n \times n}$ (see [DUN I], Ch. VII, § 1).

5.1. Theorem. Suppose $A(\lambda)$ is a versal deformation of $A_0$ with base $A$. Let $f \in F(A_0)$, with

5.1.1. $f$ 1-1 on $\sigma(A_0)$
5.1.2. $f'(A) \neq 0$ if $A \in \sigma(A_0)$

Then $f(A(\lambda))$ is a versal deformation of $f(A_0)$ with base $A$.

Proof. Note that $f \in F(A(\lambda))$, if $\lambda$ is small enough, and hence $f(A(\lambda))$ is well defined for small $\lambda$. Let $\sigma(A_0) = \{\lambda_1, \ldots, \lambda_p\}$. From the spectral mapping theorem it (see [DUN I], Ch. VII, § 3, Th. 11) follows that

$$\sigma(f(A_0)) = \{f(\lambda_1), \ldots, f(\lambda_p)\}.$$  

Since $f \in F(A_0)$ there are disjoint open sets $\Omega_1, \ldots, \Omega_p$ in $\mathbb{C}$ such that $\lambda_i \in \Omega_i$ $i = 1, \ldots, p$ and $f$ is locally holomorphic on $\bigcup_{i=1}^{p} \Omega_i$. Since $f'(\lambda_i) \neq 0$ and $f$ is 1-1, it follows from the inverse function theorem for holomorphic functions that we can also find disjoint open sets $\omega_1, \ldots, \omega_p$ such that $f(\lambda_i) \in \omega_i$ and $f^{-1}: \bigcup_{i=1}^{p} \omega_i \rightarrow \bigcup_{i=1}^{p} \Omega_i$ is locally holomorphic and satisfies

$$(f^{-1} \circ f)(z) = z \text{ if } z \in \Omega_1 \cup \ldots \cup \Omega_p.$$  

(One could also use the Bührman Lagrange theorem - applied p-times - to prove this.)

Hence $f^{-1} \in F(f(A_0))$ because $f^{-1}$ is holomorphic on an open neighbourhood of $\sigma(f(A_0))$. 

Now let $B(\mu)$ be any deformation of $f(A_0)$ with base $\Gamma$. If $\mu$ is small we have $f^{-1} \in F(B(\mu))$ and then $f^{-1}(B(\mu))$ is well defined and is a holomorphic function of $\mu$ with $f^{-1}(B(0)) = A_0$. From the versality of $A$ it follows that there is a deformation $C(\mu)$ of the identity matrix and a map $\varphi: \Gamma \to \Lambda$ with $\varphi(0) = 0$ such that

$$f^{-1}(B(\mu)) = C(\mu)A(\varphi(\mu))C^{-1}(\mu).$$

Applying $f$ to both sides we obtain

$$B(\mu) = C(\mu)f(A(\varphi(\mu)))C^{-1}(\mu)$$

and therefore $f(A(\lambda))$ is a versal deformation of $f(A_0)$ with base $\Lambda$. (If $A = CBC^{-1}$ then $f \in F(A)$ if $f \in F(B)$; $f(A) = Cf(B)C^{-1}$ in that case).

5.2. Remark. Condition 5.1.1 and 5.1.2 are both necessary. Take $D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ and $f_1(z) = (z - 1)(z - 2)$. Then $D$ has a 2-parameter versal deformation, but $f_1(D) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and therefore any versal deformation of $f_1(D)$ depends at least on 4-parameters. This proves that condition 5.1.1 is necessary. Taking

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and $f_2(z) = z^2$

we see that condition 5.1.2 is also necessary.
II. On Orbits and Centralizers of Operators

§ 0. Introduction

It is our aim to study deformations of Hilbert-Schmidt operators on an infinite dimensional separable Hilbert space $H$. The reader of chapter I may expect that the orbit and the centralizer of an operator must be studied in some detail first.

Only a part of the results seems to be new, many of them are quite standard (e.g. § 4, § 7).

We conclude this chapter with a heuristic approach to the topic of chapter III. The appendix is devoted to some isolated results on projectors in Hilbert space.

§ 1. The orbit and the centralizer in a Banach algebra

The definitions in this section are generalisations of the corresponding definitions in Chapter I, § 3. Let $B$ be a complex Banach algebra with identity $e$ (see [LAR]). The group $G$ of non-singular elements in $B$, is open in $B$ and contains $e$.

1.1. Remark. If $B = \mathcal{L}(H)$ then the set $G$ is connected, even if $H$ is infinite dimensional (see [KUI]).

1.2. Definition. If $a \in B$ is fixed we define the map $\alpha_a : G \to B$ by

$$\alpha_a(g) := g a g^{-1}.$$ 

$\alpha_a(G)$ is called the orbit of $a \in B$ under the action of the group $G$.

$\alpha_a$ is a $C^\infty$-map (the proof is similar to the proof of lemma 5.5 of this chapter) and the derivative at the identity $e : D_{\alpha_a} = : A_{\alpha_a}$ is a linear map from $B$ into $B$. (Since $G$ is open in $B$, the tangent space at the identity $T_e G$ is $B$ itself). Note that, in contrast with $\alpha_a$, the map $A_{\alpha_a}$ can also be defined in a Banach algebra without identity by putting $A_{\alpha_a}(g) := g a - a g$ for $g \in B$. Clearly $A_{\alpha_a}$ is a bounded linear operator whose norm in $\mathcal{L}(B)$ does not exceed $\|a\|$. 
1.3. Definition. \( \text{Ker}(\text{Ad}_a) \) is a closed subalgebra of \( B \) which will be called the centralizer of \( a \) in \( B \), notation \( Z_B(a) \). \( Z_B(a) \) consists of all elements in \( B \) which commute with \( a \). If no ambiguity is caused we sometimes write \( Z(a) \) instead of \( Z_B(a) \).

If \( B = \mathbb{F}^{n \times n} \) the linear manifold \( \text{Ad}_a(\mathbb{F}^{n \times n}) \) is necessarily closed and it is the tangent space to the orbit of \( a \). However, if \( B \) is infinite dimensional, for example \( B = \mathcal{L}(H) \), \( \text{Ad}_a(B) \) is not necessarily closed (see § 6). We shall only consider the special cases \( B = \mathcal{L}(H) \) and \( B = \mathcal{B} \).

§ 2. The centralizer of an operator in \( \mathcal{L}(H) \)

The main result of this section is our theorem 2.5 which states that for every operator \( A \in \mathcal{L}(H) \) the centralizer is infinite dimensional. As a preparation we start with some well known facts about minimal polynomials of operators.

2.1. Lemma. If \( A \in \mathcal{L}(H) \) and \( \psi \) is a polynomial with complex coefficients of degree \( n \geq 1 \) such that \( \psi(A) = 0 \), then there is a unique polynomial \( \varphi_0 \) of degree \( k \geq 1 \) such that

2.1.1. \( \varphi_0(A) = 0 \).
2.1.2. There is no polynomial with \( 1 \leq \text{degree} < k \) that annihilates \( A \).
2.1.3. The coefficient of \( z^k \) in \( \varphi_0 \) equals 1.

Proof. Suppose \( \psi \) annihilates \( A \). Obviously there is a polynomial \( \varphi \) of minimal degree \( k \geq 1 \) such that \( \varphi(A) = 0 \). Multiply \( \varphi \) by a complex constant \( \neq 0 \) such that in the resulting polynomial \( \varphi_0 \), the coefficient of \( z^k \) equals 1.

Now \( \varphi_0 \) clearly satisfies 2.1.1, 2.1.2 and 2.1.3.

The only thing left to prove is the uniqueness. Suppose \( \varphi_1 \) is a polynomial of degree \( k \) such that 2.1.1, 2.1.2 and 2.1.3 are satisfied. Then \( \varphi_0 - \varphi_1 \) still annihilates \( A \) and degree \( (\varphi_0 - \varphi_1) \leq k - 1 \). Since \( k \) is minimal it follows that \( \varphi_0 - \varphi_1 = 0 \) and hence \( \varphi_0 = \varphi_1 \).

\( \varphi_0 \) is called the minimal polynomial of the operator \( A \). Unlike finite matrices, most operators on \( H \) do not have a minimal polynomial.

This follows from:
2.2. Lemma. If $A \in \mathcal{L}(H)$ and has minimal polynomial $\phi_0$, then the spectrum of $A$, $\sigma(A)$, consists exactly of the zero's of $\phi_0$.

Proof. Write $\phi_0(z) = \prod_{i=1}^{p} (z - \lambda_i)^{h_i}$ with $h_i \geq 1$; $\lambda_i$'s complex and distinct.

From the spectral mapping theorem (see [DUN I], Ch. VII, § 3, Th. 11) it follows that

$$\phi_0(\sigma(A)) = \sigma(\phi_0(A)) = \{0\}.$$ 

Hence $\sigma(A) \subset \{\lambda_1, \ldots, \lambda_p\}$.

On the other hand if $1 \leq i \leq p$ the operator $A - \lambda_i I$ must be singular, since if $A - \lambda_i I$ is non-singular, $\psi_j := \prod_{j \neq i} (z - \lambda_j)^{h_j}$ still annihilates $A$ which contradicts the minimality of $\phi_0$.

Hence $\lambda_i \in \sigma(A)$, which completes the proof.

2.3. Corollary. If $A$ is quasinilpotent (that is $\sigma(A) = \{0\}$) and has an annihilating polynomial, then $A$ is nilpotent.

Lemma 2.2 and corollary 2.3 enable us to prove that for every $A \in \mathcal{L}(H)$ the dimension of the centralizer is infinite. We shall first prove this for a nilpotent operator.

2.4. Lemma. If $A \in \mathcal{L}(H)$ is nilpotent then $\dim Z(A) = \infty$.

Proof. Let $p \in \mathbb{N}$ be the smallest number for which $A^p = 0$, and define

$$N_j := \ker(A^j); \quad j = 1, 2, \ldots, p.$$ 

Every $N_j$ is a subspace of $H$ and $N_1 \subset N_2 \subset \ldots \subset N_p = H$.

It is easily seen that $\dim(N_j) = \infty; j = 1, 2, \ldots, p$. For, if $\dim(N_1)$ is finite then it follows that $\dim(N_p)$ is finite but this contradicts $N_p = H$. Since $N_{j-1} \neq N_j$ for $j = 2, \ldots, p$ there are non-trivial subspaces $M_1, \ldots, M_p \subset H$ such that

$$N_j = M_1 \oplus \ldots \oplus M_j; \quad j = 1, 2, \ldots, p$$

and

$$A(M_1) = \{0\} \text{ and } A(M_j) \subset M_{j-1} \text{ for } j \geq 2.$$ 

We now define a subspace $M \subset \mathcal{L}(H)$ as follows:

$$C \subset \mathcal{L}(H),$$

$$C \subset M \iff \begin{cases} C = 0 \text{ on } N_{p-1}, \\ C(M_p) \subset N_1. \end{cases}$$
\[ M \text{ is the intersection of two closed subsets of } \mathcal{L}(H) \text{ and therefore is closed in } \mathcal{L}(H). \text{ Since } M \text{ is non-trivial (this follows from the minimality of } p) \text{ and } N_1 \text{ has infinite dimension, } M \text{ is infinite dimensional. Each operator in } M \text{ commutes with } A. \text{ If } C \subset M \text{ we have } AC = CA = 0. \text{ To prove this write } x = x_1 + x_2 + \ldots + x_p \text{ with } x_j \in M_j \text{ and compute } ACx \text{ and } CAx. \text{ Hence } M \subset Z(A) \text{ and therefore } \dim Z(A) = \infty. \]

2.5. Theorem. For every \( A \in \mathcal{L}(H) \): \( \dim Z(A) = \infty. \)

Proof. Let \( A \in \mathcal{L}(H) \) and suppose \( \dim Z(A) \) is finite. Since \( Z(A) \) contains all powers \( A^k \) of \( A \) (\( k = 0, 1, 2, \ldots \)) we can find \( n \geq 1 \) and \( \alpha_0, \ldots, \alpha_n \in \mathbb{C} \) such that

\[ A^n + \alpha_{n-1} A^{n-1} + \ldots + \alpha_0 I = 0. \]

Hence, by lemma 2.1, \( A \) has a minimal polynomial \( \varphi_0 \) of degree \( \leq n. \) From lemma 2.2 it follows that \( \sigma(A) \) consists of the zero's of \( \varphi_0 \) and hence is a finite set say \( \sigma(A) = \{ \lambda_1, \ldots, \lambda_p \}. \)

Define the operators \( E_j, j = 1, \ldots, p \) by

2.5.1. \[ E_j := \frac{1}{2\pi i} \oint_{(\lambda_j)} R(z, A) \, dz \]

where \((\lambda_j)\) is a small circle centered at \( \lambda_j. \) Then \( E_j \) is the projection operator on the invariant subspace \( X_j = E_j(H) \) corresponding to the spectral point \( \lambda_j. \) The space \( H \) is the direct sum of \( p \) subspaces invariant under \( A: \)

\[ H = X_1 \oplus X_2 \oplus \ldots \oplus X_p. \]

Since \( H \) is infinite dimensional there is at least one \( j \) with \( \dim(X_j) = \infty. \)

Let \( A_j \) denote the restriction of \( A \) to the invariant subspace \( X_j, j = 1, 2, \ldots, p. \)

Every \( x \in H \) has a unique representation \( x = x_1 + \ldots + x_p \) with \( x_i \in X_i. \)

If \( P \) is a polynomial we have

2.5.2. \[ P(A)x = P(A_1)x_1 + \ldots + P(A_p)x_p \]

(because \( A^\ell x = A_1^\ell x_1 + \ldots + A_p^\ell x_p \) for \( \ell \in \mathbb{N} \)). Taking \( P = \varphi_0 \) in 2.5.2 it follows that every \( A_j \) has a minimal polynomial of degree less than degree(\( \varphi_0 \)).
From now on we fix $j$ such that $\dim(X_j) = \infty$ and we shall prove that
$\dim_L(X_j)Z(A_j) = \infty$. The spectrum $\sigma(A_j)$ consists of exactly one point $\lambda_j$ (see [DUN II], Ch. VII, § 3, Th. 20). It is no loss of generality assuming $\lambda_j = 0$, because $Z_L(X_j)(A_j - \lambda_j I_j) = Z_L(X_j)(A_j)$, where $I_j$ is the restriction of $I$ to $X_j$. According to this assumption the operator $A_j$ is quasinilpotent. Since $A_j$ also has a minimal polynomial it follows from corollary 2.3 that $A_j$ is nilpotent and hence by lemma 2.4 $\dim Z_L(X_j)(A_j) = \infty$.

If $B \in Z_L(X_j)(A_j)$ then $BE_j \in Z_L(H)(A)$, since for arbitrary $x \in H$

$$BE_j Ax = BA BE_j x = BA_j E_j x = A_j BE_j x = ABE_j x.$$ 

This contradicts the assumption that $\dim Z_L(H)(A)$ is finite. Hence $\dim Z_L(H)(A) = \infty$. 

2.6. Remark. Obviously, for every $A \in L(H)$ we have $Z(A^*) = (Z(A))^*$. If $A$ is normal then $Z(A^*) = Z(A)$. The last result is a theorem of Fuglede (see [FUG]), which has a short and elegant proof in [ROS].

§ 3. The centralizer of a normal compact operator

In general it is difficult to compute the centralizer of an operator (in the finite dimensional case, for matrices, the computation can be found in [GAN], Ch. VIII, § 2). For a certain class of operators, however, it is rather easy. This class includes the normal compact operators. We shall describe the centralizer of a normal compact operator and prove that it splits in $L(H)$. We first quote some standard results on normal compact operators.

Suppose $A$ is normal and compact. Let $\lambda_1, \lambda_2, \ldots$ be an enumeration of $\sigma(A) \setminus \{0\}$ such that $|\lambda_1| \geq |\lambda_2| \geq \ldots$. Define $X_0 := \ker(A)$ and $X_j = E_j(H)$; $j \geq 1$, where $E_j$ is the projector defined as in 2.5.1:

$$E_j = \frac{1}{2\pi I} \int_{\{\lambda_j\}} R(\zeta, A) d\zeta.$$ 

Since $A$ is normal the projections $E_j$ are orthogonal and therefore self-adjoint (see [DUN I], Ch. VI, § 3). It is well known that the space $H$ is the direct sum of the orthogonal eigenspaces $X_j$ which reduce $A$:

$$H = X_0 \oplus X_1 \oplus \ldots.$$
and that the operator $A$ has the spectral decomposition

$$A = \sum_{j=1}^{\infty} \lambda_j E_j + 0.E_0$$

where $E_0$ is the projector on $\text{Ker}(A)$. The subspaces $X_j$ are mutually orthogonal and

$$\text{Ker}(A) = \bigcap_{i=1}^{\infty} X_i^\perp$$

($X_i^\perp$ is the orthogonal complement of $X_i$).

For arbitrary $x \in \text{Ker}(A)$ and $j \in \mathbb{N}$ we have $x = x_1 + x_2$ with $x_1 \in X_j$ and $x_2 \in X_j^\perp$. Hence

$$0 = Ax = Ax_1 + Ax_2 = \lambda_j x_1 + Ax_2.$$

Since $Ax_2 \in X_j^\perp$ ($A$ is normal) we have $x_1 = 0$ because $\lambda_j \neq 0$. Hence $x \in X_j^\perp$.

$$\text{Ker}(A) = \bigcap_{i=1}^{\infty} X_i^\perp.$$ 

But, since $H = X_0 \oplus X_1 \oplus \ldots$ we have

$$X_0 = \text{Ker}(A) = \bigcap_{i=1}^{\infty} X_i^\perp.$$ 

3.1. Lemma. The centralizer of the normal compact operator $A$ is the subspace

$$\{ C \in \mathcal{L}(H) \mid CE_j = E_j C; j = 0,1,2,\ldots \}.$$ 

This lemma is a direct corollary of the preceding results of this chapter. It is a special case of a result in [HAL II].

Lemma 3.1 enables us to prove our theorem 3.2.

3.2. Theorem. If $A$ is normal and compact in $\mathcal{L}(H)$, then $Z(A)$ splits in $\mathcal{L}(H)$.

Proof. To prove this we give a closed complement of $Z(A)$ in $\mathcal{L}(H)$. Define $V \subset \mathcal{L}(H)$ by

$$V := \{ D \in \mathcal{L}(H) \mid D(X_j) \subset X_j^\perp; j = 0,1,2,\ldots \}.$$
then we shall show that $V$ satisfies

3.2.1. $V$ is a subspace of $\mathcal{L}(H)$.

3.2.2. $V \cap Z(A) = \{0\}$.

3.2.3. $V \oplus Z(A) = \mathcal{L}(H)$.

ad 3.2.1. It is obvious that $V$ is a linear space. Suppose $(D_n)_{n \in \mathbb{N}}$ is a sequence in $V$ with $\lim_{n \to \infty} D_n = D \in \mathcal{L}(H)$ in the uniform topology. If $x \in X_j$ and $y \in X_j$ we have

$$
(Dx, y) = \lim_{n \to \infty} (D_n x, y) = 0 .
$$

Hence $Dx \in X_j^\perp$ and therefore $D(X_j) \subseteq X_j^\perp$ and hence $V$ is closed.

ad 3.2.2. Suppose $T \in V \cap Z(A)$. Choose $x \in X_j$ then $T \in V$ implies $Tx \in X_j^\perp$ and $T \in Z(A)$ implies $Tx \in X_j$ (lemma 3.1) hence $Tx = 0$. Hence $T = 0$.

ad 3.2.3. Let $T \in \mathcal{L}(H)$. For $h \in H$ we define

$$
C h := \sum_{k=0}^{\infty} (E_k T E_k) h
$$

where the $E_k$'s are the projections in the spectral decomposition of $A$. This definition makes sense because

$$
\| \sum_{k=0}^{n} E_k T E_k h \|^2 = \sum_{k=0}^{n} \| E_k T E_k h \|^2 \leq \sum_{k=0}^{n} \| E_k h \|^2 \leq \| T \|^2 \sum_{k=0}^{n} \| E_k h \|^2 \leq \| T \|^2 \| h \|^2 .
$$

The sequence $h_n := \sum_{k=0}^{n} (E_k T E_k) h$, $n \in \mathbb{N}$, is a Cauchy-sequence in $H$ and therefore convergent with limit $C h \in H$. Clearly $C$ is linear and its norm does not exceed $\| T \|$. Hence, $C \in \mathcal{L}(H)$. We now prove $C \in Z(A)$. For $x \in H$ we have

$$
C E_j x = \sum_{k=0}^{\infty} E_k T E_j E_j x = E_j T E_j x
$$

and

$$
E_j C x = \sum_{k=0}^{\infty} E_j E_k T E_j x = E_j T E_j x .
$$

Hence by lemma 3.1 $C \in Z(A)$. 

Define $D := T - C$ then $D \in \mathcal{L}(H)$ and $T = D + C$. $D \in V$ because for $x \in X_j$ and $y \in X_j$ we have

$$(Dx, y) = (Tx - Cx, y) = (Tx - E_j Tx, y) =
= (Tx, y) - (Tx, E_j^* y) = (Tx, y) - (Tx, E_j y) = 0.$$ (Note that $E_j^* = E_j$).

Hence $DX_j \subset X_j^*$ and therefore $D \in V$ which completes the proof.

§ 4. Commutators of operators

If $A \in \mathcal{L}(H)$ the set $\text{Ad}_A(\mathcal{L}(H))$ consists of commutators of the form $CA - AC$ where $C \in \mathcal{L}(H)$. If the space $H$ is infinite dimensional, it is an interesting question whether a given operator can be written as a commutator or not. Commutators have been investigated by Halmos, Putnam, Brown and Pearcy in [HAL II], [PUT] and [BRO]. The most important result in this direction is that an operator $A \in \mathcal{L}(H)$ can be written as a commutator iff $A \neq \lambda I + C$, $\lambda \neq 0$ and $C$ compact (see [BRO]).

In § 8 we shall use a theorem which can be deduced from the Kleinecke-Shirokov theorem.

4.1. Theorem (Kleinecke-Shirokov). If $C = PQ - QP$ and $CP = PC$ then $\sigma(C) = \{0\}$.

Proof. See [HAL I], problem 184.

4.2. Theorem (Putnam). If $C = PQ - QP; CP = PC$ and $P$ is normal then $C = 0$.

Proof. See [PUT].

4.3. Corollary. If $A \in \mathcal{L}(H)$ is normal then

$$Z(A) \cap \text{Ran } \text{Ad}_A = \{0\}.$$ 

§ 5. Hilbert-Schmidt operators

In this section we quote some standard results from the theory of Hilbert-Schmidt operators. Furthermore we shall prove some new theorems, which will be useful later on. Our theorem 5.7 is a straightforward generalization of lemma 4.2 in chapter I.
5.1. Definition. An operator $A \in \mathcal{L}(H)$ is a Hilbert-Schmidt operator iff there is an orthonormal basis $(e_n)_{n \in \mathbb{N}}$ (which is fixed from now on) for the space $H$ such that

$$\|A\| := \left( \sum_{k=1}^{\infty} \|Ae_k\|^2 \right)^{\frac{1}{2}} < \infty.$$ 

The sum, which is independent of the chosen basis, is called the double-norm of $A$. The set of all Hilbert-Schmidt operators on $H$ is denoted by $\text{HS}$. For properties of Hilbert-Schmidt operators see [SCH], Ch. II or [DUN II], Ch. XI, § 6.

5.2. Remark. There is a one-to-one correspondence between the set $\text{HS}$ and the set of all 2-sided infinite complex matrices $\{a_{ij}\}$ with

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}|^2 < \infty$$

(with respect to the basis $(e_n)_{n \in \mathbb{N}}$).

If $A \in \text{HS}$ the corresponding matrix $\{a_{ij}\}$ has entries $a_{ij} = (Ae_j, e_i)$. Since $A \in \text{HS}$, $(\sum_{i=1}^{\infty} |a_{ij}|^2)\frac{1}{2}$ converges and equals $\|A\|$. To the matrix $\{a_{ij}\}$ corresponds the Hilbert-Schmidt operator $A$ defined by $Ae_j = \sum_{i=1}^{\infty} a_{ij}e_i$. The correspondence from operators to matrices has the usual algebraic properties.

5.3. Definition. In $\text{HS}$ we define an innerproduct as follows

$$(A,B) := \sum_{n=1}^{\infty} (Ae_n, Be_n).$$

The innerproduct is independent of the chosen basis and makes $\text{HS}$ into a Hilbert space (see [SCH], Ch. II). The innerproduct has three useful properties:

5.3.1. $(A,A) = \|A\|^2$,

5.3.2. $(A,B) = (A^*, B^*)$,

5.3.3. $(XA,B) = (A,X^*B)$,

for all $A$, $B$ and $X \in \text{HS}$. From 5.3.2 and 5.3.3 we can deduce

$$(AX,B) = (X^*A^*, B^*) = (A^*, XB^*) = (A, BX^*).$$
The innerproduct is a generalization of the innerproduct defined in Chapter I, definition 4.1. For, if \( A \in \mathcal{H} \) then \( AB^* \) is in the trace class (see [SCH]) and

\[
(A, B) = \text{trace}(AB^*) = \sum_{n=1}^{\infty} (Ae_n, Be_n).
\]

5.4. Remark. Since \( \mathcal{H} \) provided with double-norm is a B-algebra without identity (with involution) the map \( \alpha_A^* \), in the sense of definition 1.2, is not defined (it is easily seen that there are no invertible elements in \( \mathcal{H} \)). However, if \( A \in \mathcal{H} \), the map \( \alpha_A^* : \mathbb{G} \to \mathcal{L}(\mathcal{H}) \), defined in definition 1.2, can be considered as a map from \( \mathbb{G} \) into \( \mathcal{H} \). This follows from the fact that \( \mathcal{H} \) is a two-sided ideal in \( \mathcal{L}(\mathcal{H}) \).

5.5. Lemma. Let \( A \in \mathcal{H} \). The map \( \alpha_A^* : \mathbb{G} \to \mathcal{H} \) defined by \( \alpha_A^*(g) := gAg^{-1} \) is a \( C^0 \) map considered as a map from \( (\mathbb{G}, \| \|) \) into \( (\mathcal{H}, \| \|) \). The derivative at the identity operator \( I \in \mathcal{L}(\mathcal{H}) \) is the map \( \text{Ad}_A : \mathcal{L}(\mathcal{H}) \to \mathcal{H} \) which maps \( g \) into \( [g, A] \).

Proof. We shall first prove that \( \alpha_A^* \) is differentiable at \( I \) with derivative \( \text{Ad}_A^* \). If \( \|h\| < 1 \) we have

\[
\alpha_A^*(I + h) = (I + h)A(I + h)^{-1} = (I + h)A \sum_{n=0}^{\infty} (-1)^n h^n =
\]

\[
= A + hA - Ah + A \sum_{n=2}^{\infty} (-1)^n h^n + hA \sum_{n=1}^{\infty} (-1)^n h^n.
\]

Hence

\[
\|\alpha_A^*(I + h) - \alpha_A^*(I) - \text{Ad}_A^*(h)\| = \|A \sum_{n=2}^{\infty} (-1)^n h^n + hA \sum_{n=1}^{\infty} (-1)^n h^n\| \leq 2\|A\| \|h\|^2 \text{ if } \|h\| < \frac{1}{2}.
\]

Hence

\[
\lim_{\|h\| \to 0} \frac{\|\alpha_A^*(I + h) - \alpha_A^*(I) - \text{Ad}_A^*(h)\|}{\|h\|} = 0,
\]

and therefore \( \alpha_A^* \) is differentiable at \( I \) with derivative \( \text{Ad}_A^* \). With an analogous computation it can be shown that \( \alpha_A^* \) is differentiable at any point \( g_0 \in \mathbb{G} \) with derivative

\[
(D_{g_0} \alpha_A^*)(h) = g_0^{-1} [g_0^{-1} h, A] g_0^{-1}.
\]

It is obvious that \( g_0 \to D_{g_0} \alpha_A^* \) is \( C^0 \) and hence \( \alpha_A^* \) is \( C^0 \). \( \square \)
The first part of the lemma has a shorter proof as follows. Note that the map $g \mapsto g^{-1}$ is a $C^\infty$-map from $G$ onto $G$ and the mappings $g \mapsto gA$ and $g \mapsto Ag$ are $C^\infty$-mappings from $(G,\|\|)$ into $(HS,\|\|)$. Hence $\alpha_A$ is a $C^\infty$-map.

The map $\Ad_A$ can also be considered as a map from $HS$ into $HS$ and then $\Ker \Ad_A$ is a $\|\|$-(closed) subspace of $HS$ which is the centralizer of $A$ in $HS$: $Z_{HS}(A)$. Thus by $Z_{HS}(A)$ is always meant the set

$$\{C \in HS \mid CA = AC\}.$$ 

5.6. Remark. If $A \in HS$, the centralizer of $A$ in $HS$, $Z_{HS}(A)$, is infinite dimensional. To prove this consider $A$ as an operator in $L(H)$ and copy the proof of theorem 2.5. The only modification that has to be made is the following: replace the subspace $M \subseteq L(H)$ defined in 2.4.1 by a subspace $M' \subseteq HS$ having exactly the same properties as $M$ except that $M'$ consists only of Hilbert-Schmidt operators.

We now prove a generalization of lemma 4.2 of Chapter I.

5.7. Theorem. Let $A \in HS$. The orthogonal complement of $\Ad_A(HS)$ in $HS$ is the centralizer of the adjoint of $A$: $Z_{HS}(A^*)$.

Proof. Let $X \in (\Ad_A(HS))^\perp$ then for all $Y \in HS$ we have $(X,\Ad_A(Y)) = 0$ and hence

$$(X,AY) - (X,YA) = 0 .$$

If we use 5.3.2 and 5.3.3 we obtain

$$(AX - XA^*,Y) = 0 \quad \text{for } Y \in HS$$

and hence $X \in Z_{HS}(A^*)$.

If $X \in Z_{HS}(A^*)$ the proof goes the other way around. Hence $(\Ad_A(HS))^\perp = Z_{HS}(A^*)$. 

5.8. Remark. The proof of theorem 5.7 can be formulated in another way. Note that the map $\Ad_A^*$ is the adjoint of $\Ad_A$ in $L(HS)$. (Just as in the given proof this is a direct consequence of 5.3.2.) Hence

$$(\Ran(\Ad_A))^\perp = (\Ad_A(HS))^\perp = \Ker(\Ad_A^*) = Z_{HS}(A^*) .$$
Theorem 5.7 plays an important role in the construction of weakly versal de­formations of Hilbert-Schmidt operators (see Ch. III, § 5).

The map $\text{Ad}_A$ can also be considered as a map from $\mathcal{L}(H)$ into $\text{HS}$. Obviously, we have

$$\text{Ad}_A (\mathcal{L}(H)) \supset \text{Ad}_A (\text{HS}).$$

In the next theorem we shall prove that $\text{Ad}_A (\text{HS})$ is double-norm dense in $\text{Ad}_A (\mathcal{L}(H))$.

5.9. Theorem. Let $A \in \text{HS}$. Then

$$\overline{\text{Ad}_A (\text{HS})} = \overline{\text{Ad}_A (\mathcal{L}(H))}.$$  

(The double bar denotes the double-norm closure.)

Proof. Suppose $A$ has the matrix \{a_{ij}\}, with respect to the basis $(e_j)_{j \in \mathbb{N}}$, then $A^*$ has the matrix $\{\overline{a}_{ji}\}$. Hence, for $i \in \mathbb{N}$ we have

$$Ae_i = \sum_{k=1}^{\infty} a_{ki} e_k \quad \text{and} \quad A^* e_i = \sum_{k=1}^{\infty} \overline{a}_{ik} e_k.$$ 

If $B \in Z_{\text{HS}}(A^*)$ and $g \in \mathcal{L}(H)$ we have

$$(\text{Ad}_A(g), B) = \sum_{k=1}^{\infty} (gAe_k, Be_k) - \sum_{k=1}^{\infty} (Ae_k, Be_k) =$$

$$= \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} a_{jk} (ge_j, Be_k) - \sum_{k=1}^{\infty} (ge_k, A^* Be_k).$$

Since $BA^* = A^* B$ we obtain

$$(\text{Ad}_A(g), B) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} a_{jk} (ge_j, Be_k) - \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} a_{kj} (ge_k, Be_j).$$

We now prove that both sums are absolutely convergent. Both proofs are alike, so we give only one of them. Applying the Cauchy-Schwarz inequality to the first sum we find:
\[\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |a_{jk}| |(ge_j, Be_k)| \leq \left( \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |a_{jk}|^2 \right)^{1/2} \left( \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |(ge_j, Be_k)|^2 \right)^{1/2} = \|A\| \|g^* B\|.\]

For,
\[\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |(ge_j, Be_k)|^2 = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |(e_j, g^* Be_k)|^2 = \sum_{k=1}^{\infty} \|g^* Be_k\|^2 = \|g^* B\|^2.\]

Hence both sums are absolutely convergent and therefore we may change the order of summation. Hence
\[(\text{Ad}_A(g), B) = 0.\]

Since \(g \in \mathcal{L}(H)\) was arbitrary this proves \(B \in (\text{Ad}_A(\mathcal{L}(H)))^\perp\). Since \(B \in Z_{HS}(A^*)\) was arbitrary, we may conclude
\[Z_{HS}(A^*) \subseteq (\text{Ad}_A(\mathcal{L}(H)))^\perp.\]

In theorem 5.7 we already have proved that \(Z_{HS}(A^*) = (\text{Ad}_A(\mathcal{L}(H)))^\perp\). Hence
\[Z_{HS}(A^*) = (\text{Ad}_A(\mathcal{L}(H)))^\perp \supseteq (\text{Ad}_A(\mathcal{L}(H)))^\perp \supseteq Z_{HS}(A^*) ,\]
and therefore \((\text{Ad}_A(\mathcal{L}(H)))^\perp = (\text{Ad}_A(\mathcal{L}(H)))^\perp \supseteq Z_{HS}(A^*)\),
\[\text{Ad}_A(\mathcal{L}(H)) = \text{Ad}_A(\mathcal{L}(H)).\]

5.10. Remark. Theorem 5.9 shows that for our purpose, which will become clear in chapter III, we can disregard the difference between \(\text{Ad}_A(\mathcal{L}(H))\) and \(\text{Ad}_A(\mathcal{L}(H))\).

5.11. Remark. In many examples the linear manifolds \(\text{Ad}_A(\mathcal{L}(H))\) are not closed. We shall give two examples in the next section.

§ 6. Examples

In chapter I \(\mathcal{H}\) and \(\mathbb{C}^{n \times n}\) coincide, so \(\text{Ad}_A(\mathcal{H}) = \text{Ad}_A(\mathbb{C}^{n \times n})\) and these finite dimensional linear manifolds are necessarily closed. In the infinite dimensional case there are many examples in which \(\text{Ad}_A(\mathcal{H})\) and \(\text{Ad}_A(\mathcal{L}(H))\) are not closed. This fact forms an additional complication to the theory in chapter III.
6.1. **Remark.** If $A \in \text{HS}$ both $\text{Ad}_A(\text{HS})$ and $\text{Ad}_A(\mathcal{L}(H))$ are subsets of $\text{HS}$ satisfying:

$$\text{Ad}_A(\text{HS}) \subset \text{Ad}_A(\mathcal{L}(H))$$

and

$$\text{Ad}_A(\text{HS}) = \text{Ad}_A(\mathcal{L}(H))$$

(see theorem 5.9).

If $\text{Ad}_A(\text{HS})$ is $\|\|\|$-closed, $\text{Ad}_A(\mathcal{L}(H))$ must be $\|\|\|$-closed. On the other hand if $\text{Ad}_A(\mathcal{L}(H))$, considered as a subset of $\mathcal{L}(H)$, is $\|\|\|$-closed it is necessarily $\|\|\|$-closed, since every $\|\|\|$-closed set in $\text{HS}$ is $\|\|\|$-closed ($\|A\| \leq \|A\|$).

6.2. **Corollary.** If $\text{Ad}_A(\mathcal{L}(H))$ is not $\|\|\|$-closed then

$$\text{Ad}_A(\text{HS})$$

is not $\|\|\|$-closed.

$$\text{Ad}_A(\mathcal{L}(H))$$

is not $\|\|\|$-closed.

We shall give two examples of an operator $A \in \text{HS}$ for which $\text{Ad}_A(\text{HS})$ and $\text{Ad}_A(\mathcal{L}(H))$ are not closed (in $\|\|$ or $\|\|\|$). The first example deals with a diagonal operator, the second with a monotone $\ell_2$-shift. In the first example we shall give two different proofs to show that $\text{Ad}_A(\text{HS})$ is not $\|\|\|$-closed. In both examples we are able to compute the centralizers $Z_{\text{HS}}(A)$. The examples are described with respect to the basis $(e_n)_{n \in \mathbb{N}}$.

6.3. **Example.** Let $D$ be a diagonal operator in $\text{HS}$

$$D = \text{diag}\{\lambda_1, \lambda_2, \ldots\}$$

with $\lambda_i$'s distinct and \(\sum_{j=1}^{\infty} |\lambda_j|^2 < \infty\). It follows from an easy computation that $Z_{\mathcal{L}(H)}(D)$ consists of all diagonal operators in $\mathcal{L}(H)$ and therefore $Z_{\text{HS}}(D)$ consists of all diagonal Hilbert-Schmidt operators (with respect to the basis $(e_n)_{n \in \mathbb{N}}$). To compute $\text{Ad}_D(\text{HS})$ we use theorem 5.7, which implies $\text{Ad}_D(\text{HS}) = (Z_{\text{HS}}(D))^\perp$. (Note that $Z_{\text{HS}}(D^*) = Z_{\text{HS}}(D)$). Suppose $X \in (Z_{\text{HS}}(D))^\perp$ then for all diagonal operators $\Lambda \in \text{HS}$ we have $(\Lambda, X) = 0$ and therefore $\forall i \in \mathbb{N}$ $(Xe_i, e_i) = 0$. On the other hand if $\forall i \in \mathbb{N}$ $(Xe_i, e_i) = 0$ it follows that $X \in (Z_{\text{HS}}(D))^\perp$. Hence

$$\text{Ad}_D(\text{HS}) = \{X \in \text{HS} \mid \forall i \in \mathbb{N} \ (Xe_i, e_i) = 0\}.$$
We now show that there is an operator $F \in \overline{\text{Ad}_D(\mathcal{H})}\setminus \text{Ad}_D(\mathcal{H})$ and therefore $\text{Ad}_D(\mathcal{H})$ cannot be \(\|\|\)-closed in $\mathcal{H}$. Let $F \in \mathcal{H}$ be the shift operator defined by $F_i = \mu_i e_{i+1}$ with $\sum_{i=1}^{\infty} |\mu_i|^2 < \infty$. Obviously, $F \in \overline{\text{Ad}_D(\mathcal{H})}$. Suppose $F \in \text{Ad}_D(\mathcal{H})$ then for some $C \in \mathcal{H}$ we have $CD - DC = F$.

This implies

$$\forall i \in \mathbb{N} \forall j \in \mathbb{N} \quad ([C,D]e_i,e_j) = (F_i,e_j) .$$

Hence

$$\forall i \in \mathbb{N} \quad (Ce_i,e_{i+1}) = \frac{\mu_i}{\lambda_i - \lambda_{i+1}} .$$

Hence

6.3.1. $\|C\| \geq \|C\| \geq \|Ce_i\| = \left( \sum_{j=1}^{\infty} |(Ce_i,e_j)|^2 \right)^{1/2} \geq |(Ce_i,e_{i+1})| = \left| \frac{\mu_i}{\lambda_i - \lambda_{i+1}} \right| .$

Since $\lim_{i \to \infty} |\lambda_i - \lambda_{i+1}| = 0$ we can find a subsequence $(\lambda_{i_k})_{k \in \mathbb{N}}$ of $(\lambda_i)_{i \in \mathbb{N}}$ such that

$$\forall k \in \mathbb{N} \quad |\lambda_{i_k} - \lambda_{i_k+1}| \leq 2^{-k} .$$

We now make a special choice for the weights $(\mu_j)_{j \in \mathbb{N}}$ of the shift $F$. Take

$$\mu_{i_k} = 2^{-k}$$

$$\mu_i = 0 \text{ if } i \neq i_k \text{ for all } k .$$

Then $\sum_{i=1}^{\infty} |\mu_i|^2 < \infty$ and

$$\lim_{k \to \infty} \frac{|\lambda_{i_k} - \lambda_{i_k+1}|}{|\mu_{i_k}|} = \infty .$$

Hence, by 6.3.1, there is no operator $C \in \mathcal{H}$ (neither in $\mathcal{L}(\mathcal{H})$) such that $CD - DC = F$ and therefore $F \not\in \text{Ad}_D(\mathcal{H})$. Hence $\text{Ad}_D(\mathcal{H})$ is not \(\|\|\)-closed in $\mathcal{H}$. The same arguments prove that $\text{Ad}_D(\mathcal{L}(\mathcal{H}))$ is not \(\|\|\)-closed.

There is another way of proving that $\text{Ad}_D(\mathcal{L}(\mathcal{H}))$ is not \(\|\|\)-closed. Suppose $\text{Ad}_D(\mathcal{L}(\mathcal{H}))$ is closed in \(\|\|\). Define

$$V := \{X \in \mathcal{L}(\mathcal{H}) \mid (Xe_i,e_i) = 0, i \in \mathbb{N} \} .$$
It follows from the proof of theorem 3.2 that \( V \) is a closed complement of \( Z_L(H)(D) \) in \( L(H) \):
\[
V \oplus Z_L(H)(D) = L(H).
\]
Consider the restriction \( \Delta \) of the map \( Ad_D \) to the subspace \( V \). Then \( \Delta \) is a bounded linear operator from the Banach space \( (V, \| \cdot \|) \) onto the Banach space \( (Ad_D(L(H)), \| \cdot \|) \). (The norm of \( \Delta \) does not exceed \( 2\|D\| \)). \( \text{Ker}(\Delta) = \text{Ker}(Ad_D) \cap V = Z_L(H)(D) \cap V = \{0\} \) and hence \( \Delta \) is 1-1. By the closed graph theorem \( \Delta \) is invertible with bounded inverse, so there is a \( \delta > 0 \) such that
\[
6.3.2. \quad \forall x \in V \quad \| \Delta(x) \| \geq \delta \| x \|.
\]
Define the sequence \( (x_n)_{n \in \mathbb{N}} \subset V \) as follows:
\[
x_n e_j = 0 \text{ if } j \neq n + 1; \quad x_{n+1} = e_n
\]
extend \( x_n \) linearly to the whole space \( H \). Now \( \| x_n \| = 1 \) and
\[
\lim_{n \to \infty} \| \Delta(x_n) \| = \lim_{n \to \infty} \| \lambda_n - \lambda_{n+1} \| = 0.
\]
This contradicts 6.3.2 and therefore \( Ad_D(L(H)) \) is not \( \| \cdot \| \)-closed.

6.4. Remark. If \( A \in HS \) is normal (not necessarily diagonal) the same arguments (theorem 3.2 and the closed graph theorem) show that \( Ad_A(L(H)) \) is not closed.

6.5. Example. Let \( (\alpha_n)_{n \in \mathbb{N}} \) be a sequence in \( \mathbb{R} \) such that

6.5.1. \( \alpha_1 > \alpha_2 > \ldots > 0 \).

6.5.2. \( \sum_{i=1}^{\infty} \alpha_i^2 < \infty \).

Let the operator \( U \in HS \) be defined by
\[
U e_n = \alpha_{n} e_{n+1} \quad n \in \mathbb{N}.
\]
\( U \) is called a monotone \( \ell_2 \)-shift with weights \( (\alpha_n)_{n \in \mathbb{N}} \). \( U^* \) satisfies
\[
U^* e_1 = 0
\]
and
\[
U^* e_n = \alpha_{n-1} e_{n-1}, \quad n \geq 2.
\]
Note that $U^*e_1 = 0$ and $U^*u_1 = \alpha_1 u^*_2 = \alpha_1^2 e_1$ and therefore $U$ is not normal. We first compute $Z_H(U^*)$. Note that the only non-trivial invariant subspaces under $U^*$ are

$$M_n := \text{span}(e_1, \ldots, e_n); \quad n \in \mathbb{N}$$

(see [HAL I], problem 151).

Suppose $U^*R = RU^*$ then

$$U^*R(M_n) = RU^*(M_n) \subset R(M_n).$$

Hence $R(M_n)$ is an invariant subspace under $U^*$, and therefore there is an integer $k \in \mathbb{N}$ such that

$$R(M_n) = M_k \quad \text{(or } R(M_n) = \{0\}).$$

Clearly $\dim(R(M_n)) \leq n$ and hence $k \leq n$. This proves that $M_n$ is an invariant subspace under $R$. Hence every $R \in Z_H(U^*)$ must be uppertriangular (with respect to the basis $(e_n \mid n \in \mathbb{N})$). We now make a special choice for the weights $\alpha_n$. Let $\alpha_n$ be given by $\alpha_n = \alpha^n$ where $0 < \alpha < 1$.

Further computation shows that $R \in Z_H(U^*)$ iff

6.5.3. $R$ is uppertriangular

6.5.4. $R_{i,i+k} = \beta_k \alpha^{k(i-1)}; \quad i \in \mathbb{N}, \quad k \in \mathbb{N} \cup \{0\}$

6.5.5. $\sum_{k=0}^{\infty} \sum_{i=1}^{k} |R_{i,i+k}|^2 < \infty$ where $\beta_k \in \mathbb{C}$ for $k = 0, 1, 2, \ldots$.

Condition 6.5.5 implies $\beta_0 = 0$. Combining 6.5.4, 6.5.5 and $\beta_0 = 0$ we obtain

$$\sum_{k=1}^{\infty} \sum_{i=1}^{k} |R_{i,i+k}|^2 = \sum_{k=1}^{\infty} |\beta_k|^2 \sum_{i=1}^{\infty} \alpha^{2k(i-1)} = \sum_{k=1}^{\infty} \frac{|\beta_k|^2}{1 - \alpha^{2k}} < \infty.$$

Note that for all $k \in \mathbb{N}$ we have

$$|\beta_k|^2 \leq \frac{|\beta_k|^2}{1 - \alpha^{2k}} \leq \frac{|\beta_k|^2}{1 - \alpha^2}.$$

Hence

$$\sum_{k=1}^{\infty} \frac{|\beta_k|^2}{1 - \alpha^{2k}} < \infty \quad \text{iff} \quad \sum_{k=1}^{\infty} |\beta_k|^2 < \infty$$

and therefore $R \in Z_H(U^*)$ iff
6.5.6. $R$ is uppertriangular

6.5.7. $R_{i,i+k} = \beta_k a^{k(i-1)}$, $i \in \mathbb{N}$ and $k \in \mathbb{N} \cup \{0\}$.

6.5.8. $\beta_0 = 0$ and $\sum_{k=1}^{\infty} |\beta_k|^2 < \infty$.

The double norm of an operator $R \in Z_{HS}(U^*)$ is given by

$$\|R\|_d = \left\{ \sum_{k=1}^{\infty} \frac{|\beta_k|^2}{1 - \alpha^{2k}} \right\}^{1/2}$$

and $R$ has the matrix

$$
\begin{pmatrix}
0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \ldots \\
0 & \alpha \beta_1 & \alpha^2 \beta_2 & \alpha^3 \beta_3 & \ldots \\
0 & \alpha^2 \beta_1 & \alpha^4 \beta_2 & \ldots \\
0 & \alpha^3 \beta_1 & \ldots \\
0 & \\
\end{pmatrix}
$$

We are now able to compute $Z_{\mathcal{L}(H)}(U^*)$. Condition 6.5.6 and 6.5.7 still hold if $R \in Z_{\mathcal{L}(H)}(U^*)$, and if $R \in \mathcal{L}(H)$ we have

$$\sum_{k=1}^{\infty} |R_{1k}|^2 = \|R^* e_1\| = \sum_{k=0}^{\infty} |\beta_k|^2$$

so also $R \in Z_{\mathcal{L}(H)}(U^*)$ implies

$$\sum_{k=0}^{\infty} |\beta_k|^2 < \infty.$$ 

The only difference between $Z_{\mathcal{L}(H)}(U^*)$ and $Z_{HS}(U^*)$ is the condition $\beta_0 = 0$. Hence

$$Z_{\mathcal{L}(H)}(U^*) = \{ \lambda I + R \mid \lambda \in \mathbb{C}, R \in Z_{HS}(U^*) \}$$

and therefore

$$Z_{\mathcal{L}(H)}(U) = \{ \lambda I + R^* \mid \lambda \in \mathbb{C}, R \in Z_{HS}(U^*) \}.$$ 

(In this example $Z(U) \neq Z(U^*)$.)
$Z_{L}(H)(U)$ splits in $L(H)$ and a complement is given by the subspace

$$V := \{ X \in L(H) \mid X e_1 = 0 \}.$$ 

For $B \in V \cap Z_{L}(H)(U)$ implies $B = 0$ and if $B \in L(H)$ we have

$$B = (B_{11} I + C) + (B - B_{11} I - C)$$

where $B_{11} = (B e_1, e_1)$ and $C$ is an operator in $Z_{HS}(U)$ with first column equal to the first column of $B$ except $C_{11} = 0$. Thus $C_k = B_k'$, $k \geq 2$ and $C_{11} = 0$. Then $B_{11} I + C \in Z_{L}(H)(U)$ and $B - B_{11} I - C \in V$ hence any operator in $L(H)$ is the sum of an operator in $Z_{L}(H)(U)$ and an operator in $V$.

Suppose now $A_{\infty}(L(H))$ is $\| \| \| -$closed in $HS$. Exactly the same arguments as in example 6.3 show that this implies

6.5.9. \( \exists \delta > 0 \forall X \in V \|\| XU - UX \|\| \geq \delta \| X \| \).

Let $X_n \in L(H)$ be given by

$$X_n = \text{diag}(0, \ldots, 0, 1, 0, \ldots); \ n \geq 2$$

$n$th component.

Then $X_n \in V$, $n \geq 2$, $\| X_n \| = 1$ and

$$\| X_n U - UX_n \| = (\alpha^{2(n-1)} + \alpha^{2n})^{1 \over 2}.$$

Hence

$$\lim_{n \to \infty} \| \text{Ad}_{U}(X_n) \| = 0.$$

This contradicts 6.5.9 and therefore $A_{\infty}(L(H))$ is not $\| \| \| -$closed. From corollary 6.2 it follows that also $A_{\infty}(HS)$ is not $\| \| \| -$closed.

§ 7. The embedding of $HS$ in $HS^+$

In the previous section we have seen that the space of Hilbert-Schmidt operators equipped with $\| \| \|$ is a Banach algebra without identity. In this section we "adjoin" an identity element and describe the standard embedding of $HS$ in the extended space $HS^+$ (see [DUN II], Ch. XI, § 6).

7.1. Definition.

$$HS^+ := \{ \langle \alpha, A \rangle \mid \alpha \in \mathbb{C}, \ A \in HS \},$$

and the operations on $HS^+$ are the following:
addition: \( <\alpha, A> + <\beta, B> := <\alpha + \beta, A + B> \).
scalar multiplication: \( \lambda <\alpha, A> := <\lambda \alpha, \lambda A> \).
multiplication: \( <\alpha, A> <\beta, B> := \alpha\beta A + \beta A + AB \).
involution: \( ( <\alpha, A> )^* := <\bar{\alpha}, A^* > \).
inner product: \( <\alpha, A> , <\beta, B> := \alpha \bar{\beta} + (A,B) \).
1-norm: \( \| <\alpha, A> \|_1 := |\alpha| + \| A \| \).
2-norm: \( \| <\alpha, A> \|_2 := (|\alpha|^2 + \| A \|^2)^{1/2} \).

The following lemma holds

7.2. Lemma. \( H^+ \) provided with the defined algebraic operations and \( \| \cdot \|_1 \) is a Banach algebra with identity \( e = <1,0> \) and involution. \( H^+ \) equipped with \( \| \cdot \|_2 \) is a Hilbert space and the norms \( \| \cdot \|_1 \) and \( \| \cdot \|_2 \) are equivalent on \( H^+ \).

Proof. The first part of the lemma is a standard result (see [DUN II], Ch. XI, § 6). We only prove the equivalence of \( \| \cdot \|_1 \) and \( \| \cdot \|_2 \). If \( <\alpha, A> \in H^+ \) we have
\[
|\alpha|^2 + \| A \|^2 \leq (|\alpha| + \| A \|)^2 \leq 2(|\alpha|^2 + \| A \|^2)
\]
and hence
\[
\| <\alpha, A> \|_2 \leq \| <\alpha, A> \|_1 \leq \sqrt{2}\| <\alpha, A> \|_2 . \]

7.3. Corollary. Lemma 7.2 shows that any \( \| \cdot \|_1 \)-open (closed) set is a \( \| \cdot \|_2 \)-open (closed) set and vice versa, and therefore every subspace (\( \| \cdot \|_1 \) or \( \| \cdot \|_2 \)-closed) has a closed complement, namely the orthogonal complement in the space \( H^+ \), and this complement is also \( \| \cdot \|_1 \)-closed.

7.4. Remark. Note that \( H^+ \) provided with \( \| \cdot \|_1 \) is not a \( B^* \)-algebra, because in general
\[
\| <\alpha, A>^* <\alpha, A> \|_1 \neq \| <\alpha, A> \|^2 .
\]

The natural embedding map \( \text{Emb}: H \rightarrow H^+ \), which maps \( A \) into \( <0, A> \), is an isometric \( * \) isomorphism from \( H \) onto \( \text{Emb}(H) \), which is subalgebra of \( H^+ \). For example we have:
\[
\text{Emb}(A + B) = \text{Emb}(A) + \text{Emb}(B)
\]
\[
\text{Emb}(A^*) = (\text{Emb}(A))^*
\]
(A, B) = (Emb(A), Emb(B))

\|Emb(A)\|_{1,2} = \|A\|.

7.5. Definition. Since \( HS^+ \) with \( \| \cdot \| \) is a Banach algebra with identity \( e = <1, e> \) the set of non-singular elements \( G^+ \) is an open set in \( HS^+ \) containing \( e \) (see § 1). As in § 1 definition 1.2 we can define the mappings \( \alpha^+_a: G^+ \to HS^+ \) by \( \alpha^+_a(g) = g a g^{-1} \) and \( Ad^+_a: HS^+ \to HS^+ \), being the derivative of \( \alpha^+_a \) at \( e \). The kernel of \( Ad^+_a \) is the centralizer of \( a \) in \( HS^+ \); notation \( Z_{HS^+}(a) \).

7.6. Remark. If \( a = <a, A> \), the map \( Ad^+_a \) and the set \( Z_{HS^+}(a) \) are closely related to \( Ad_A \) respectively \( Z_{HS}(A) \). It is easily seen that

7.6.1. \( Ad^+_a(HS^+) = \{<0, B> \mid B \in Ad_A(HS)\} \)

7.6.2. \( Z_{HS^+}(a) = \{<\gamma, C> \mid \gamma \in C, C \in Z_{HS}(A)\} \).

7.7. Corollary. From 7.6 it follows that \( Ad^+_a(HS^+) \) is \( \| \|_1 \)-closed iff \( Ad_A(HS) \) is \( \| \|_1 \)-closed.

7.8. Theorem. Let \( a \in HS^+ \). Then

\((Ad^+_a(HS^+))^{-1} = Z_{HS^+}(a^*)\).

Proof. Use 7.6.1, 7.6.2 and theorem 5.7.

We now define the map \( \theta: HS^+ \to \mathcal{L}(H) \) by

\( \theta(<a, A>) := aI + A \)

(see [DUN II], Ch. XI, § 6). Then \( \theta \) is an injective, continuous, homomorphism from \( HS^+ \) into \( \mathcal{L}(H) \). We only prove the continuity of \( \theta \) (the rest of this statement is also easy to verify)

\[ \| \theta(<a, A>) \| = \| aI + A \| < |a| + \| A \| \leq |a| + \| A \| = \| <a, A> \|_1. \]

Note that \( <a, A> \in G^+ \) (is invertible in \( HS^+ \)) iff \( aI + A \in G \) (is invertible in \( \mathcal{L}(H) \)) and

\[ \theta(<a, A>^{-1}) = (aI + A)^{-1} \]

(see [DUN II], Ch. XI, § 6).
Finally we prove two lemmas which show the relationship between similarity in $\mathcal{HS}^+$ and the induced relation in $\mathcal{HS}$ (note that similarity in $\mathcal{HS}$ is not yet defined).

7.9. Lemma. Let $\langle \alpha, A \rangle, \langle \beta, B \rangle ∈ \mathcal{HS}^+$ and $\langle \gamma, C \rangle ∈ G^+$. Then

$$\langle \beta, B \rangle = \langle \gamma, C \rangle \langle \alpha, A \rangle \langle \gamma, C \rangle^{-1}$$

iff

$$\begin{cases}
\alpha = \beta \\
\langle 0, B \rangle = \langle \gamma, C \rangle \langle 0, A \rangle \langle \gamma, C \rangle^{-1}.
\end{cases}$$

Proof. Note that $\langle \gamma, C \rangle ∈ G^+$ implies $\gamma \neq 0$. The rest of the proof is computation.

7.10. Lemma. Let $\langle \alpha, A \rangle ∈ \mathcal{HS}^+$ and $\langle \gamma, C \rangle ∈ G^+$, then

$$\langle \gamma, C \rangle \langle \alpha, A \rangle \langle \gamma, C \rangle^{-1} = \langle \alpha, (\gamma I + C)A(\gamma I + C)^{-1} \rangle .$$

Proof. Put $\langle \beta, B \rangle := \langle \gamma, C \rangle \langle \alpha, A \rangle \langle \gamma, C \rangle^{-1}$. Applying the preceding lemma we have

$\beta = \alpha$ and $\langle 0, B \rangle = \langle \gamma, C \rangle \langle 0, A \rangle \langle \gamma, C \rangle^{-1}$. Hence (using that $\theta$ is a homomorphism) we find

$$\theta(\langle 0, B \rangle) = \theta(\langle \gamma, C \rangle) \theta(\langle 0, A \rangle) \theta(\langle \gamma, C \rangle^{-1})$$

and therefore

$$B = (\gamma I + C)A(\gamma I + C)^{-1}$$

which completes the proof.

§ 8. Heuristics

In this section we discuss the possible extension of theorem 1.3.6 to deformations of operators defined on an infinite dimensional Hilbert space $\mathcal{H}$. The natural relation with regard to which versality, of deformations of operators is considered is the relation of similarity. If two operators are similar the only difference between them lies in the chosen basis of the underlying Hilbert space $\mathcal{H}$. For example, all spectral properties of two similar operators are the same.
By a deformation of an operator $A_0 \in \mathcal{L}(H)$ we mean a differentiable mapping $A$ from an open neighbourhood $U$ of the origin in a Banach space $E$ into $\mathcal{L}(H)$ with $A(0) = A_0$ and double splitting at 0 (see definition B4). As in definition I.2.1 the space $E$ will be called the base of the deformation. A straightforward generalization of the definition of versal deformation (see definition I.2.4) runs as follows: A deformation $A$ of an operator $A_0 \in \mathcal{L}(H)$ with base $E$ is versal iff for every deformation $B$ of $A_0$ with base $F$ we have

$$B(s) = C(s) A(\varphi(s)) C^{-1}(s)$$

for small $s \in F$; where $C$ is a deformation of the identity operator $I \in \mathcal{L}(H)$ and $\varphi$ is a differentiable map from $F$ into $E$ with $\varphi(0) = 0$. Suppose $A$ is a versal deformation of $A_0$ then by taking the derivatives at $t = 0$ at both sides of 8.1 we obtain an equation analogous to I.3.6.1:

$$[D_0 C, A_0] + [D_0 A, D_0 \varphi] \xi$$

for all $\xi \in T_0 F$.

This implies, just as in the proof of theorem I.3.6, that every operator in $\mathcal{L}(H)$ is the sum of a commutator of the form $[C, A_0]$ and an operator in the image of $D_0 A$. Suppose $A_0$ is normal. Then by corollary 4.3 we have 

$$Z(A_0) \cap \text{Ad} A_0 (\mathcal{L}(H)) = \{0\}.$$ 

Since by theorem 2.5 $Z(A_0)$ is always infinite dimensional, versality of $A$ implies that $\text{Ran}(D_0 A)$ is infinite dimensional. It is not difficult to prove, with the aid of the Kleinecke-Shirokov theorem (theorem 4.1) and theorem 2.5, that a complement of $\text{Ad} A_0 (\mathcal{L}(H))$ is always infinite dimensional (even if $A_0$ is not normal) and therefore there are no versal deformations with finite dimensional base.

Suppose the original operator $A_0$ is Hilbert-Schmidt. Let $S$ denote the norm closure of the set $\text{Ad} A_0 (\mathcal{L}(H))$ in $\mathcal{L}(H)$. Since $\text{HS}$ is a two sided ideal in $\mathcal{L}(H)$ every operator in $\text{Ad} A_0 (\mathcal{L}(H))$ is Hilbert-Schmidt and therefore $S$ is a subset of the set of compact operators on $H$. Hence versality of $A$ implies that $\text{Ran}(D_0 A)$ contains at least a complement of the subspace of compact operators in $\mathcal{L}(H)$. For this reason we only study deformations in a smaller class of operators: not in $\mathcal{L}(H)$ but in the space of Hilbert-Schmidt operators which is still a large and important class. So, we shall consider deformations of Hilbert-Schmidt operators in the space $\text{HS}$. In this case we have two possible ways to define similarity and the orbit. Let $A, B \in \text{HS}$. 

Let $A, B \in \text{HS}$. 

i) Similarity induced from $\mathcal{L}(H)$.
A $\sim B$ iff there is a $C \in G \subset \mathcal{L}(H)$ such that $B = CAC^{-1}$.
The corresponding orbit is $N_1 := a_A(G)$ (see remark 5.4).

ii) Similarity induced from $HS^+$ (see definition 7.5).
A $\sim B$ iff $\langle 0, A \rangle \sim \langle 0, B \rangle$ in $HS^+$ which by lemma 7.9 and lemma 7.10 is equivalent to

$$B = (\gamma I + C)A(\gamma I + C)^{-1} \quad \text{with} \quad \langle \gamma, C \rangle \in G^+.$$ 

The corresponding orbit is

$$N_2 := \{(\gamma I + C)A(\gamma I + C)^{-1} \mid \langle \gamma, C \rangle \in G^+\}$$

(the orbits $N_1$ and $N_2$ need not to be submanifolds of $HS^+$).

8.3. Remark. As defined in § 7 of this chapter, the set $G^+$ is a subset of
$HS^+$. In the heuristic approach of this section, however, we consider $G^+$ as
a subset of $G$:

$$G^+ = G \cap \{\lambda I + C \mid \lambda \in \mathfrak{g}, C \in HS\}.$$ 

Of course we want to keep the base of our versal deformations as "small" as
possible and therefore the orbits as "large" as possible. Obviously $N_2 \subset N_1$,
but by theorem 5.9 we have $\text{Ad}_{A_0}(\mathcal{L}(H)) = \text{Ad}_{A_0}(HS)$. $(A_0 + \text{Ad}_{A_0}(\mathcal{L}(H))$ and

$A_0 + \text{Ad}_{A_0}(HS)$ can be considered as linear approximations of $N_1$ respectively

$N_2$ at $A_0$). This means that for the "size" of the base of a versal deformation it makes
no difference for our theory whether we consider the action of $G$ or $G^+$
(case i) or case ii)) on $HS$ because we shall prove the equivalence of ver-
sality (in fact weak-versality) and transversality to the space $\text{Ad}_{A_0}(HS)$.

In case i) (if we consider the action of the group $G$ on $HS$) it is not gua-
ranteed that there is a submanifold of $G$ minimal transversal to $Z_{\mathcal{L}(H)}(A_0)$
at $I$ because it is not guaranteed that $Z_{\mathcal{L}(H)}(A_0)$ splits in $\mathcal{L}(H)$ (although
it does so when $A_0$ is normal (see theorem 3.2)). This submanifold plays an
important role in the proof of theorem 1.3.6 as well as in the proof of
theorem III.4.2). In case ii) we can always find a submanifold of $G^+$ mini-
mal transversal to $Z_{HS^+}(a_0)$ (where $a_0 = \langle 0, A_0 \rangle$) because $HS^+$ is a Hilbert
space and therefore every subspace splits. Therefore we choose case ii).

Suppose $A$ is a versal deformation of $A_0 \in HS$. Then condition 8.2 is still
valid for deformations of $A_0$ in $HS$. Since, by theorem 5.7 $(\text{Ad}_{A_0}(HS))^\perp =
$$Z_{HS}(A_0^*)$ and $\dim Z_{HS}(A_0^*) = \infty$ (see remark 5.6) the subspace $\text{Ran}(D_0A) \subset HS$
must be infinite dimensional (under the assumption that $A$ is versal). Hence every versal deformation depends on infinitely many (one dimensional) complex parameters (i.e. the base of the deformation is infinite dimensional). A straightforward generalization of theorem 1.3.6 is still impossible. In § 6 we have seen that there are many operators for which $\text{Ad}_{A_0}(\text{HS})$ is not $\| \cdot \|$-closed, (e.g. all normal Hilbert-Schmidt operators). Let $A_0$ be such an operator. Suppose $A$ is a deformation of $A_0$ minimal transversal to $\text{Ad}_{A_0}(\text{HS})$ at 0 that is

$$\text{Ran}(D_0A) \oplus \text{Ad}_{A_0}(\text{HS}) = \text{HS}$$

(see definition B5).

Since $\text{Ad}_{A_0}(\text{HS})$ is not $\| \cdot \|$-closed we can choose $X \in \text{Ad}_{A_0}(\text{HS}) \setminus \text{Ad}_{A_0}(\text{HS})$ and consider the 1-dimensional deformation $B$ of $A_0$ defined by

$$B(t) := A_0 + tX; \quad t \in \mathbb{C}.$$  

The derivative $D_0B: \mathbb{C} \to \text{HS}$ is the linear map $t \mapsto tX; \quad t \in \mathbb{C}$, and therefore $(D_0B)(1)$ cannot be written as the sum of an operator in $\text{Ad}_{A_0}(\text{HS})$ and an operator in $D_0A$ and hence $A$ is not versal (see 8.2). This means that transversality does not imply versality in the sense defined in this section. In chapter III we shall define weak-versality which is equivalent to transversality.
Appendix.

Some lemmas in Hilbert space

Before starting with chapter III we shall give some standard lemmas on projections in Hilbert space. These lemmas are used in the proof of theorem III.4.2 to get round the difficulties of the infinite dimensional case. Let \( \mathcal{H} \) denote a Hilbert space.

1. **Definition.** Let \((V_n)_{n \in \mathbb{N}}\) and \(V\) be subspaces of \( \mathcal{H} \). We define

\[
V_n \lesssim V \iff \begin{cases} 
V_1 \subseteq V_2 \subseteq \cdots \subseteq V \\
\{P_n \lesssim P\}
\end{cases}
\]

where \(P_n\) is the orthogonal projection onto \(V_n\) and \(P\) is the orthogonal projection onto \(V\). \(P_n \lesssim P\) means convergence in the strong operator topology of \(\mathcal{L}(\mathcal{H})\).

2. **Lemma.** Let \((V_n)_{n \in \mathbb{N}}\) and \(V\) be subspaces of \( \mathcal{H} \) such that \(V_1 \subseteq V_2 \subseteq \cdots \subseteq V\). Then \(V_n \lesssim V\) iff for every \(x \in V\)

\[
\lim_{n \to \infty} \min_{v \in V_n} \|x - v\| = 0.
\]

**Proof.** Only the non-trivial if-part is proved here. Choose \(v \in V\) and select a sequence \(v_n \subseteq V_n\) with \(v_n \to v\), then

\[
\|P_n x - P x\| = \|P_n (x - v_n) + P_n x - P_n (x - v_n)\| \leq \leq (\|P\| + \|P_n\|) \|x - v_n\| + \|P_n x - P x\|. \]

Since both \(P\) and \(P_n\) are orthogonal projections and \(V \subseteq V_n\) we obtain

\[
\|P_n x - P x\| \leq 2 \|x - v_n\|.
\]

Hence

\[
\|P_n x - P x\| \to 0 \quad \text{if} \quad n \to \infty.
\]

If \(w \in V^\perp\) then \(P w = P_{\perp w} = 0\). Hence \(P_n \lesssim P\). \(\square\)
3. **Lemma.** Let \( L \in L(h) \) and \( V_n \subseteq V \) in \( h \). Then \( \overline{L(V_n)} \subseteq \overline{L(V)} \).

**Proof.** Clearly

\[
\overline{L(V_1)} \subseteq \overline{L(V_2)} \subseteq \ldots \subseteq \overline{L(V)}.
\]

Let \( x \in \overline{L(V)} \). We first prove that if \( \epsilon > 0 \) there is a \( z \in L(V_n(\epsilon)) \) such that

\[
\| x - z \| < \epsilon.
\]

To do so select \( y = Lv, v \in V \) with

\[
\| x - y \| < \frac{1}{2}\epsilon.
\]

next choose \( n(\epsilon) \) and \( w \in V_n(\epsilon) \), with

\[
\| w - v \| < \frac{\epsilon}{2\| L \| + 1}.
\]

Define \( z := Lw \), then \( z \in L(V_n(\epsilon)) \) and

\[
\| x - z \| \leq \| x - y \| + \| y - z \| \leq \| x - y \| + \| L \| \cdot \| v - w \| < \epsilon.
\]

This proves

\[
\min_{z \in L(V_n(\epsilon))} \| x - z \| < \epsilon.
\]

and hence, since \( \overline{L(V_1)} \subseteq \overline{L(V_2)} \subseteq \ldots \subseteq \overline{L(V)} \)

\[
\min_{z \in L(V_n)} \| x - z \| < \epsilon \quad \text{if } n \geq n(\epsilon).
\]

So

\[
\lim_{n \to \infty} \min_{z \in L(V_n)} \| x - z \| = 0.
\]

Since \( x \in \overline{L(V)} \) is arbitrary the previous lemma proves

\[
\overline{L(V_n)} \subseteq \overline{L(V)}.
\]

We quote a standard result on the sum of subspaces (see [HAL 1], problem 8).
4. **Lemma.** If $V$ and $W$ are subspaces on $H$ with $V \cap W = \{0\}$ and if $V$ has finite dimension, then $V \oplus W$ is closed (equal to $\text{span}(V \cup W)$) and the canonical projection operators

$$
P_V : V \oplus W \to V \quad P_W : V \oplus W \to W
$$

are bounded. (Considered as operators in $\mathcal{L}(V \oplus W)$).

5. **Corollary.** If $V$ and $W$ satisfy the assumptions of lemma 4 and if $N := (V \oplus W)^\perp$ then bounded projections $P_V$, $P_W$ and $P_N$ (onto $V$, $W$ and $N$) exist, such that

$$
P_V + P_W + P_N = \text{id}_H$$

$$
\text{Ker}(P_V) = W \oplus N$$

$$
\text{Ker}(P_W) = V \oplus N$$

$$
\text{Ker}(P_N) = V \oplus W$$

Note that all direct sums are equal to the span and hence are closed.
III. Deformations of Hilbert-Schmidt Operators

§ 0. Introduction

In this chapter we shall consider deformations of Hilbert-Schmidt operators and we shall prove the main theorem of this paper (theorem 5.5) which is the extension of theorem I.3.6.

As pointed out in chapter II, § 8, transversality of a deformation to the closure of \( \text{Ad} (\text{HS}) \) does not imply versality in the sense of chapter I. In § 4 we shall define weak-versality, which, as proved in that section, is equivalent to transversality.

The theory is first developed in the Banach algebra \( \text{HS}^+ \), but with the aid of lemma II.7.9 and lemma II.7.10 the theory can be translated immediately to Hilbert-Schmidt operators (see § 5).

Before starting with § 1 we choose an arbitrary element \( x_0 \in \text{HS}^+ \) which remains fixed throughout the sections 1, 2, 3, 4. In these sections we shall use the shorter notations: \( S^+ \) for the \( \| \|_1 \)-closure of \( \text{Ad}^+_{x_0} (\text{HS}^+) \), \( Z^+ \) for \( Z_{\text{HS}^+}(x_0) \) and \( \text{Ad}^+ \) for the map \( \text{Ad}^+_{x_0} \) (see II, § 7).

§ 1. Slices in \( G^+ \)

In this section we define submanifolds of \( G^+ \subset \text{HS}^+ \) of a simple form, which are called slices. Note that the set \( G^+ \) is a submanifold of \( \text{HS}^+ \) (proof: \( G^+ \) is open).

1.1. Lemma. Suppose \( V \) is a finite dimensional subspace of \( \text{HS}^+ \). Let \( B^+ \) denote the \( \| \|_1 \)-open unit ball in \( \text{HS}^+ \):

\[
B^+ := \{ a \in \text{HS}^+ \mid \|a\|_1 < 1 \}.
\]

Define

\[
G^+(V) := e + (B^+ \cap V) := \{ e + a \mid a \in B^+ \cap V \}
\]

then \( G^+(V) \subset G^+ \) and \( G^+(V) \) is a finite dimensional submanifold of \( G^+ \). The tangent space of \( G^+(V) \) at \( x \) equals \( V \): \( T_{x}G^+(V) = V \).

A submanifold of this type is called a finite dimensional slice.

Proof. Let \( W \) denote the orthogonal complement of \( V \) in \( \text{HS}^+ \). Define

\[
B_{e,1} := e + B^+ \quad \text{and} \quad V_1 := V \cap B^+ \quad \text{and} \quad W_1 := W \cap B^+.
\]

are open sets in the relative topology, induced by \( \| \|_1 \), of \( V \) respectively \( W \).
Let \( P_1 \) and \( P_2 \) denote orthogonal projection on \( V \) respectively \( W \). Define
\[
\psi: B_{e,1} \to V_1 \times W_1
\]
by
\[
\psi(x) := (P_1(x - e), P_2(x - e)).
\]
Then \( \psi \) is a \( C^\infty \)-diffeomorphism from \( B_{e,1} \) onto \( V_1 \times W_1 \) and \( \psi(G^+(V)) = V_1 \times \{0\} \). Hence, by definition B2, \( G^+(V) \) is a submanifold of \( G^+ \) which is diffeomorphic to an open set in \( \mathbb{C}^k \), where \( k = \dim(V) \). The statement about the tangent space is obvious.

1.2. Definition. If \( N_1 \) and \( N_2 \) are \( (C^p, p \geq 1) \) submanifolds of a manifold \( M \) then we say \( N_1 \) intersects \( N_2 \) at \( x \) iff \( x \in N_1 \cap N_2 \) and \( T_x N_1 \cap T_x N_2 = \{0\} \).

1.3. Lemma. Let \( V \) be a finite dimensional subspace of \( HS^+ \) such that \( V \cap Z^+ = \{0\} \). Then the slice \( G^+(V) \) intersects \( Z^+ \) at \( e \).

Proof. This is a trivial consequence of definition 1.2 and the previous lemma.

§ 2. Deformations in \( HS^+ \). Versality in a submanifold

2.0. Notation. The letters \( H \) and \( K \) will denote Hilbert spaces and \( \Omega_H, \Omega_K \) will always denote open neighbourhoods of the origin in \( H \) respectively \( K \).

2.1. Definition. A deformation of an element \( x_0 \in HS^+ \) is a map \( x \in C^1(\Omega_H \to HS^+) \) such that \( x(0) = x_0 \) and \( x \) is double splitting at \( 0 \) (see definition B4). The space \( H \) is called the base of the deformation.

2.2. Remark. Since \( x \) is a map from an open subset of a Hilbert space into a Hilbert space, double splitting at \( 0 \) is equivalent to \( \text{Ran}(D_0 x) \) is closed (see definition B4).

In the following lemma we introduce a submanifold of \( HS^+ \).

2.3. Lemma. Suppose \( G^+(V) \) is a finite dimensional slice intersecting \( Z^+ \) at \( e \), i.e. \( V \cap Z^+ = \{0\} \) (see definition 1.2).

Let \( x \) be a deformation of \( x_0 \) with base \( H \), transversal to \( x_0 + S^+ \) at \( 0 \). Assume furthermore that
2.3.1. \( x_* := D_0 x \) is injective.

2.3.2. \( \text{Ran} \ x_* \cap \text{Ad}^+(V) = \{0\} \).

Define
\[
M_0 := \{gx(t)g^{-1} \mid g \in G^+(V), t \in \Omega_H\}
\]
(where \( \Omega_H \) is the open set on which \( x \) is defined).

Then there is an open ball \( B_{\text{Ad}^+(V)} \) centered at \( x_0 \) such that
\[
M := M_0 \cap B_{x_0}
\]
is a submanifold of \( \text{Ad}^+(V) \).

Proof.

Since \( \text{Ad}^+(V) \) is finite dimensional and \( \text{Ran} \ x_* \cap S^+ \) is closed (\( \text{Ran} \ x_* \) and \( S^+ \) are closed) it follows from II, appendix, lemma 4 that the space
\( (\text{Ran} \ x_* \cap S^+) \oplus \text{Ad}^+(V) \) is closed (note that we may write \( \oplus \) since by 2.3.2 \( \text{Ran} \ x_* \cap \text{Ad}^+(V) = \{0\} \)). Let \( N \) denote a closed complement of the subspace
\( (\text{Ran} \ x_* \cap S^+) \oplus \text{Ad}^+(V) \) in \( S^+ \) (e.g. the orthogonal complement in \( S^+ \)). We shall prove that
\[(0,e,0) \in U_1 \times U_2 \times U_3 \]
\[x_0 \in U_0 \]
and \(\gamma\) is a diffeomorphism from \(U_1 \times U_2 \times U_3\) onto \(U_0\). Now
\[\gamma^{-1}(x_0) = (0,e,0)\]
and
\[\gamma^{-1}(U_0 \cap M_0) = U_1 \times U_2 \times \{0\}.\]
Hence, by definition B2, \(M = M_0 \cap B(x_0)\) is a submanifold of \(HS^+\) if \(B(x_0) \subset U_0\).

The tangent space at \(x_0\) is the subspace
\[T_{x_0} M = \text{Ran } x_0 \oplus \text{Ad}^+(V).\]

The following lemma deals with deformations of \(x_0\) with values in \(M\).

2.4. Lemma. Let \(G^+(V)\), \(x\) and \(M\) be defined as in lemma 2.3. Suppose \(y\) is a deformation of \(x_0\) with base \(K\) and values in \(M\), i.e.
\[y \in C^1(\Omega^1_K \rightarrow M);\]
\(y\) is double splitting at 0 and \(y(0) = x_0\).
Then there is an open neighbourhood \(\Omega^1_K\) of the origin in \(K\) and there are mappings
\[c \in C^1(\Omega^1_K \rightarrow G^+(V))\]
\[\varphi \in C^1(\Omega^1_K \rightarrow \Omega_H)\]
with \(c(0) = e\) and \(\varphi(0) = 0\) such that
\[y(t) = c(t)x(\varphi(t))c^{-1}(t); \quad t \in \Omega^1_K.\]

Proof. (The proof of this lemma is analogous to the proof of theorem I.3.6).
The set \(B^+ \cap V\) is open in the relative topology of \(V\) and contained in \(G^+\).
Define
\[\beta: (B^+ \cap V) \times \Omega_H \rightarrow M\]
by
\[\beta(v,t) := (e+v)x(t)(e+v)^{-1}.\]
This definition makes sense since \(e + v \in G^+\) if \(v \in B^+ \cap V\). With the same arguments as used in the proofs of theorem I.3.6 and lemma 2.3 it can be
Ran \( x_* \oplus \text{Ad}^+(V) \oplus N = HS^+ \).

Let \( y \in HS^+ \) then, by the transversality of the deformation \( x \) to \( x_0 + S^+ \) we have \( y = y_1 + y_2 \) with \( y_1 \subset \text{Ran} \ x_* \) and \( y_2 \in S^+ \oplus \text{Ran} \ x_* \). Since

\[
(Ran \ x_* \cap S^+) + \text{Ad}^+(V) + N = S^+
\]

we have

\[
y_2 = 0 + y_3 + y_4
\]

with \( y_3 \subset \text{Ad}^+(V) \) and \( y_4 \subset N \). Hence \( y = y_1 + y_3 + y_4 \) with \( y_1 \subset \text{Ran} \ x_* \), \( y_3 \subset \text{Ad}^+(V) \), \( y_4 \subset N \). We leave it to the reader to notice that \( \text{Ran} \ x_* \), \( \text{Ad}^+(V) \) and \( N \) are mutually independent.

We now define \( y: \Omega_H \times G^+(V) \times N \to HS^+ \) by

\[
y(t,g,n) := gx(t)g^{-1} + n.
\]

Then \( y(0,e,0) = x_0 \) and \( y \) is differentiable in \( \Omega_H \times G^+(V) \times N \) (see [LAN], Ch. I, § 3, prop. 11). The derivative at \( (0,e,0) \)

\[
\gamma_* := D(0,e,0) \gamma: H \times V \times N \to HS^+
\]

is given by

\[
\gamma_*(t,g,n) = x_*(t) + [g,x_0] + n.
\]

(The space \( H \times V \times N \) becomes a Banach space in one of the usual ways; by defining \( \|t,g,n\| := \max(\|t\|_H,\|g\|_1,\|n\|_1) \) and then the map \( \gamma_* \) is a bounded linear operator from \( H \times V \times N \) into \( HS^+ \)).

Since

\[
\text{Ran} \ x_* \oplus \text{Ad}^+(V) \oplus N = HS^+
\]

and

\[
V \cap Z^+ = \{0\}
\]

we may conclude

\[
\ker \gamma_* = (0,0,0)
\]

and

\[
\text{Ran} \ \gamma_* = HS^+.
\]

Hence, since \( \gamma_* \) is bounded, it follows from the closed graph theorem of Banach that \( \gamma_* \) is invertible as a linear operator. Hence by the inverse function theorem (see [LAN], Ch. I, § 5, Th. 1). \( \gamma \) is a local diffeomorphism at \( (0,e,0) \) and therefore there are open sets \( U_1 \subset \Omega_H \), \( U_2 \subset G^+(V) \), \( U_3 \subset N \) and \( U_0 \subset HS^+ \) such that:
shown that $\beta$ determines a $C^1$-diffeomorphism from an open neighbourhood $\Omega^0_V \times \Omega^0_H$ of $(0,0)$ in $V \times H$ onto an open subset $\Omega_0^0 \subset M$ containing $x_0$ ($\Omega^0_V$ is open in the relative topology of $V$ and $\Omega^0_H$ is a set of the form $\sigma \cap M$ where $\sigma$ is open in $HS^+$; $M$ is given the relative topology induced by $\| \cdot \|_1$).

Let $\pi_1$ and $\pi_2$ denote the canonical projections of $\Omega^0_V \times \Omega^0_H$ onto $\Omega^0_V$ respectively $\Omega^0_H$. Obviously there is an open set $\Omega_1^1 \subset \Omega^0_K$ such that $y(\Omega_1^1) \subset \Omega^0_K$. Hence if $t \in \Omega_1^1$ we have:

$$y(t) = \beta(w, s)$$

for some $w \in V$ and $s \in H$. Hence

$$y(t) = c(t)x(\phi(t))c^{-1}(t); \quad t \in \Omega_1^1$$

where

$$c(t) := e + \pi_1\beta^{-1}(y(t))$$

and

$$\phi(t) := \pi_2\beta^{-1}(y(t)).$$

Since $\beta^{-1}$ is $C^1$ on $\Omega^0_M$ and $\pi_1$ and $\pi_2$ are both $C^\infty$, it follows from [LAN], Ch. I, § 3, prop. 7 that $c$ and $\phi$ are $C^1$ on $\Omega_1^1_K$. $\square$

2.5. Remark. As in chapter I the theory in this chapter is essentially local. We do not care how small $\Omega_1^1_K$ is.

§ 3. An exponential map

Let $V$, $G^+(V)$ and $M$ be defined as in lemma 2.3 and $\beta$ as in the proof of lemma 2.4.

3.1. Definition. The mapping $\text{EXP}: T_{x_0}M \to M$ is defined by

$$\text{EXP} := \beta \circ \beta^{-1}_*$$

where $\beta_* := D(0,0)\beta$.

3.2. Lemma. If $a \in T_{x_0}M$ and $\|a\|_1$ is sufficiently small we have

3.2.1. $\|\text{EXP}(a) - (x_0 + a)\|_1 = o(\|a\|_1)$. 
Proof. We compute \( \text{EXP}(0) \) and \( D_0 \text{EXP} \):

\[
\text{EXP}(0) = \beta(\beta^{-1}_*(0)) = \beta(0) = x_0
\]

\[
D_0 \text{EXP} = D_0(\beta \circ \beta^{-1}) = \text{id}_{T_{x_0} M}
\]

and this implies 3.2.1.

§ 4. Weakly versal deformations. Weak-versality \( \Leftrightarrow \) transversality

4.1. Definition. A deformation \( x \) of \( x_0 \in HS^+ \) with base \( H \) is weakly versal iff for every deformation \( y \) of \( x_0 \) with finite dimensional base \( K \) there exists a map \( \varphi \in C^1(\Omega^*_K \to \Omega^*_H) \), with \( \varphi(0) = 0 \), such that for every \( \varepsilon > 0 \) there is a deformation \( c_\varepsilon \) of the identity \( e \in HS^+ \) with base \( K \) such that:

4.1.1. \[ \| y(s) - c_\varepsilon(s)x(\varphi(s))c^{-1}_\varepsilon(s) \|_1 \leq \varepsilon \|s\|; \quad s \in \Omega^*_K \]

where \( \Omega^*_K \) is open in \( K \) and depends on \( \varepsilon \). (note that if \( \Omega^*_K \) is small enough \( c_\varepsilon(s) \in G^+ \).

In the next theorem we shall prove the equivalence of weak-versality and transversality to the set \( x_0 + S^+ \). The proof of the implication weak-versality \( \Rightarrow \) transversality is rather easy. The proof of the implication the other way around is based on the following idea. The map \( y - x_0 \) splits into two parts (depending on \( \varepsilon \)) \( y_1 \) and \( y_2 \), \( y_1 \) with values in \( T_{x_0}^* M \), and \( y_2 \) with values in the orthogonal complement of \( T_{x_0}^* M \). \( M \) is a submanifold of \( HS^+ \) of the type described in lemma 2.3). The map \( x_0 + y_1 \) is close enough to a deformation described in lemma 2.4 and \( \|y_2\|_1 \) is small. At the end of the proof we shall see that the transformation of the base, \( \varphi_\varepsilon \), can be chosen independently of \( \varepsilon \).

4.2. Theorem. (Weak-versality \( \Leftrightarrow \) transversality). \( x \) is a weakly versal deformation of \( x_0 \) iff \( x \) is transversal to \( x_0 + S^+ \) at 0.

Proof. A) weak-versality \( \Rightarrow \) transversality. Suppose \( y \in C^1(\Omega^*_K \to HS^+) \) is an arbitrary deformation of \( x_0 \) with base \( K \). Then by the weak-versality of \( x \) we have

4.2.1. \[ \|y(s) - c_\varepsilon(s)x(\varphi(s))c^{-1}_\varepsilon(s)\|_1 \leq \varepsilon \|s\|; \quad s \in \Omega^*_K \]

Define
4.2.2. \[ z_\varepsilon(s) := y(s) - c_\varepsilon(s)x(\varphi(s))c^{-1}_\varepsilon(s) \]

for \( s \in \Omega_\varepsilon \). Then \( z_\varepsilon(0) = 0 \) and \( z_\varepsilon \) is \( C^1 \) on \( \Omega_\varepsilon \).

The derivative of \( z_\varepsilon \) at \( s = 0 \)

\[ z_{\varepsilon, *} := D_0z_\varepsilon : K \to HS^+ \]

is given by

4.2.3. \[ z_{\varepsilon, *}(\xi) = y_*(\xi) - ([c_{\varepsilon, *}(\xi), x_0] + x_*\varphi_*(\xi)) \]

for all \( \xi \in K \), where \( y_* = D_0y, x_* := D_0x, \varphi_* := D_0\varphi \) and \( c_{\varepsilon, *} := D_0c_\varepsilon \). From 4.2.1 and 4.2.3 we derive \( \|z_{\varepsilon, *}\| < 2\varepsilon \) where the norm is the norm of \( \mathcal{L}(K \to HS^+) \).

Using 4.2.3 we obtain

\[ y_* = x_* \varphi_* + \lim_{\varepsilon \to 0} [c_{\varepsilon, *}, x_0] \]

where the limit is taken in the norm topology of \( \mathcal{L}(K \to HS^+) \). Hence any vector in \( \text{Ran} \ y_* \) can be written as the sum of a vector in \( \text{Ran} \ x_* \) and a vector in \( S^+ \). Since \( y \) is arbitrary it follows that \( \text{Ran} \ x_* + S^+ = HS^+ \) and therefore \( x \) is transversal to \( x_0 + S^+ \) (see definition B5).

B) Transversality \( \Rightarrow \) weak-versality. Let \( x \) be a deformation of \( x_0 \) with base \( H \) transversal to \( x_0 + S^+ \) at 0. Then \( \text{Ran} \ x_* \) is closed and contains a complement of \( S^+ \) in \( HS^+ \).

We shall assume that \( x_* \) satisfies the conditions

4.2.4. \( x_* \) is injective .

4.2.5. \( \text{Ran} \ x_* \cap S^+ = \{0\} \).

These assumptions imply the conditions 2.3.1 and 2.3.2 of lemma 2.3 for every finite dimensional \( V \). On the other hand these assumptions cause no loss of generality. Since, if \( x_* \) is not injective we replace the base \( H \) by \( H' \) (e.g. the orthogonal complement of \( \text{Ker} \ x_* \)) such that the derivative \( x_*' \) at 0 of the restriction \( x' \) of \( x \) to \( H' \) is injective and the deformation \( x_*' \) is still transversal to the manifold \( x_0 + S^+ \) at 0. Obviously weak-versality of \( x_*' \) implies weak-versality of \( x \). Moreover, if condition 4.2.5 is not satisfied we can use similar arguments: since \( S^+ \) is closed and \( x_* \) is continuous \( x_*'(S^+) \) is closed in \( H \). Therefore it is possible to replace \( H \) by \( H' \) (a complement of the space \( x_*'(S^+) \cap \text{Ran} \ x_* \)) such that the restriction \( x' \) of \( x \) to \( H' \) satisfies 4.2.5 and is still transversal to \( x_0 + S^+ \) at 0. If \( x' \) is weakly versal then \( x \) itself is certainly weakly versal.
Now if these assumptions are fulfilled we choose a sequence of finite dimensional subspaces \((V_n)_{n \in \mathbb{N}}\) with \(V_n \subset \text{HS}^+\) such that \(V_n \varpropto V\) where \(V\) is the orthogonal complement of \(Z^+\) in \(\text{HS}^+\) and \(\varpropto\) is defined in II, appendix, definition 1. The \(V_n\)'s can be chosen as follows:

\[
V_n := \text{span}\{f_1, \ldots, f_n\}
\]

where \(f_1, f_2, \ldots\) is an orthonormal basis for \(V \subset \text{HS}^+\). Applying lemma 3 of the appendix of chapter II we may conclude

\[
\text{Ad}^+(V_n) \varpropto \text{Ad}^+(V) = S^+.
\]

Define \(M_n\) as follows

\[
M_n := \{g \in G(V_n) \mid g \in G(V_n), t \in \Omega_n \cap B_n\}
\]

where \(B_n\) is an open ball centered at \(x_0\) and \(G(V_n)\) is a finite dimensional slice (see § 1). Then by lemma 2.3 \(M_n\) is a submanifold of \(\text{HS}^+\) if the ball \(B_n\) is small enough.

For every \(n\) we have

\[
T_n M_n = \text{Ran} x^* \oplus \text{Ad}^+(V_n)
\]

and this space is closed by II, appendix, lemma 4. Since \(x\) satisfies 4.2.5 the sum is a direct sum. Let \(N_n\) denote the orthogonal complement of \(\text{Ran} x^* \oplus \text{Ad}^+(V_n)\) in \(\text{HS}^+\). Then

\[
\text{Ran} x^* \oplus \text{Ad}^+(V_n) \oplus N_n = \text{HS}^+
\]

(compare the proof of lemma 2.3). From II, appendix, corollary 5 it follows that bounded projections \(P, Q_n\) and \(R_n\) exist onto \(\text{Ran} x^*, \text{Ad}^+(V_n)\) and \(N_n\) respectively such that

\[
P + Q_n + R_n = \text{id}_{\text{HS}^+}
\]

\[
\text{Ker}(P) = \text{Ad}^+(V_n) \oplus N_n
\]

\[
\text{Ker}(Q_n) = \text{Ran} x^* \oplus N_n
\]

\[
\text{Ker}(R_n) = \text{Ran} x^* \oplus \text{Ad}^+(V_n).
\]

Define

\[
L_n := P + Q_n
\]

then \(L_n\) is the projector onto \(\text{Ran} x^* \oplus \text{Ad}^+(V_n)\) with kernel \(N_n\). Since \(\text{Ad}^+(V_n) \varpropto S^+\) and \(\text{Ran} x^* \oplus S^+ = \text{HS}^+\) we have \(\text{Ran} x^* \oplus \text{Ad}^+(V_n) \varpropto \text{HS}^+\) and therefore

\[
4.2.6. \quad L_n \varpropto \text{id}_{\text{HS}^+}.
\]
Since $R_n + L_n = \text{id}_{HS^+}$ we also have

$$4.2.7. \quad R_n \nrightarrow 0 \quad (N_n \nrightarrow [0]).$$

Now let $y$ be any deformation of $x_0$ with finite dimensional base $K$. Write

$$4.2.8. \quad y(s) - x_0 = y_1(s) + y_2(s)$$

with

$$y_1(s) := L_n(y(s) - x_0) \quad \text{and} \quad y_2(s) := R_n(y(s) - x_0).$$

Then $y_1 \in C^1(\Omega_K \to T_{x_0^0} M)$ and $y_2 \in C^1(\Omega_K \to N_n)$ (where $\Omega_K$ is the open set on which $y$ is defined). Since $y(s) - x_0 = y_1(s) + o(\|s\|)$ (where $y_1 := D_0y$) we have

$$4.2.9. \quad y_2(s) = R_n y_1(s) + o(\|s\|); \quad s \in \Omega_K$$

(note that $\|R_n\| \leq 1$ and therefore the $o$ term is uniform in $n$).

Suppose $\dim K = m$. Since $y_1$ is linear and bounded the image of the closed unit ball $B_K^*$ in $K$ is contained in an $m$-dimensional disc $D \subset HS^+$, that is the intersection of an $m$-dimensional subspace and a closed ball centered at 0 with radius $\|y_1^*\|$. By 4.2.7 we have

$$\forall f \in D \lim_{n \to \infty} R_n f = 0.$$

Since $D$ is finite dimensional and $R_n$ is linear we have

$$\lim_{n \to \infty} \max_{s \in B_K^*} \|R_n y_1(s)\| = 0.$$

Hence, if $\epsilon > 0$ is fixed, we can choose $n$ so large that

$$\|R_n y_1(s)\|_1 \leq \frac{\epsilon}{8}\|s\|; \quad s \in K.$$

Combining this with 4.2.9 we obtain

$$4.2.10. \quad \|y_2(s)\|_1 \leq \frac{\epsilon}{4}\|s\|$$

on a sufficiently small subset $\Omega_K^0$ of $\Omega_K$. Since the image of $\Omega_K^0$ under $y_1$ is contained in $T_{x_0^0} M$ and $y_1(0) = 0$ it is possible to define

$$z(s) := \text{EXP}(y_1(s))$$
on an open set in $K$ containing 0 (depending on $n$), where $\text{EXP}$ is defined in definition 3.1.

Now $z(s) \in M_n$ and $z(0) = \text{EXP}(0) = x_0$. Hence $z$ is a deformation of $x_0$ with values in $M_n$. By lemma 2.4 there are mappings

$$\varphi_n \in C^1(\mathcal{J}_n \to \Omega_n)$$

$$c_n \in C^1(\mathcal{J}_n \to G^+(V_n))$$

with $\varphi_n(0) = 0$, $c_n(0) = e$ and

$$z(s) = c_n(s)x(\varphi_n(s))c_n^{-1}(s); \quad s \in \mathcal{J}_n.$$  

$\mathcal{J}_n$ is open in $K$.

Since $y_1(s) = \text{EXP}^{-1}(z(s))$ it follows from 3.2.1 that

$$y_1(s) = -x_0 + c_n(s)x(\varphi_n(s))c_n^{-1}(s) + o(\|s\|).$$

Combining this with 4.2.8 and 4.2.10 we obtain

$$4.2.11. \quad \|y(s) - c_n(s)x(\varphi_n(s))c_n^{-1}(s)\|_1 \leq \varepsilon\|s\|$$

for $s \in \mathcal{J}_n$, where $\Omega_n$ is sufficiently small and open. (note that $n$ depends on $\varepsilon$). The only thing left to prove is that $\varphi_n$ can be chosen independently of $n$ (of $\varepsilon$).

Let $\pi_2: V_n \times H \to H$ denote the canonical projection on the second factor.

Let $\beta_n$ denote the diffeomorphism defined in lemma 2.4. Define $\beta_n, = D(0,0)^\beta_n$ and

$$T_n := \pi_2 \circ \beta_n^{-1} \circ L_n$$

then $T_n$ is a linear map from $HS^+$ into $H$ and the following diagram commutes.
By lemma 2.4 we have
\[
\varphi_n = \pi_n^2(\beta_n^{-1}(z)) = \pi_n(\beta_n^{-1}(\text{EXP}(y_n))) = \\
\pi_n(\beta_n^{-1} \circ \beta_n \circ \beta_n^{-1} y_n) = \\
= \pi_n(\beta_n \circ \beta_n^{-1} y_n) = \\
= \pi_n(\beta_n^{-1} \circ (L_n(y - x_0))).
\]
Hence
\[4.2.12. \quad \varphi_n(s) = T_n(y(s) - x_0)\]

\( \varphi_n \) is only defined on a small neighbourhood of \( 0 \in K \), depending on \( n \), but we can extend \( \text{dom} \ \varphi_n \) to \( \text{dom} \ \text{y} \) by 4.2.12 because \( \text{dom} \ T_n = HS^+ \). From 4.2.12 we can deduce \( \varphi_{n,*} = T_n v \) where \( \varphi_n := D_0 \varphi \). We shall prove that \((\varphi_{n,*})_{n \in \mathbb{N}}\) is a Cauchy sequence in \( L(K \rightarrow H) \). To do this consider first the composition
\[
\pi_n^2 \circ \beta_n^{-1}_{n,*} : \text{Ran} \ \text{x} \oplus \text{Ad}^+(V_n) \rightarrow H
\]
\[
x_n(t) + [v, x_0] \xrightarrow{\beta_{n,*}} (v, t) \xrightarrow{\pi_n^2} t.
\]
Since \( x_n \) is injective and \( \text{Ran} \ \text{x} \) is closed it follows from the closed graph theorem that there is a \( \delta > 0 \) such that
\[
\|x_n(t)\|_1 \geq \delta \|t\|; \quad t \in H
\]
and therefore \( \|\pi_n^2 \circ \beta_n^{-1}_{n,*}\| \) is bounded by a constant independent on \( n \) say \( A \). If \( f \in HS^+ \) and \( n > m \) we have
\[
T_n f - T_m f = \pi_n^2 \circ \beta_{n,*}^{-1} \circ L_n(f) - \pi_m^2 \circ \beta_{m,*}^{-1} \circ L_m(f) = \\
= \pi_n^2 \circ \beta_{n,*}^{-1} \circ (L_n - L_m)(f),
\]
because
\[
\pi_n^2 = \pi_m^2 \text{ on } V_m \times H \quad (n > m)
\]
\[
\beta_{n,*} = \beta_{m,*} \text{ on } \text{Ran} \ \text{x} \oplus \text{Ad}^+(V_m) \quad (n > m).
\]
Hence
\[
\|T_n f - T_m f\| \leq \|\pi_n^2 \circ \beta_{n,*}^{-1}\| \| (L_n - L_m) f\| \leq A \| (L_n - L_m) f\|.
\]
Since $L_n \frac{S}{\nabla} \text{id}_{HS^+}$ we may conclude

$$\forall f \in HS^+ \lim_{n,m \to \infty, (n>m)} \|T_n f - T_m f\| = 0 .$$

Hence

$$\forall s \in K \lim_{n,m \to \infty, (n>m)} \|\phi_{n,*}(s) - \phi_{m,*}(s)\| = 0 .$$

Since $K$ is finite dimensional and $\phi_{n,*}$ is linear we may conclude:

$$\lim_{n,m \to \infty, (n>m)} \max \frac{\|\phi_{n,*}(s) - \phi_{m,*}(s)\|}{\|s\| = 1} = 0 .$$

Hence

$$\lim_{n,m \to \infty, (n>m)} \|\phi_{n,*} - \phi_{m,*}\| = 0$$

and therefore the sequence $(\phi_{n,*})_{n \in \mathbb{N}}$ is a Cauchy sequence in the Banach space $L(K \to H)$. Define $\psi \in L(K \to H)$ by

$$\psi := \lim_{n \to \infty} \phi_{n,*} .$$

We shall prove that if $n$ is large enough we may replace $\phi_n$ by $\psi$ in 4.2.11 if $h_\epsilon$ is replaced by $\epsilon$. First we choose $n$ so large that

4.2.13. $\|x_x(\psi(s)) - x_*(\phi_{n,*}(s))\|_1 \leq \frac{1}{24} \epsilon \|s\| .$

Furthermore, we choose a small open set in $K$ on which

4.2.14. $\|x(\phi_n(s)) - (x_0 + x_*(\phi_{n,*}(s)))\|_1 \leq \frac{1}{24} \epsilon \|s\|$

(note that this is possible since $x(0) = x_0$ and $\phi_{n,*}(0) = 0$).

Finally we restrict ourselves to an open set such that

4.2.15. $\|x(\psi(s)) - (x_0 + x_*(\psi(s)))\|_1 \leq \frac{1}{24} \epsilon \|s\| .$

Combining 4.2.13, 4.2.14 and 4.2.15 we obtain

4.2.16. $\|x(\phi_n(s)) - x(\psi(s))\|_1 \leq \frac{1}{8} \epsilon \|s\|$

on a (small) open set in $K$. 

Now let \( \Omega^n_K \) be open in \( K \) such that
\[ \max(\|c_n(s)\|,\|c_n^{-1}(s)\|) < 2 \]
for \( s \in \Omega^n_K \) and also 4.2.11 and 4.2.16 hold on \( \Omega^n_K \). Then
\[ \|y(s) - c_n(s)x(\psi(s))c_n^{-1}(s)\| \leq \varepsilon \| s \| \]
for \( s \in \Omega^n_K \), and the proof is complete.

\[ \square \]

§ 5. Deformations of Hilbert-Schmidt operators

5.0. Introduction

In this section we employ the theory developed in § 4 to study deformations of Hilbert-Schmidt operators. For the transition of deformations in \( HS^+ \) to deformations in \( HS \) we use lemma II.7.9 and lemma II.7.10. For an arbitrary operator \( A_0 \in HS \) a minimal weakly versal deformation is constructed in theorem 5.6 (by minimal weakly versal we mean minimal transversal). As an example we shall give a weakly versal deformation of a diagonal operator.

From now on \( A_0 \in HS \) is fixed and we shall use the shorter notations \( Z(A^*_0) \), \( Ad \) and \( S \) for respectively \( Z_{HS}(A^*_0) \), \( Ad_{A_0} \) and the \( \|\|\)-closure of \( Ad_{A_0} (HS) \).

5.1. Definition. A deformation of an operator \( A_0 \in HS \) is a map \( A \in C^1(\Omega_H \rightarrow HS) \) such that \( A(0) = A_0 \) and \( A \) is double splitting at 0. As usual \( \Omega_H \) is open in \( H \), the base of the deformation.

5.2. Definition. A deformation of an operator \( A_0 \in HS \) with base \( H \) is weakly versal iff for every deformation \( B \) of \( A_0 \) with finite dimensional base \( K \) there exists a map \( \varphi \in C^1(\Omega_K \rightarrow H) \), with \( \varphi(0) = 0 \), such that for every \( \varepsilon > 0 \) there is a deformation \( C_\varepsilon(s) \) of the identity operator \( I \in \mathcal{L}(H) \) of the form
\[ C_\varepsilon(s) = \gamma_\varepsilon(s)I + D_\varepsilon(s), \quad s \in K, \]
where \( D_\varepsilon \in HS \) is a deformation of \( \delta \in HS \) and \( \gamma_\varepsilon \) is a deformation of \( 1 \in \mathcal{C} \), such that
\[ \|B(s) - C_\varepsilon(s)A(\varphi(s))C_\varepsilon^{-1}(s)\| \leq \varepsilon \| s \|; \quad s \in \Omega^\varepsilon_K. \]

The reader may have noticed that definition 5.1 and 5.2 are analogous to definition 2.1 and 4.1.
5.3. Lemma. If $A$ is a deformation of $A_0 \in HS$ with base $H$ and $x_0 := \langle 0, A_0 \rangle$ then the map $x: C \oplus H \to HS^+$ defined by

$$x(a,t) := \langle a, A(t) \rangle$$

is a deformation of $x_0 \in HS^+$ in the sense of definition 2.1 with base $C \oplus H$. We shall say that the deformation $x$ corresponds to the deformation $A$.

The proof of this lemma is left to the reader.

5.4. Lemma. Let $A$ be a deformation of $A_0$ and $x$ the corresponding deformation of $x_0 := \langle 0, A_0 \rangle$ (see 5.3). Then $A$ is weakly versal iff $x$ is weakly versal and

$A$ is transversal to $A_0 + S$ iff $x$ is transversal to $x_0 + S$.

Proof. We shall only prove

$A$ is weakly versal only if $x$ is weakly versal.

$A$ is transversal to $A_0 + S$ if $x$ is transversal to $x_0 + S$.

The remainder of the proof is left to the reader. Suppose $A$ is a weakly versal deformation of $A_0$ with base $H$ and let $y := \langle \beta, Y \rangle$ be any deformation of $x_0$ with finite dimensional base $K$. Then $Y$ is a deformation of $A_0$ and hence, since $A$ is weakly versal, there is a map $\varphi \in C^1(\Omega_K \to H)$ such that for every $\epsilon > 0$ there is a deformation of $I$ of the form $C_\epsilon = \gamma_\epsilon I + D_\epsilon$ such that

$$\|Y(s) - C_\epsilon(s)A(\varphi(s))C^{-1}_\epsilon(s)\| \leq \epsilon\|s\|; \quad s \in \Omega_K^\epsilon.$$

Hence by lemma II.7.9 and lemma II.7.10 we have

$$\|\langle \beta(s), Y(s) \rangle - \langle \gamma_\epsilon(s), D_\epsilon(s) \rangle \langle \beta(s), A(\varphi(s)) \rangle \langle \gamma_\epsilon(s), D_\epsilon(s)^{-1} \rangle\| \leq \epsilon\|s\|$$

$s \in \Omega_K^\epsilon$.

which can be written as

$$\|y(s) - C_\epsilon(s)x(\psi(s))C^{-1}_\epsilon(s)\| \leq \epsilon\|s\|; \quad s \in \Omega_K^\epsilon.$$

where $C_\epsilon(s) := \langle \gamma(s), D_\epsilon(s) \rangle$ is a deformation of $\epsilon \in HS^+$ and $\psi(s) := (\beta(s), \psi(s)) \in C \oplus H$ satisfies $\psi(0) = (0,0)$. This proves the weak-versality of the deformation $x$ (see definition 4.1).

Suppose $x$ is transversal to $x_0 + S$ at $0$. Then $\text{Ran } x_*$ contains a closed complement of $S^+$ in $HS^+$.

Since $\text{Ran } x_* = \text{Ran } \langle \text{id}, A_* \rangle = C \oplus \text{Ran } A_*$

(by $\langle \text{id}, A_* \rangle$ we mean the map $(a,t) \to \langle a, A_*(t) \rangle$ and since $S^+ = \{0, B \mid B \in S\}$)
(see remark II.7.6) this implies that \( \text{Ran} \ A_* \) contains a closed complement of \( S \) in \( \text{HS} \) and hence \( A \) is transversal to \( A_0 + S \) at 0.

The following theorem, which is the main theorem of this paper follows from lemma 5.4 and theorem 4.2.

5.5. Theorem. A deformation \( A \) of an operator \( A_0 \in \text{HS} \) is weakly versal iff \( A \) is transversal to \( A_0 + S \) at 0.

Construction of weakly versal deformations.

5.6. Theorem. Every operator \( A_0 \in \text{HS} \) has a (minimal) weakly versal deformation. It can be given the following form

\[
A(X) := A_0 + X; \quad X \in Z(A_0^*).
\]

The base of this deformation is \( Z(A_0^*) \).

Proof. Note that \( A_* \) is the linear embedding map from \( Z(A_0^*) \) into \( \text{HS} \). Hence \( A \) is double splitting at 0 and \( \text{Ran} \ A_* = Z(A_0^*) \). Since \( Z(A_0^*) \) is the orthogonal complement of \( S \) in \( \text{HS} \) (see theorem II.5.7) the deformation \( A \) is (orthogonal) transversal to \( A_0 + S \) at 0 and hence by theorem 5.5 \( A \) is weakly versal. \( A \) is minimal weakly versal because \( A \) is minimal transversal (see definition B5).

5.7. Corollary. If \( B \) is a deformation of \( A_0 \in \text{HS} \) with finite dimensional base \( K \) then there is a map \( \varphi: K \to Z(A_0^*) \) with \( \varphi(0) = 0 \) such that for every \( \varepsilon > 0 \) there is a deformation \( C_\varepsilon \) of \( I \in \mathcal{L}(\mathcal{H}) \) of the form described in definition 5.2 such that

\[
\|B(s) - C_\varepsilon(s)(A_0 + \varphi(s))C_\varepsilon^{-1}(s)\| \leq \varepsilon \| s \|, \quad s \in 0^K_\varepsilon.
\]

5.8. Example. Let \( D \) be a diagonal operator in \( \text{HS} \)

\[
D = \text{diag}(\lambda_1, \lambda_2, \ldots) \text{ with } \lambda_i \text{'s complex},
\]

distinct and \( \sum_{i=1}^{\infty} |\lambda_i|^2 < \infty. \)

Then \( Z(D^*) = Z(D) \) is the set of all diagonal operators in \( \text{HS} \). It follows from corollary 5.7 that if \( B \) is a deformation of \( D \) with finite dimensional base \( K \) we have
\[ \|B(s) - C_\varepsilon(s)A(s)C_\varepsilon^{-1}(s)\| \leq \varepsilon\|s\| \quad s \in \Omega^c_K \]

where \( A(s) \) is diagonal for all \( s \in \Omega^c_K \) and \( A(0) = D \).

5.9. Remark. If the space \( H \) is finite dimensional (\( H = \mathbb{C}^n \)) then \( HS = \mathbb{C}^{n \times n} \) and then theorem 5.5 of this section is equivalent to theorem 3.6 of chapter I. Because a weakly versal deformation \( A \) of the matrix \( A_0 \) is transversal to orbit of \( A_0 \) by theorem 5.5, and hence by theorem I.3.6 \( A \) is a versal deformation of the matrix \( A_0 \) in the sense of definition I.2.4. So, if \( H \) is finite dimensional we have

\( A \) is weakly versal \( \iff \) \( A \) is versal.
References


[SCH] Schatten, R., Norm Ideals of Completely Continuous Operators; Springer-Verlag; Berlin (1960).