International Journal of Pure and Applied Mathematics

Volume 53 No. 1 2009, 95-107

DIFFERENTIAL SYSTEMS ON SPACES OF BOUNDED LINEAR OPERATORS

Fernanda Botelho[1 §], Annita Davis[2]

[1,2]Department of Mathematical Sciences
The University of Memphis
373, Dunn Hall, 3721, Norriswood St.,
Memphis, TN 38152-3240, USA
[1]email: mbotelho@memphis.edu
[2]email: adavis2@memphis.edu

Abstract: We study systems of differential equations in $\mathcal{B}(\mathcal{H})$, the Banach space of bounded linear operators on a separable complex Hilbert space $\mathcal{H}$, equipped with the standard operator norm. The systems considered in this paper are infinite dimensional generalizations of mathematical models of learning. We use the polar decomposition of operators to find an explicit form for solutions. We also discuss the usual questions of existence and uniqueness of solutions, as well as their stability behavior.

AMS Subject Classification: 34G20, 47J25
Key Words: nonlinear systems, bounded operators, complex Hilbert spaces, polar decomposition of operators

1. Introduction

Component analysis algorithms are powerful algorithms used to extract relevant features from a large multi-dimensional array of data. Among such algorithms is the well-known Oja and Karhunen's algorithm (see Oja [19]), translated by the $n \times k$ matrix differential equation:

$$\dot{Z} = M\,Z \,-\, Z\,Z^T\,M\,Z,$$

where Z^T denotes the transpose of Z. This system also mimics the evolution of certain intrinsic parameters of an artificial neural network with n input and

Received: April 13, 2009

§Correspondence author

k output neurons. The input correlation matrix M is symmetric, with the $(i, j)-$entry equal to the expected value of the correlation between the inputs into neurons i and j, respectively. The $(i, j)-$entry of the matrix Z represents the connecting weight between the input neuron i and the output neuron j. Connecting weights are responsible for the changes that quantifiable information undergoes while traveling through channels connecting the neurons in a network. Such a device, when presented with a collection of inputs and with an initial assignment for the weights, generates a flow of changes of its connecting weights. It is of great interest to know whether and when such a process stabilizes. In such a case, the resulting outcome is interpreted as the ultimate learning from the exposure of that particular network to a set of inputs. In this paper, we consider the following generalization of the Oja and Karhunen model:

$$\begin{cases} \dot{Z} = M\,Z \,-\, Z\,Z^*\,M\,Z\,, \\ Z(0) = Z_0\,, \end{cases} \tag{1}$$

in which the time dependent variables are bounded linear operators on $\mathcal{B}(\mathcal{H})$. The operator Z^* is the adjoint of Z and M is a normal operator on $\mathcal{H}$. Infinite dimensional systems are better equipped to represent the complexity of human intelligence and hence may perform more realistically through a learning process. We apply the polar decomposition of operators to solve system (1), provided Z_0 is invertible and commutes with M. This approach allows a decomposition of (1) into a scalar system and a polar system. The scalar system is an Euler type equation, for which well-known techniques can be extended in order to derive an explicit form for solutions. The polar system is a first order non autonomous linear differential equation, that can also be solved explicitly. These two components of the solution are combined to define the polar representation for local solutions of (1). This representation allows us to study the long-term and the stability behavior of the flow.

2. Local Existence and Uniqueness of Solutions

In this section we establish the local existence and uniqueness of solutions of system (1). We use a version of the classical fixed point theorems due to Tychonov to prove the existence of a positive number ϵ and a unique differentiable path $Z : (-\varepsilon, \varepsilon) \to \mathcal{B}(\mathcal{H})$ so that $\dot{Z} = MZ - ZZ^*MZ$ and $Z(0) = Z_0$, cf. Botelho et al [6]. We start by recalling the Tychonov's Fixed Point Theorem, as stated in Hartman [15].

Theorem 2.1. *Let $\mathcal{D}$ be a Banach space of elements x, y, ... with norms $|x|, |y|, ...$ and T_0 be a map from the ball $|x| \leq \rho$, in $\mathcal{D}$, into $\mathcal{D}$ satisfying $|T_0[x] - T_0[y]| \leq \theta|x-y|$, for some θ, $0 < \theta < 1$. If $|T_0[0]| \leq \rho(1-\theta)$, then there exists a unique fixed point x_0 of T_0.*

We first recall that given an operator Z in $\mathcal{B}(\mathcal{H})$, we have $\|Z^*\| = \|Z\|$ and $\|Z^*Z\| = \|ZZ^*\| = \|Z\|^2$ (Furuta [11], or Zimmer [24]). Given $\rho > 0$, we set

$$B_\rho(Z_0) = \{Z \in \mathcal{B}(\mathcal{H}) : \ \|Z - Z_0\| \leq \rho\},$$

and we denote by $\mathcal{C}\left([-\varepsilon, \varepsilon], B_\rho(Z_0)\right)$ the space of all continuous functions defined on the interval $[-\varepsilon,\, \varepsilon]$ with values in $B_\rho(Z_0)$. This space is equipped with $\|\cdot\|_\infty$.

Theorem 2.2. *If M and Z_0 are bounded operators on a complex Hilbert space $\mathcal{H}$, then there exist positive numbers ε and ρ, and a unique differentiable map $Z : (-\varepsilon,\, \varepsilon) \to B_\rho(Z_0)$ such that $\dot{Z}(t) = MZ(t) - Z(t)Z^*(t)MZ(t)$ and $Z(0) = Z_0$.*

Proof. The map $T : B_\rho(Z_0) \to \mathcal{B}(\mathcal{H})$, given by

$T(Z) = MZ - ZZ^*MZ$, satisfies a Lipschitz condition, since

$$\begin{aligned}\|TZ_1 - TZ_2\| &= \|MZ_1 - Z_1Z_1^*MZ_1 - (MZ_2 - Z_2Z_2^*MZ_2)\| \\ &\leq \|M\| \ \left(1 + \|Z_1\|^2 + \|Z_1\|\,\|Z_2\|\right)\|Z_1 - Z_2\| \\ &\quad + \|M\|\,\|Z_1 - Z_2\|\,\|Z_2\|^2 \\ &\leq \|M\| \ \left[1 + 3(\|Z_0\| + \rho)^2\right] \ \|Z_1 - Z_2\|,\end{aligned}$$

for $Z_1, Z_2 \in B_\rho(Z_0)$. If we set $\rho = 2\|Z_0\|$, then

$$\|TZ_1 - TZ_2\| \leq \|M\| \ \left[1 + 27\|Z_0\|^2\right] \ \|Z_1 - Z_2\|.$$

We now choose $\varepsilon > 0$ so that $\theta = \|M\| \ \left[1 + 27\|Z_0\|^2\right] \varepsilon < \frac{1}{2}$. We define the function F on $\mathcal{C}\left([-\varepsilon, \varepsilon], B_\rho(Z_0)\right)$ as follows: $F(Z)(t) = Z_0 + \int_0^t T(Z(\xi))\, d\xi$. This function is a contraction, since

$$\begin{aligned}\|F(Z_1) - F(Z_2)\|_\infty &= \left\| \int_0^t (T(Z_1) - T(Z_2))\, d\xi \right\|_\infty \\ &\leq \ \varepsilon\, \|T(Z_1) - T(Z_2)\|_\infty < \theta\|Z_1 - Z_2\|.\end{aligned}$$

Theorem 2.1 asserts that F has a unique fixed point, i.e., there exists $Z \in$

$\mathcal{C}([-\varepsilon,\varepsilon], B_\rho(Z_0))$ such that

$$Z(t) = Z_0 + \int_0^t (MZ(\xi) - Z(\xi)Z^*(\xi)MZ(\xi))\, d\xi.$$

Therefore, $Z(0) = Z_0$ and $\dot{Z}(t) = MZ(t) - Z(t)Z^*(t)MZ(t)$. □

3. The Scalar System

In this section we use the polar decomposition (Ringrose [22]) of operators to construct the "scalar" system associated with (1). We use a generalization of the polar representation of a complex number in order to define a system with solutions equal to the scalar factor of the referred decomposition. We first recall the definition of partial isometry and the polar decomposition for an operator.

Definition 3.1. An operator U on a Hilbert space $\mathcal{H}$ is a *partial isometry* if there exists a closed subspace D of $\mathcal{H}$ such that

$$\|Ux\| = \|x\| \text{ for any } x \in D,$$

$$\text{and } Ux = 0 \text{ for any } x \in D^\perp = \{x \in \mathcal{H} : \langle x, y\rangle = 0 \forall\, y \in D\}.$$

Remark 3.2. It is shown in Furuta [11] that an operator U is a partial isometry if and only if $UU^*U = U$.

Lemma 3.3. (see Furuta [11], or Retherford [21]) *If T is a bounded operator on a Hilbert space $\mathcal{H}$, with the inner product $\langle\, ,\, \rangle$, then:*

1. There exists a unique positive bounded operator S so that $S^2 = T^*T$. The operator S is self-adjoint (i.e. $S^* = S$) and is denoted by $|T|$.

2. $T = P\,|T|$, with P a partial isometry.

The decomposition of T stated in Lemma 3.3(2) is unique and is called the polar decomposition of the operator T. In particular, Lemma 3.3(2) also implies that $Z_0^* = Q_0\sqrt{Z_0 Z_0^*}$, where Q_0 is a partial isometry. Therefore we have that $Z_0 = \sqrt{Z_0 Z_0^*}P_0$, with $P_0 = Q_0^*$. For $t \in (-\varepsilon, \varepsilon)$ we denote by $Z(t)$ a local solution of (1) and we set $V(t) = Z(t)Z(t)^*$. It is a straightforward calculation to verify that $V(t)$ is a local solution of the system

$$\begin{cases} \dot{V} = MV + VM^* - VMV - VM^*V\,, \\ V(0) = Z_0 Z_0^*\,. \end{cases} \tag{2}$$

The following proposition is a consequence of Theorem 2.1. The proof is omitted since it follows similar arguments to those given for Proposition 2.2.

Proposition 3.4. *If M and Z_0 are bounded operators on a complex Hilbert space $\mathcal{H}$, then there exist a positive number ε and a unique differentiable map $V : (-\varepsilon, \varepsilon) \to \mathcal{B}(\mathcal{H})$ such that*

$$\dot{V}(t) = MV(t) + V(t)M^* - V(t)MV(t) - V(t)M^*V(t) \text{ and } V(0) = Z_0Z_0^*.$$

The family $\{V(t)\}_{t\in(-\varepsilon,\,\varepsilon)}$ defines a path of positive operators, since

$$\langle Z(t)Z(t)^*\nu, \nu\rangle = \langle Z^*(t)\nu, Z^*(t)\nu\rangle = \|Z(t)^*\nu\|^2 \geq 0.$$

Lemma 3.5. *If M is a normal operator that commutes with Z_0, then for $t \in (-\varepsilon, \varepsilon)$, $V(t)$ commutes with both M and M^*, and $\{V(t)\}_{t\in(-\varepsilon,\,\varepsilon)}$ is a family of commuting operators.*

Proof. Fuglede-Putman Theorem (cf. Furuta [11], p. 67) asserts that if a normal operator M commutes with an operator Z_0, then M^* also commutes with Z_0. This implies that Z_0 and Z_0^* commute with both M and M^*. Therefore $V_0 = Z_0Z_0^*$ also commutes with M and M^*.

We recall that $V(t)$ satisfies system (2). Piccard's iterative method (see Hartman [15]) asserts that $V(t)$ can be obtained as the limit of a sequence of operators $V_n(t, V_0)$ defined iteratively as follows:

$$V_1(t, V_0) = V_0 \quad \text{and}$$

$$V_{n+1}(t, V_0) = V_0 + \int_0^t [MV_n(\xi, V_0) + V_n(\xi, V_0)M^* - V_n(\xi, V_0)MV_n(\xi, V_0) - V_n(\xi, V_0)M^*V_n(\xi, V_0)]\, d\xi.$$

It follows from Fuglede-Putman Theorem that if W is an operator that commutes with M, then $MW + WM^* - WMW - WM^*W$ commutes with both M and M^*. Since $V_n(t, V_0)$ is a polynomial in t with commuting operators as coefficients, it follows from an induction argument that

$$V_n(t_1, V_0)\, V_n(t_2, V_0) = V_n(t_2, V_0)\, V_n(t_1, V_0), \text{ for } t_1 \text{ and } t_2$$

in the interval $(-\epsilon, \epsilon)$. Therefore $V(t_1)$ and $V(t_2)$ commute. □

Lemma 3.6. *If $Z_0 \in \mathcal{B}(\mathcal{H})$ is invertible, then Z_0^* and $Z_0Z_0^*$ are invertible.*

Proof. Since Z_0 is invertible, then

$$Z_0^*(Z_0^{-1})^* = (Z_0^{-1}Z_0)^* = I = (Z_0Z_0^{-1})^* = (Z_0^{-1})^*Z_0.$$

We also have that $(Z_0Z_0^*)\,[(Z_0^*)^{-1}Z_0^{-1}] = I$ and

$$[(Z_0^*)^{-1}Z_0^{-1}]\,(Z_0Z_0^*) = I.$$ □

If follows from Theorem in Douglas [10] (p. 32) that the set of invertible operators is open in the Banach algebra $\mathcal{B}(\mathcal{H})$. Furthermore if f is in $\mathcal{B}(\mathcal{H})$ and $\|I-f\|<1$, then f is invertible (see Halmos [14]). We assume that $V(t)$ ($|t|<\varepsilon$) is an invertible operator that commutes with M.

The operator $V = ZZ^*$ is positive, for simplicity of exposition we now denote the unique positive square root of V by $V^{1/2}$ (cf. Lemma 3.3). We now list additional properties of the local family of operators $\{V(t)\}_{t\in(-\varepsilon,\,\varepsilon)}$.

Lemma 3.7. *If Z_0 is invertible, M is normal, $M\,Z_0 = Z_0\,M$, and $\{V(t)\}_{t\in(-\varepsilon,\,\varepsilon)}$ is a local solution of (2), then:*

1. $(V^{-1})^* = V^{-1}$.
2. $V^{-1}M^* = M^*V^{-1}$.
3. $\dot{V}V^{-1} = V^{-1}\dot{V}$.
4. $\dot{V}V^{-1/2} = V^{-1/2}\dot{V}$.

Proof. 1. The invertibility of V implies that

$$(V^{-1}V)^* = (VV^{-1})^* = I. \text{ Therefore } (V^{-1})^* = V^{-1}.$$

2. The equation $V^{-1}M^* = M^*V^{-1}$ follows from the commutativity of M and V^{-1}.

3. Since $\dot{V} = MV + VM^* - V(M+M^*)V$ and $VM = MV$, we have that $\dot{V} = MV + M^*V - (M+M^*)V^2 = (M+M^*)V - V(M+M^*)V$. Hence $\dot{V}V^{-2} = (M+M^*)V^{-1} - (M+M^*)$ and $V^{-1}\dot{V}V^{-1} = V^{-1}(M+M^*) - (M+M^*) = (M+M^*)V^{-1} - (M+M^*)$. Therefore $V^{-1}\dot{V}V^{-1} = \dot{V}V^{-2}$, and $V^{-1}\dot{V} = V^{-1}\dot{V}$.

4. The derivative of $VV^{-1/2} = V^{1/2}$ yields

$$\dot{V}V^{1/2} - \frac{1}{2}VV^{-3/2}\dot{V} = \frac{1}{2}V^{-1/2}\dot{V}.$$

Thus, $\dot{V}V^{-1/2} = V^{-1/2}\dot{V}$. □

Proposition 3.8. *Let Z_0 be an invertible operator in $\mathcal{B}(\mathcal{H})$, M normal, and $MZ_0 = Z_0M$. If for $t\in(-\varepsilon,\,\varepsilon)$, $V(t)$ is a solution of the system*

$$\begin{cases} \dot{V} = MV + VM^* - VMV - VM^*V \\ V(0) = Z_0Z_0^*, \end{cases} \tag{3}$$

then

$$V(t) = (I + C\,\exp(-(M+M^*)t))^{-1}, \tag{4}$$

with $C = V^{-1}(0) - I$.

Proof. We select $\varepsilon > 0$ so that $V(t)$ is invertible for $|t| < \varepsilon$. Then we have $\lim_{h\to 0} \frac{V^{-1}(t+h)-V^{-1}(t)}{h} = V^{-1}(t+h)V^{-1}(t)\frac{V(t)-V(t+h)}{h} = -V^{-2}\dot{V}$. Therefore $\frac{d}{dt}(V^{-1}) = -V^{-2}\dot{V}$. This implies that

$$\tfrac{d}{dt}(V^{-1}) = (M + M^*) - (M + M^*)V^{-1}.$$

Hence $V^{-1}(t) = I + C\,\exp(-(M + M^*)t)$, or equivalently

$$V(t) = (I + C\,\exp(-(M + M^*)\,t))^{-1}\,,\ \text{with } C = V^{-1}(0) - I. \qquad \square$$

It is a straightforward calculation to verify that

$$V(t) = (I + C\,\exp(-(M + M^*)t))^{-1} \text{ satisfies (2).}$$

The family $\{V(t)\}_{t\in(m,\,M)}$ is a maximal solution of system (2), provided that:

$$M = \sup\{t :\, I + C\,\exp(-(M + M^*)t) \text{ is invertible}\}$$

and

$$m = \inf\{t :\, I + C\,\exp(-(M + M^*)t) \text{ is invertible}\}.$$

4. The Polar System

In this section we derive the polar system associated with (1). For every $t \in (-\varepsilon,\,\varepsilon)$ we have that $Z(t) = V^{1/2}(t)\,P(t)$, with $V^{1/2}$ representing the unique positive operator so that $V^{1/2}\,V^{1/2} = Z\,Z^*$ and P is a partial isometry. Differentiating the equation

$$Z(t) = V^{1/2}(t)\,P(t), \text{ we obtain } \dot{Z} = V^{1/2}\dot{P} + \tfrac{1}{2}V^{-1/2}\dot{V}P = (I - V)MZ.$$

Therefore $\dot{P} = V^{-1/2}(I-V)MV^{1/2}P - \frac{1}{2}V^{-1}\dot{V}P$. The commutativity of V and M implies that

$$\dot{P} = \left[M - VM - \frac{1}{2}V^{-1}\dot{V}\right]P. \tag{5}$$

We set $A(t) = M - V(t)M - \frac{1}{2}V^{-1}(t)\dot{V}(t)$. Equation (5) now becomes $\dot{P} = A(t)P$. Since $V(t) = [I + C\,\exp(-(M + M^*)t)\,]^{-1}$ we have that $\dot{V} = [I + C\,\exp(-(M+M^*)t)\,]^{-2}C\,\exp(-(M+M^*)t)\,(M+M^*)$, and thus $\dot{V} = V^2C(M+M^*)\exp(-(M + M^*)t)$. Therefore

$$A(t) = M - VM - \frac{1}{2}VC\,\exp(-(M + M^*)t)\,(M + M^*)$$

$$
\begin{aligned}
&= M - V\left(M + \frac{1}{2}C(M+M^*)\exp(-(M+M^*)t)\right)\\
&= M - [I + C\ \exp(-(M+M^*)t)\,]^{-1}\\
&\qquad\qquad \left(M + \frac{1}{2}C(M+M^*)\exp(-(M+M^*)t)\right)\\
&= M - \left(M + \frac{1}{2}C(M+M^*)\exp(-(M+M^*)t)\right)\\
&\qquad\qquad\qquad \left(\sum_{n=0}^{\infty}(-1)^n\, C^n \exp(-n(M+M^*)t)\right)\\
&= \frac{1}{2}(M-M^*)\left(I - \sum_{n=0}^{\infty}(-1)^n C^n\ \exp(-n(M+M^*)t)\right)\\
&= \frac{1}{2}(M-M^*)[I - (I + C\ \exp(-(M+M^*)t))^{-1}]\\
&= -\frac{1}{2}(M-M^*)(V(t)-I).
\end{aligned}
$$

Thus equation (5) reduces to $\dot{P} = -\frac{1}{2}(M-M^*)(V(t)-I)P$. It follows from Lemma (3.5) that for every t_1 and t_2 in the interval $(-\varepsilon, \varepsilon)$, we have that $A(t_1)A(t_2) = A(t_2)A(t_1)$. Therefore

$P(t) = \exp\left(\int_0^t A(\xi)d\xi\right) P_0(\,|t| < \varepsilon)$ is a solution of the polar system

$$
\begin{cases} \dot{P} = \frac{1}{2}(M-M^*)(V(t)-I)\,P\,, \\ P(0) = P_0. \end{cases} \tag{6}
$$

Remark 4.1. For every t, $P(t) = \exp\left(\int_0^t\ A(\xi)d\xi\right) P_0$ is a partial isometry. This follows from Remark 3.2, since

$$
\begin{aligned}
&P(t)\,P(t)^*\,P(t)\\
&= \exp\left(\int_0^t\ A(\xi)d\xi\right) P_0\,P_0^* \exp\left(-\int_0^t\ A(\xi)d\xi\right)\ \exp\left(\int_0^t\ A(\xi)d\xi\right)\ P_0 = P(t).
\end{aligned}
$$

These considerations prove the following proposition.

Proposition 4.2. *If Z_0 is an invertible operator in $\mathcal{B}(\mathcal{H})$, M is normal, $MZ_0 = Z_0M$, then $P(t)$ (for $|t| < \varepsilon$) is a solution of system (6) if and only if*

$$
P(t) = \exp\left(\int_0^t\ A(\xi)d\xi\right) P_0.
$$

Theorem 4.3. *If Z_0 is invertible and commutes with the normal operator*

M, then there exist $\varepsilon > 0$ and a unique differentiable mapping $Z : (-\varepsilon, \varepsilon) \to \mathcal{B}(\mathcal{H})$ such that

$$\dot{Z} = M\,Z \,-\, Z\,Z^*\,M\,Z \text{ and } Z(0) = Z_0, \textit{ if and only if}$$

$$Z(t) = V(t)^{1/2}\;P(t),\;\; V(t) = \left[I + (V_0^{-1} - I)\exp(-(M+M^*)t)\,\right]^{-1}$$

$$\textit{and} \quad P(t) = \exp\left(\int_0^t -\tfrac{1}{2}(M-M^*)(V(\xi)-I)d\xi\right)P_0.$$

Proof. The statement follows from Propositions 3.8, 4.2 and Remark 3.2. □

5. Stability Analysis

In this section we establish the long term behavior of solutions. Theorem (4.3) states that

$$Z(t) = \left[I + (V_0^{-1} - I)\exp(-(M+M^*)t)\right]^{-1/2} \exp\left(\int_0^t -\tfrac{1}{2}(M-M^*)(V(\xi)-I)d\xi\right)P_0,$$

is a solution of (1), provided that $I + (V_0^{-1} - I)\,\exp(-(M+M^*)t)$ is invertible. We consider additional assumptions on M to ensure the existence of solutions for $t \in (-\varepsilon, \infty)$.

Lemma 5.1. *If Z_0 is an invertible operator in $\mathcal{B}(\mathcal{H})$, M is a normal operator that commutes with Z_0, $\|(Z_0\,Z_0^*)^{-1} - I\| < 1$, and the spectrum of $M + M^*$ is strictly positive, then there exists $\epsilon > 0$ so that*

$$I + \left[(Z_0\,Z_0^*)^{-1} - I\right]\,\exp\left(-(M+M^*)\,t\,\right)$$

is invertible on the interval $(-\varepsilon,\,\infty)$ and

$$\lim_{t\to\infty}\,\left[I + ((Z_0\,Z_0^*)^{-1} - I)\,\exp\left(-(M+M^*)\,t\right)\right] = I.$$

Proof. Proposition 3.8 implies that

$$I + [(Z_0\,Z_0^*)^{-1} - I]\,\exp\left(-(M+M^*)\,t\,\right)$$

is invertible on the interval $(-\varepsilon,\,\varepsilon)$. Since the spectrum of $M + M^*$ is strictly positive, i.e. $\sigma(M+M^*) \geq \lambda > 0$, then

$$\|[(Z_0Z_0^*)^{-1} - I]\exp(-(M+M^*)\,t)\| \leq \|(Z_0Z_0^*)^{-1} - I\|\exp(-\lambda t) < 1.$$

Therefore we have $I + [(Z_0\,Z_0^*)^{-1} - I]\,\exp\left(-(M+M^*)\,t\,\right)$ is invertible for $t \in (-\varepsilon,\,\infty)$ and

$$\lim_{t\to\infty}\,\left[I + ((Z_0\,Z_0^*)^{-1} - I)\,\exp\left(-(M+M^*)t\right)\right] = I. \qquad \square$$

In the following proposition we use the logarithmic function of an operator. We refer the reader to Baker [5] and Conway [8] for more details. Given a self-adjoint operator A, the spectral theorem and the spectral calculus allow us to write

$$\log(A) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\,(A-I)^n}{n}, \text{ provided } \|A-I\| < 1.$$

We first recall some basic properties of the log function.

Proposition 5.2. (cf. Baker [5], p. 48) *If* $\|A-I\| < 1$, *then* $\exp(\log(A)) = A$. *If* $\|\exp(A) - I\| < 1$, *then* $\log(\exp(A)) = A$.

Proposition 5.3. *Let Z_0 be an invertible operator in $\mathcal{B}(\mathcal{H})$, M a normal operator that commutes with Z_0, $\|(Z_0\,Z_0^*)^{-1} - I\| < 1$, and the spectrum of $M + M^*$ strictly positive. If $Z(t)$ is the maximal solution of system (1) with $Z(0) = Z_0$, then*

$$\lim_{t\to\infty} Z(t) = \exp\left[\frac{1}{2}(M-M^*)(M+M^*)^{-1}\log((Z_0 Z_0^*)^{-1})\right] P_0,$$

where P_0 is the partial isometry in the polar decomposition of Z_0.

Proof. Theorem 4.3 states that $Z(t) = V(t)^{1/2}P(t)$ with

$$V(t) = \left[I + (V_0^{-1} - I)\,\exp\left(-(M+M^*)\,t\right)\right]^{-1}, \text{ and}$$

$$P(t) = \exp\left(\int_0^t -\frac{1}{2}(M-M^*)(V(\xi)-I)d\xi\right) P_0.$$

Therefore, with

$$C = V_0^{-1} - I, \text{ we have that } \int_0^t -\tfrac{1}{2}(M-M^*)(V(\xi)-I)d\xi$$

$$= -\frac{1}{2}(M-M^*)\int_0^t \left[[I + C\,\exp\left(-(M+M^*)\xi\right)]^{-1} - I\right] d\xi$$

$$= -\frac{1}{2}(M-M^*)\int_0^t \sum_{n=1}^{\infty}(-1)^n C^n \exp\left[-n(M+M^*)\xi\right] d\xi$$

$$= -\frac{1}{2}(M-M^*)\sum_{n=1}^{\infty}\int_0^t (-1)^n C^n \exp\left[-n(M+M^*)\xi\right] d\xi$$

$$= \frac{1}{2}(M - M^*)(M + M^*)^{-1} \left[\sum_{n=1}^{\infty} \frac{[-C \exp[-(M + M^*)t]]^n}{n} - \sum_{n=1}^{\infty} \frac{(-C)^n}{n} \right].$$

Since $\|C \exp(-(M + M^*)t)\| \le \|C\| < 1$, for $t \ge 0$, we have that

$$\sum_{n=1}^{\infty} \frac{[-C \exp[-(M + M^*)\,t\,]]^n}{n} = -\log\{C \exp[-(M + M^*)t] + I\}$$

$$\text{and} \quad \sum_{n=1}^{\infty} \frac{(-C)^n}{n} = -\log(C + I).$$

Therefore we have that

$$\int_0^t -\frac{1}{2}(M - M^*)(V(\xi) - I)d\xi = \frac{1}{2}(M - M^*)(M + M^*)^{-1}$$
$$\times [\log(C + I) - \log[C \exp(-(M + M^*)t) + I\,]].$$

We also observe that

$$\begin{aligned} &\exp[\log(C + I) - \log(C \exp(-(M + M^*)\,t\,) + I)] \\ &\quad = \exp(\log(C + I)) \cdot [\exp[\log(C \exp(-(M + M^*)\,t\,) + I\,)\,]\,]^{-1} \\ &\quad\quad = (C + I)\,[C \exp[-(M + M^*)\,t\,] + I\,]^{-1}, \end{aligned}$$

hence

$$\lim_{t\to\infty} \exp[\log(C + I) - \log[C \exp(-(M + M^*)\,t\,) + I\,]\,] = C + I = V_0^{-1}.$$

For large values of t we have that

$$\|\exp[\log(C + I) - \log[C \exp(-(M + M^*)\,t\,) + I\,]\,] - I\| < 1,$$

and thus

$$\lim_{t\to\infty} \log[\exp[\log(C + I) - \log(C \exp(-(M + M^*)t) + I)]\,] = \log(V_0^{-1}).$$

This implies that

$$\lim_{t\to\infty} P(t) = \lim_{t\to\infty} \exp\left(\int_0^t -\frac{1}{2}(M - M^*)(V(\xi) - I)d\xi \right) P_0$$
$$= \exp\left(\frac{1}{2}(M - M^*)(M + M^*)^{-1} \log(V_0^{-1}) \right) P_0$$

and

$$\lim_{t\to\infty} Z(t) = \exp\left(\frac{1}{2}(M - M^*)(M + M^*)^{-1} \log(V_0^{-1}) \right) P_0\,. \qquad \square$$

Remark 5.4. Under the assumptions listed in Proposition 5.3, if in addition M is self-adjoint then the limit of $Z(t)$ as t goes to infinity is equal to P_0.

References

[1] S. Amari, Dynamics of pattern formation in lateral-inhibition type neural fields, *Biological Cybernetics*, **27** (1977), 77-87.

[2] S. Amari, *Mathematical Theory of Neural Networks*, Sangyo-Tosho, Tokyo (1998).

[3] D. Arrowsmith, C. Place, *An Introduction to Dynamical Systems*, Cambridge University Press (1990).

[4] W. Arverson, *A Short Course on Spectral Theory*, Cambridge University Press (2000).

[5] A. Baker, *Matrix Groups. An Introduction to Lie Group Theory*, Springer Undergraduate Mathematics Series (2001).

[6] F. Botelho, J. Jamison, Differential equations in spaces of compact operators, *J. of Math. Analysis and Appl.*, **308** (2005), 105-120 .

[7] J. Conway, *A Course in Functional Analysis*, Graduate Texts in Mathematics, Springer-Verlag, **96** (1990).

[8] J. Conway, B. Morrel, Roots and logarithms of bounded operators on Hilbert spaces, *Journal of Functional Analysis,* **70** (1987), 171-193.

[9] K. Deimling, Ordinary differential equations in Banach spaces, *Lecture Notes in Mathematics*, Springer-Verlag, **596** (1977).

[10] R. Douglas, *Banach Algebra Techniques in Operator Theory, 2-nd Edition*, Graduate Texts in Mathematics, Springer-Verlag, **179** (1998).

[11] T. Furuta, *Invitation to Linear Operators*, Taylor & Francis. (2001).

[12] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan College Publ. Co. (1994).

[13] K. Kingsley, P. Adams, *Formation of New Connections by a Generalisation of Hebbian Learning* (2001), Preprint.

[14] P. Halmos, *Introduction to Hilbert Spaces*, Chelsea Publishing Company (1957).

[15] P. Hartman, *Ordinary Differential Equations*, Wiley (1964).

[16] J. Hertz, A. Krogh, R. Palmer, Introduction to the theory of neural computation, *A Lecture Notes Volume, Santa Fe Institute Studies in the Sciences of Complexity,* (1991).

[17] A. Mukherjea, K. Pothoven, *Real and Functional Analysis*, Plenum Press, New York (1978).

[18] E. Oja, A Simplified Neuron Model as a Principal Component Analyzer, *J. of Math. Biology,* **15** (1982), 267-273 .

[19] E. Oja, J. Karhunen, A stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix, *J. of Math. Analysis and Appl.,* **106** (1985), 69-84 .

[20] N. Pavel, Differential equations, flow invariance and applications, *Research Notes in Mathematics*, Pitman Publishing Inc., **113** (1994).

[21] J. Retherford, *Hilbert space: Compact operators and the trace theorem*, London Math. Society Student Text, **27**, Cambridge University Press (1994).

[22] J. Ringrose, *Compact Non-Self-Adjoint Operators*, Van Nostrand ReinHold Mathematical Studies, **35** (1994).

[23] R. Schatten, *Norm Ideals of Completely Continuous Operators*, Springer Verlag (1960).

[24] Robert Zimmer, *Essential Results of Functional Analysis*, The University of Chicago Press (1990), 28-29.