ON AN EIGENFLOW EQUATION AND ITS LIE ALGEBRAIC GENERALIZATION*

CHRISTIAN EBENBAUER† AND ALESSANDRO ARSIE‡

Abstract. This paper deals with a dynamical system of the form $\dot{A} = [[N, A^T + A], A] + \nu[[A^T, A], A]$, where A is an $n \times n$ real matrix, N is a constant $n \times n$ real matrix, ν is a positive constant and $[A, B] = AB - BA$. In particular, the purpose of this paper is to establish a sorting behavior of the dynamical system and to represent it in a general Lie algebraic setting. Moreover, some applications of the dynamical system are presented.

Keywords: Control theory, dynamical systems, ordinary differential equations

1. Introduction. Brockett introduced in [6] the so-called double-bracket equation

$$(1.1) \qquad \dot{H} = [[N, H], H],$$

where H is an $n \times n$ real symmetric matrix, N is a constant $n \times n$ real symmetric matrix, and $[A, B] = AB - BA$. This dynamical system has several remarkable properties. For example (1.1) can be used to sort lists or to diagonalize symmetric matrices. In the recent paper [13], a dynamical system of the form

$$(1.2) \qquad \dot{A} = [[N, A^T + A], A] + \nu[[A^T, A], A],$$

where A is an $n \times n$ real matrix, N is a constant $n \times n$ real matrix, and ν is a positive constant, has been introduced. The dynamical system (1.2) has similar properties as (1.1), as will be shown in this paper, but it also diagonalizes and computes eigenvalues of *nonsymmetric* matrices. In the case of A being symmetric, the self-commutator $[A^T, A]$ vanishes and (1.2) reduces to (1.1). Thus, the flow in the space of nonsymmetric matrices described by (1.2) can be considered as a generalization of the flow in the space of symmetric matrices described by (1.1). Due to the ability to simultaneously compute all eigenvalues of nonsymmetric matrices, one may call (1.2) eigenflow equation.

The motivation to design and study dynamical systems like (1.1) and (1.2) has several roots. For example, solving computational problems with the help of continuous-

---

time dynamical systems can be seen as a way to design and implement analog algorithms. Analog algorithms and analog computation have been investigated in various fields or research, including neuroscience, biology, informatics, mathematics [3, 14, 26, 27, 30]. One main research goal in this area is to build devices which can perform massively parallel computations and/or which function with minimal power supply. These are abilities which are mastered very well in living organisms, like in the human brain, but which have not been successfully realized yet in terms of digital computation. Analog computation is one alternative to digital computation, with promising potentials to achieve this goal. Another reason stems from the wish to obtain a new way to design numerical algorithms and to analyze their behavior and the underlying geometry of these algorithms [2, 9–12, 17, 29]. For example, in order to apply ideas from the theory of dynamical systems and differential geometry, it is often easier to study continuous algorithms rather then discrete algorithms, i.e. algorithms which are described by differential equations rather then by difference equations.

The purpose of this paper is to analyze and to design dynamical systems of the form (1.2). In particular, the purpose is to provide an abstract version of (1.2) in a Lie algebraic setting and to establish a sorting property of (1.2) analogous to (1.1). Moreover, some applications of (1.2) in the area of analog computation are presented.

The content of this paper is as follows: In Section 2, some preliminary results are summarized. In Section 3, the dynamical system (1.2) as well as the underlying idea behind (1.2) is generalized to a Lie algebraic setting. Moreover, an analogous sorting behavior of (1.2) as known from (1.1) is established. In Section 4, applications of the dynamical system (1.2) are discussed, including sorting roots of polynomials and the spectral factorization of polynomials. Finally, a summary of the results is given in Section 5.

Notation: Let $A = (a_{ij}) = (A)_{ij} \in \mathbb{R}^{n \times n}$ be a real $n \times n$ matrix. Then $\lambda(A) = \{\lambda_1, \ldots, \lambda_n\}$ denotes the spectrum of A and $<A, B> = \text{trace}(A^T B)$, $\|A\|_F^2 = \text{trace}(A^T A)$ the Frobenius norm, where $\text{trace}(A) = a_{11} + \ldots + a_{nn}$ and A^T denotes the transposed of A. Moreover, A^* denotes the conjugate transposed of a complex matrix A. $N = \text{diag}(n_1, \ldots, n_n)$ denotes a diagonal matrix with diagonal elements n_i, $i = 1 \ldots n$ and $I, 0$ denote an identity matrix respectively a zero matrix of appropriate dimension. Furthermore, $\pi \in S_n$ denotes a permutation of the set $\{1, \ldots, n\}$. Let V be a finite dimensional vector space over a field K, then $\mathfrak{gl}(V)$ ($\mathfrak{sl}(V)$) denotes the space of all endomorphisms of V (with zero trace). Moreover, $\mathfrak{gl}_n(\mathbb{R})$ ($GL_n(\mathbb{R})$, $\mathfrak{sl}_n(\mathbb{R})$) is the set of real $n \times n$ matrices (with nonzero determinant, with zero trace). Let $W : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function, then ∇W denotes the gradient (row vector) of W.

2. Preliminaries. In the following, some auxiliary results are summarized. Consider the dynamical system

$$\dot{A} = [U(A), A], \tag{2.1}$$

where A is a real $n \times n$ matrix, $U : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ is a continuous matrix-valued function, and $[A, B] = AB - BA$ is the commutator. One of the most important properties of dynamical systems in Lax form, i.e. dynamical systems of form (2.1), is the fact that they preserve the spectrum of $A(t)$, i.e. the spectrum of $A(t)$ is equal to the spectrum of $A(0)$ for any $t \in [0, T_{sup})$, where $[0, T_{sup})$ is the maximal interval of existence of the solution $A = A(t)$, see e.g. [25, 28, 32].

THEOREM 1. *The flow described by* (2.1) *is isospectral.*

In order to analyze and design flows like (2.1), it is sometimes convenient to consider U as a control input. In doing so, isospectrality means then that the spectrum of $A(t)$ is invariant under any feedback $U = U(A)$ and (1.1) and (1.2) can be considered as feedback systems defined by the control system (2.1) and by the feedbacks

$$U(A) = [N, A], \tag{2.2}$$

$$U(A) = [N, A^T + A] + \nu[A^T, A], \tag{2.3}$$

respectively (see Fig. 1). Later on, this control systems point of view is used to analyze the convergence behavior of vector fields of type (2.1) on Lie algebras.
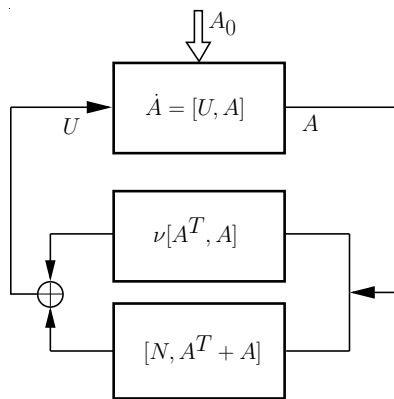

FIG. 1. *Equation* (1.2) *represented as feedback loop.*

The following lemma is a standard result about invariant sets (see e.g. [24], Chapter 2, Theorem 5.2).

LEMMA 2. *Let* $\Omega(x(0))$ *denote the positive limit set* (ω-*limit set*) *of a bounded solution* $x = x(t)$ *of*

$$\dot{x} = f(x), \tag{2.4}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous. Then $\Omega(x(0))$ is nonempty, compact, connected, (positively and negatively) invariant with respect to (2.4), and it is the smallest closed set that $x = x(t)$ approaches as $t \to \infty$, i.e. if $x = x(t)$ converges to a closed (compact) set which contains $\Omega(x(0))$, then $x = x(t)$ converges to $\Omega(x(0))$.

The following lemma is Chetaev's instability theorem [15, 22].

LEMMA 3. Let x_E be an equilibrium point of (2.4) and let $W : \mathcal{D} \to \mathbb{R}$ be a continuously differentiable function on a neighborhood $\mathcal{D}$ of $x = x_E$, such that $W(x_E) = 0$. Suppose that the set $U = \{x \in \mathcal{D} : \|x - x_E\| < r, W(x) > 0\}$ is nonempty for every $r > 0$. If $\dot{W}(x) > 0$ in $U \setminus \{x_E\}$, then the equilibrium point $x = x_E$ is unstable.

The next result is taken from [20] (Corollary 1) and states, roughly speaking, that an equilibrium point x_E in a positively invariant set $\mathcal{E}$, where $\mathcal{E}$ itself is assumed to be locally asymptotically stable, is locally asymptotically stable if and only if the equilibrium point x_E is locally asymptotically stable on $\mathcal{E}$, which means $\forall \delta > 0 \; \exists \epsilon > 0$ such that if $\|x(0) - x_E\| \leq \epsilon$ and $x(0) \in \mathcal{E}$ then $\|x(t) - x_E\| \leq \delta$ and $\lim_{t \to \infty} x(t) = x_E$.

LEMMA 4. Let x_E be an equilibrium point of (2.4) and let $\mathcal{D} \subset \mathbb{R}^n$ be a neighborhood of x_E. Assume that there exists a continuously differential function $V : \mathcal{D} \to \mathbb{R}$ such that $V(x) \geq 0$, $V(x_E) = 0$, and $\dot{V}(x) = \nabla V(x) f(x) \leq 0$ on $\mathcal{D}$. Then x_E is asymptotically stable if and only if x_E is locally asymptotically stable on the largest positively invariant set contained in $\mathcal{E} = \{x \in \mathcal{D} : \dot{V}(x) = 0\}$.

Notice that Lemma 4 is also useful in combination with Chetaev's instability theorem on invariant sets $\mathcal{E}$, since an equilibrium point x_E is asymptotically stable if and only if x_E is asymptotically stable on $\mathcal{E}$. Hence, by showing instability on $\mathcal{E}$ using Lemma 3, instability of the equilibrium point follows from Lemma 4.

3. Main Results.

3.1. The convergence behavior of (1.2). Brockett obtained the double-bracket equation (1.1) by recasting a continuous-time steepest descent algorithm for solving a least-squares matching problem in a Lie algebraic setting [6–8]. As already mentioned, (1.1) has many interesting properties, including the ability to diagonalize symmetric matrices, to sort lists, or to solve various combinatorial optimization problems. The double-bracket equation is also related to other well-known dynamical systems, like the Toda lattice flow or projected gradient flows, see [4, 5, 10]. The literature concerning the double-bracket equation and gradient flows has been rapidly growing. For example, [17] provides a good introduction to the double-bracket equation and other dynamical systems, e.g. the QR flow or Oja's flow, which can be used to solve computational problems in an analog fashion (see also [2, 11]).

Notice that "to compute" or "to solve" by means of continuous-time dynamical systems is here understood in the following way: Starting with a certain initial condition (input data), the solution of the computation (output data) is given by

the equilibrium point to which the flow converges. Thus, the computation process (solving process, analog algorithm) is described by the flow of the dynamical system.

Since (1.1) can diagonalize symmetric matrices and compute its eigenvalues, it is natural to search for a dynamical system which can do the same for nonsymmetric matrices. This was the motivation which has led to the dynamical system (1.2). In contrast to (1.1), the basic idea behind (1.2) is not based on a gradient flow (steepest descent) argument but rather then on an orthogonality argument. However, (1.2) is naturally related to (1.1), since for symmetric initial data, i.e. $A(0) = A(0)^T$, the flows defined by (1.1) and (1.2) are equivalent.

The next theorem summarizes the convergence behavior of (1.2) established in [13].

THEOREM 5. *For any initial condition $A(0) = A_0 \in \mathbb{R}^{n \times n}$, the solution $A = A(t)$ of (1.2) is well-defined for all $t \geq 0$ and it converges to a set of normal matrices which have the same spectrum as A_0. Moreover, if N is a diagonal matrix with pairwise distinct diagonals and if A_0 is a matrix with eigenvalues $\lambda_i = \sigma_i + i\omega_i$, $i = 1 \ldots n$, which have pairwise distinct real parts except for complex conjugate pairs[1], then the solution $A = A(t)$ of (1.2) converges to an equilibrium point $A(\infty)$, i.e.*

$$(3.1) \qquad\qquad [A(\infty)^T, A(\infty)] = 0,$$

$$(3.2) \qquad\qquad [N, A(\infty)^T + A(\infty)] = 0.$$

In particular, there exist $n!$ (isolated) equilibrium points $A(\infty)$ and they are of the form

$$(3.3) \qquad (A(\infty))_{ij} = \begin{cases} \sigma_{\pi(i)} & \text{if } i = j \\ \pm\omega_{\pi(i)} & \text{if } i \neq j \text{ and } \sigma_{\pi(i)} = \sigma_{\pi(j)} \\ 0 & \text{else} \end{cases},$$

where $(\pi(1), \ldots, \pi(n))$ is a permutation of $(1, \ldots, n)$.

Proof. The steps of the proof are as follows: Step 1: It is shown that any solution $A = A(t)$ of (1.2) satisfy $[A(t)^T, A(t)] \to 0$ for $t \to \infty$, which means $A(t)$ converges to the set of normal matrices, that is the set of matrices which satisfy $A = U^*\Lambda U$, $U^*U = I$. Step 2: It is shown that $[N, A(t)^T + A(t)] \to 0$ for $t \to \infty$. Step 3: It is shown that there exist $n!$ equilibrium points and (3.3) is satisfied.

Step 1: In a first step, it is shown that the derivative of the function

$$(3.4) \qquad V(A) = \|A\|_F^2 - \sum_{i=1}^{n} |\lambda_i|^2 = \text{trace}(A^T A) - \sum_{i=1}^{n} |\lambda_i|^2$$

is monotonically decreasing along the flow (1.2) with the initial condition $A(0) = A_0$ as long as $[A(t)^T, A(t)] \neq 0$. Notice that differentiating the sum of squares of eigenvalue

---

[1] That means two eigenvalues of the form $a + ib$, a or $a + ib, a + ic$, $b \neq -c$ are not allowed.

in (3.4) with respect to (2.1) is zero due to isospectrality. Moreover, it is easy to see that (3.4) is zero if and only if A is normal, i.e. $[A^T, A] = 0$, because $\|A\|_F^2$ is the sum of squares of singular values of A [19]. Differentiating (3.4) with respect to (2.1) and using the facts $\text{trace}(AB) = \text{trace}(BA)$, $\text{trace}(A^T) = \text{trace}(A)$ one obtains:

$$(3.5) \qquad \dot{V}(A) = -2\text{trace}([A^T, A][N, A + A^T]) - 2\nu\text{trace}([A^T, A]^2).$$

Now observe that $\text{trace}([A^T, A][N, A + A^T]) = 0$. This follows from the fact that $[A^T, A]$ is symmetric and $[N, A + A^T]$ is skewsymmetric and the trace of a product between a symmetric matrix and a skewsymmetric matrix is zero. Thus with $\text{trace}([A^T, A]^2) = \|[A^T, A]\|_F^2 \geq 0$ one gets

$$(3.6) \qquad \dot{V}(A) = -2\nu\|[A^T, A]\|_F^2 < 0 \quad \text{as long as} \quad [A^T, A] \neq 0.$$

Therefore, all solutions $A = A(t)$ of (1.2) converge into the set where the self-commutator vanishes, i.e.

$$(3.7) \qquad \lim_{t\to\infty} [A(t)^T, A(t)] = 0.$$

Notice that all solutions are bounded, since the derivative (3.6) coincide with the derivative of (the positive definite and radially unbounded function) $\|A\|_F^2$ along (2.1).

Step 2: In a second step, it is shown that the function

$$(3.8) \qquad W(A) = \frac{1}{2}\text{trace}(N(A + A^T))$$

is monotonically increasing along all solution $A = A(t)$ of (1.2) starting in the positively invariant set defined by $\mathcal{N}_{\lambda(A(0))} = \{A \in \mathbb{R}^{n\times n} : \lambda(A) = \lambda(A(0)) \wedge [A^T, A] = 0\}$ as long as $[N, A(t)^T + A(t)] \neq 0$, i.e. the set of normal matrices which have the same spectrum as $A(0)$. Moreover, it is shown that any solution of $A = A(t)$ of (1.2) converges to an equilibrium point given by $[A(\infty)^T, A(\infty)] = 0$, $[N, A(\infty)^T + A(\infty)] = 0$.

First, notice that the set $\mathcal{N}_{\lambda(A(0))}$ is positively invariant w.r.t. (1.2), because

$$(3.9) \qquad \frac{d}{dt}[A^T(t), A(t)]|_{t=0} = [[U(A_0), A_0]^T, A_0] + [A_0^T, [U(A_0), A_0]] = 0$$

holds for $[A_0^T, A_0] = 0$, $A_0 = A(0)$, and $U(A)$ given by (2.3) In particular, (3.9) follows from $U(A_0)^T = -U(A_0)$ and the Jacobi identity, i.e. $[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0$. Differentiating (3.8) with respect to (1.2) (i.e. (2.1),(2.3)) and using $\text{trace}(AB) = \text{trace}(BA)$, $\text{trace}(A^T) = \text{trace}(A)$, $[A, B] = -[B, A]$, $[A, B]^T = [B^T, A^T]$ one obtains:

$$
\begin{aligned}
\dot{W}(A) = &\frac{1}{2}\text{trace}(N[U, A]) + \frac{1}{2}\text{trace}(N[A^T, U^T]) \\
= &\frac{1}{2}\text{trace}([A + A^T, N][N, A + A^T]) \\
&+ \nu\text{trace}([A, N][A^T, A]).
\end{aligned}
$$
$$(3.10)$$

Now, from (3.10) it follows that

$$(3.11) \qquad \dot{W}(A) = -\frac{1}{2}\text{trace}([A + A^T, N]^2) > 0 \quad \text{as long as} \quad [A + A^T, N] \neq 0$$

and $A \in \mathcal{N}_{\lambda(A_0)}$. From Lemma 2 and Step 1 it follows that $\Omega(A(0))$ is a compact and invariant set in $\mathcal{N}_{\lambda(A_0)}$. As will be shown in Step 3, there exists a finite number of *isolated* (equilibrium) points $A(\infty)$ which simultaneously satisfy $\dot{V}(A) = 0$ and $\dot{W}(A) = 0$. In order to show that $\Omega(A(0))$ is an equilibrium point, make the following observations.

Observation 1: If $\Omega(A(0))$ is a single point, then it is an equilibrium point.

Observation 2: If $\Omega(A(0))$ is contained in a level set of W, i.e. $W(\Omega(A(0))) = m \in \mathbb{R}$, then $\Omega(A(0))$ is a single point. Assume $\Omega(A(0))$ is not a single point. Then, since $\Omega(A(0))$ is connected and since the points $A(\infty)$ (where $\dot{W} = 0$) are isolated, there exists a point $P \in \Omega(A(0))$ such that $\dot{W}(P) > 0$. Hence W is increasing along a solution passing through P, but on the other hand $\Omega(A(0))$ is invariant and contained in a level set of W. Contradiction.

Observation 3: See also Figure 2. Assume now $\Omega(A(0))$ is not contained in a level set of W, i.e. $W(\Omega(A(0))) = [\underline{m}, \overline{m}] \subset \mathbb{R}$, $\underline{m} < \overline{m}$ and $A_1(\infty), \ldots, A_k(\infty)$, $k > 1$, are equilibrium points in $\Omega(A(0))$ ordered in such a way that the following holds: $W(A_1(\infty)) = \ldots = W(A_m(\infty)) > W(A_{m+1}(\infty)) \geq \ldots \geq W(A_k(\infty))$. The claim is that if a solution $A = A(t)$ is sufficiently close to one of the equilibria $A_j(\infty)$, $j \leq m$, let's say at $t_1 > 0$, then $A = A(t)$ cannot enter anymore a suitably chosen neighborhood of $W(A_l(\infty))$, $l \geq m + 1$, for $t \geq t_1$.

By (3.11), $\dot{W} > 0$ on $\mathcal{N}_{\lambda(A_0)} \setminus \{A_1(\infty), \ldots, A_k(\infty)\}$. Let $\mathcal{B}(\epsilon)$ be a closed neighborhood of $\Omega(A(0))$ such that $\text{dist}(\mathcal{B}(\epsilon), \Omega(A(0))) \leq \epsilon$. Notice that dist denotes the (Hausdorff) distance between sets and $\mathcal{B}(\epsilon)$ is a subset of the A_0-isospectral set, i.e. the ambient space of $\mathcal{N}_{\lambda(A_0)}$ is considered to be the set $\{B \in \mathbb{R}^{n \times n} : \lambda(B) = \lambda(A_0)\}$ where (2.1) flows. Moreover, let $\mathcal{U}_i$ be open neighborhoods for $A_i(\infty)$ and let $\mathcal{U}_i(\epsilon) = \mathcal{B}(\epsilon) \cap \mathcal{U}_i$. Denote with $\underline{b}_j = \inf_{A \in \mathcal{U}_j} W(A)$, and with $\overline{b}_l = \sup_{A \in \mathcal{U}_l} W(A)$. Choose now ϵ and neighborhoods $\mathcal{U}_i$, $i = 1, \ldots, k$, such that $\min_{j=1\ldots m} \underline{b}_j > \max_{l=m+1\ldots k} \overline{b}_l$ and such that

$$(3.12) \qquad \min_{A \in \mathcal{B}(\epsilon) \setminus \cup_{i=1}^{k} \mathcal{U}_i(\epsilon)} \dot{W}(A) \geq \frac{\delta}{2} > 0$$

holds for some $\delta > 0$. Notice that it is always possible to choose ϵ and neighborhoods $\mathcal{U}_i$ such that $\min_{j=1\ldots m} \underline{b}_j > \max_{l=m+1\ldots k} \overline{b}_l$, since W is continuous and $W(A_j(\infty)) > W(A_l(\infty))$, $j \leq m$, $l \geq m + 1$. Moreover

$$(3.13) \qquad \min_{A \in \Omega(A(0)) \setminus \cup_{i=1}^{k} \mathcal{U}_i(0)} \dot{W}(A) \geq \delta > 0$$

can be always be satisfied, because $\Omega(A(0)) \setminus \cup_{i=1}^{k} \mathcal{U}_i(0)$ is a compact set and $\dot{W}$ is strictly positive on $\Omega(A(0)) \setminus \cup_{i=1}^{k} \mathcal{U}_i(0)$. Hence, because W and $\dot{W}$ are continuous

functions, (3.12) must be true for a sufficiently small ϵ. In other words, for ϵ sufficiently small, the set $\{A : \ \dot{W}(A) \leq 0, \ A \in \mathcal{B}(\epsilon)\}$ is contained in $\cup_{i=1}^{k} \mathcal{U}_i(\epsilon)$.

Since $\Omega(A(0))$ is the positive limit set, there exists a $t_1 > 0$ such that $A(t_1) \in \cup_{j=1}^{m} \mathcal{U}_j(\epsilon)$ and such that $A(t) \in \mathcal{B}(\epsilon)$ for $t \geq t_1$, because of Lemma 2 (observe that $B(\epsilon)$ is closed and compact). Moreover, there must exists a $t_2 > t_1$ such that $A(t_2) \in \cup_{l=m+1}^{k} \mathcal{U}_l(\epsilon)$. However, this is impossible because whenever $\max_{l=m+1...k} \overline{b}_l < W(A(t)) < \min_{j=1...m} \underline{b}_j$, then $\dot{W} > 0$ and therefore $W(A(t)) \geq \min_{j=1...m} \underline{b}_j$ for $t \geq t_1$. But in order to reach some neighborhood $\mathcal{U}_l(\epsilon)$, $l \geq m + 1$, $W(A(t)) < \max_{l=m+1...k} \overline{b}_l$ must hold. (This leads to a contradiction, because $\min_{j=1...m} \underline{b}_j > \max_{l=m+1...k} \overline{b}_l$.) Thus, $\Omega(A(0))$ cannot contain equilibrium points $A_1(\infty)$, ..., $A_k(\infty)$, $k > 1$ with $W(A_1(\infty)) = \ldots = W(A_m(\infty)) > W(A_{m+1}(\infty)) \geq \ldots \geq W(A_k(\infty))$.

Observation 4: Assume again $\Omega(A(0))$ is not contained in a level set of W, i.e. $W(\Omega(A(0))) = [\underline{m}, \overline{m}] \subset \mathbb{R}$, $\underline{m} < \overline{m}$ and $A_1(\infty), \ldots, A_k(\infty)$, $k \geq 0$, are equilibrium points in $\Omega(A(0))$ but now $W(A_1(\infty)) = \ldots = W(A_k(\infty))$. Then there exist points $P_1, P_2 \in \Omega(A(0))$ such that $W(P_1) > W(P_2)$. If $k \geq 1$, choose $P_1 = A_1(\infty)$ or $P_2 = A_1(\infty)$. Using the same arguments as in Observation 3, where P_1, P_2 play now the role of $A_m(\infty), A_{m+1}(\infty)$, one can conclude that points $P_1, P_2 \in \Omega(A(0))$ such that $W(P_1) > W(P_2)$ holds, can not exist. Thus, $\Omega(A(0))$ must be contained in a level set W.

In summary, Observation 1 and 2 imply that $\Omega(A(0))$ is a single equilibrium point, if $\Omega(A(0))$ is contained in a level set of W and Observation 3 and 4 assure that $\Omega(A(0))$ is indeed contained in a level set of W.

Therefore, one can finally conclude that

$$(3.14) \qquad \lim_{t \to \infty} [N, A(t)^T + A(t)] = 0.$$

Thus, (3.7) together with (3.14) and the Observation 1-4 imply that any solution $A = A(t)$ converges to an equilibrium point defined by (3.1) and (3.2).

Step 3: In order to show that (3.3) holds, observe first that

$$(3.15) \qquad [N, A(\infty)^T + A(\infty)] = 0 \Leftrightarrow A(\infty)^T + A(\infty) = 2D$$

with $D = \mathrm{diag}(d_1, \ldots, d_n)$ diagonal and $N = \mathrm{diag}(n_1, \ldots, n_n)$, $n_i \neq n_j$, $i \neq j$. In particular, $([N, A + A^T])_{ij} = (a_{ij} + a_{ji})(n_j - n_i) = 0$ if and only if $a_{ij} + a_{ji} = 0$, $i \neq j$. Thus, (3.14) implies that $A(t)^T + A(t)$ converges to a real diagonal matrix $2D$. Moreover, from (3.15) follows that

$$(3.16) \qquad A(\infty) = D + S,$$

where $S = (s_{ij})$ is a skewsymmetric matrix. Due to (3.7), $A(\infty)$ is a normal matrix,

i.e. $A(\infty) = U^* \Lambda_\pi U$, $U^* U = I$, Λ_π diagonal. Hence, one obtains

$$
\begin{aligned}
\text{(3.17)} \qquad 2D &= A(\infty) + A(\infty)^T = A(\infty) + A(\infty)^* \\
&= U^*(\Lambda_\pi + \Lambda_\pi^*)U = 2U^* \Re\{\Lambda_\pi\}U.
\end{aligned}
$$

Because of the isospectral property of (1.2) (Theorem 1), the diagonals of D are the real parts of the eigenvalues of $A(0)$, i.e. $D = \mathrm{diag}(d_1, \ldots, d_n) = \Re\{\Lambda_\pi\} = \mathrm{diag}(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(n)})$. Thus, what is left is to reveal the structure of S in (3.16). Observe first that the form (3.16) and the fact that $A(\infty)$ is a normal matrix imply

$$
\text{(3.18)} \qquad [A(\infty)^T, A(\infty)] = 0 \Leftrightarrow [S, D] = 0 \Leftrightarrow s_{ij}(d_j - d_i) = 0,
$$

because $0 = [A(\infty)^T, A(\infty)] = [D - S, D + S] = 2[D, S]$ and $([D, S])_{ij} = s_{ij}(d_j - d_i)$. Notice moreover that due to the assumption that $A(0)$ has eigenvalues with pairwise distinct real parts except for complex conjugate pairs, there does not exist three pairwise distinct indices i, j, k such that $d_i = d_j = d_k$ ($\sigma_i = \sigma_j = \sigma_k$). Thus, one can distinguish between two cases: Case 1 (complex eigenvalue): If $(A(\infty))_{ii} = d_i = \sigma_{\pi(i)} = d_j$ for some i and $j(\neq i)$ in $\{1, \ldots, n\}$, then the *only* element in the ith column/row and jth row/column of S which may not vanish is s_{ij} ($s_{ji} = -s_{ij}$). In particular, for all k different from i and from j, (3.18) implies $d_j - d_k \neq 0$ and $d_i - d_k \neq 0$ and thus $s_{kj} = 0$ and $s_{ki} = 0$ for all $k \neq i, j$. Case 2 (real eigenvalue): If for some i, $d_i = \sigma_{\pi(i)} \neq d_k$ for all $k \neq i$ then (3.18) implies that the ith column/row of S vanishes. By writing down $S = (s_{ij})$ and by taking into account that the eigenvalues of $A(0)$ are identical with the eigenvalues of $A(\infty)$, it can be easily observed that the non-vanishing entries in S must be equal to $s_{ij} = \pm \omega_{\pi(i)}$. Thus, the structure of $A(\infty)$ is given by (3.3).

Finally, notice that $\dot{V}$ (see (3.6)) and $\dot{W}$ (see (3.11)) is zero if and only if (3.1) and (3.2) is satisfied. Moreover, under the assumptions made on N and A_0, any equilibrium point $A(\infty)$ must be of the form (3.3). Thus, taking into account that the eigenvalues of $A(\infty)$ must coincide with the eigenvalues of A_0, it can be observed that there exists exactly $n!$ different configurations $A(\infty)$ which satisfy (3.3). Since there is only a finite number of equilibrium points, they must be isolated. $\qquad\square$

In case the eigenvalues of $A(0)$ have not pairwise distinct real parts, then Theorem 5 does not say anything if the solution converges to an equilibrium point or not. However, all solutions of (1.2) converge to $\mathcal{N}_{\lambda(A_0)}$, see (3.7). Moreover, numerical simulations support that all solutions converge to an equilibrium point and it seems that similar arguments (Observation 1 to 4) also apply to this case. This is a future research point. Furthermore, if the spectrum of $A(0)$ is real, then (3.3) implies that $A(\infty)$ is diagonal. Notice also that the assumption on $A(0)$ having pairwise distinct real parts is *generically* true. Hence, Theorem 5 states that for almost all initial conditions $A_0 \in \mathbb{R}^{n \times n}$, $A(\infty)$ is of the form (3.3) (see examples in Section 4). Thus, to the authors best knowledge, (1.2) is the only available differential equation in
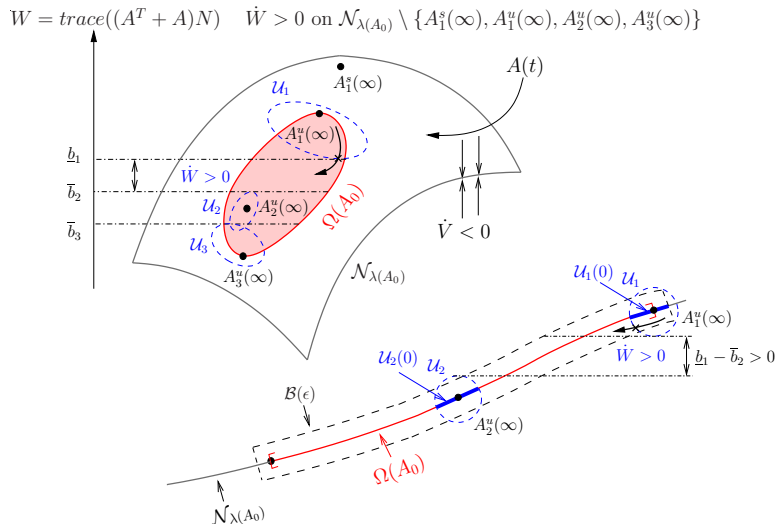
$W = trace((A^T + A)N)$ $\dot{W} > 0$ on $\mathcal{N}_{\lambda(A_0)} \setminus \{A_1^s(\infty), A_1^u(\infty), A_2^u(\infty), A_3^u(\infty)\}$

FIG. 2. *Basic idea (top) and a detailed illustration (bottom) of Observation 3 in the proof of Theorem 5 for $m = 1$. The ambient space in these figures is considered to be the A_0-isospectral set.*

the literature which allows to simultaneously compute (with guaranteed convergence) all the eigenvalues for generically any real matrix. In [29] convergence for general nonsymmetric matrices has been proved in the case of a real spectrum. If the spectrum is complex, the QR flow proposed in [29] has almost periodic solutions. The proof of Theorem 5, however, allows to obtain many other flows with the same convergence properties as (1.2). For example, the assumption in (1.2) that ν is a positive constant can be easily relaxed to $\nu = \nu(t) \geq \nu_0 > 0$.

3.2. A sorting property of (1.2). The double-bracket equation (1.1) can be seen as a dynamical system which solves the following optimization problem in an analog fashion:

$$(3.19) \qquad \Theta^o = \arg \max_{\Theta^T \Theta = I} \; trace(N\Theta^T H_0 \Theta),$$

where $N = diag(n_1, \ldots, n_n)$ is a real diagonal matrix with pairwise distinct diagonals and H_0 is a symmetric matrix with eigenvalues σ_i, $i = 1 \ldots n$. A result which dates back to von Neumann [31] (see also [6–8]) shows that the optimal solution Θ^o satisfies $(\Theta^o)^T H_0 \Theta^o = diag(\sigma_{\pi(1)} \ldots \sigma_{\pi(n)})$, where the diagonals $\sigma_{\pi(i)}$ are arranged in such a way that $trace(N(\Theta^o)^T H_0 \Theta^o) = n_1 \sigma_{\pi(1)} + \ldots + n_n \sigma_{\pi(n)}$ is maximized. Moreover, the relation between (3.19) and (1.1) is the following: $H(\infty) = (\Theta^o)^T H_0 \Theta^o$ for almost all initial data $H_0 = H(0) = H(0)^T$. Notice that a solution $H = H(t)$ of (1.1) may end up in $n!$ different equilibrium points, which are given by the $n!$ permutations $diag(\sigma_{\pi(1)} \ldots \sigma_{\pi(n)})$. However, as shown in [8], only one equilibrium point is asymptotically stable, namely that equilibrium point $H(\infty)$ which maximizes $trace(NH(\infty))$.

Since the optimal equilibrium point is the only stable one, (1.1) solves (3.19) "almost always". Exceptional cases are for example if one starts with a diagonal matrix H_0 (see [6–8] for details). The sorting behavior of (1.1) follows now from the fact that solving (3.19) can be seen as a way of sorting lists (eigenvalues). For example, if $0 \leq n_1 < n_2 < \ldots < n_n$, then the eigenvalues of $H(\infty) = \mathrm{diag}(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(n)})$ are arranged in such a way that $\sigma_{\pi(1)} \leq \sigma_{\pi(2)} \leq \ldots \leq \sigma_{\pi(n)}$.

From the above considerations, it is natural to ask if the same sorting behavior is still present in (1.2), i.e. is it true that an asymptotically stable equilibrium point $A(\infty)$ maximizes $\mathrm{trace}(NA(\infty))$? In the following, it is shown that this is indeed the case.

To prove the desired result, one needs the tangent space of a manifold which is described by the intersection of an isospectral set, lets say defined by a matrix Y (Y-isospectral manifold), with the set of normal matrices.

LEMMA 6. *Assume $Y \in \mathbb{R}^{n \times n}$ complex diagonalizable. Then the intersection of the Y-isospectral manifold with the set of normal matrices, i.e.*

$$(3.20) \qquad \mathcal{N}_{\lambda(Y)} = \{B \in \mathbb{R}^{n \times n} : \ \lambda(B) = \lambda(Y) \ \wedge \ [B^T, B] = 0\},$$

is a manifold and the tangent space $T_A \mathcal{N}_{\lambda(Y)}$ at $A \in \mathbb{R}^{n \times n}$ of $\mathcal{N}_{\lambda(Y)}$ is given by

$$(3.21) \qquad T_A \mathcal{N}_{\lambda(Y)} = \{[U, A] \in \mathbb{R}^{n \times n} : \ U = -U^T \in \mathbb{R}^{n \times n}\}.$$

Proof. Notice that the algebraic variety of all normal matrices has singularities and is not a manifold. However, the intersection of the set of normal matrices with the Y-isospectral manifold defined by a real $n \times n$ normal matrix Y is a (compact connected) embedded submanifold in the space of matrices. This follows from the fact that if Y is a normal matrix, then every other normal matrix B with $\lambda(B) = \lambda(Y)$ can be written as $B = \Theta Y \Theta^T$, Θ orthogonal (see [18], p.105, Theorem 2.5.8). Hence

$$(3.22) \qquad \mathcal{N}_{\lambda(Y)} = \{B \in \mathbb{R}^{n \times n} : \ B = \Theta Y \Theta^T, \ \Theta \text{ orthogonal}\},$$

i.e. $\mathcal{N}_{\lambda(Y)}$ is an orbit of the normal matrix Y under the action of the orthogonal group by conjugation.

By Proposition 8.1 in [17] (see p.355ff) it follows that this orbit is an embedded submanifold in the space of matrices. Moreover, from (3.22) directly follows that the tangent vectors at A are of the form $[U, A]$, $U = -U^T$, i.e. for a smooth curve $\Theta = \Theta(t)$, $\Theta(0) = I$, $t \in (-\epsilon, +\epsilon)$, $\dot{\Theta}(t) = U(t)\Theta(t)$, $U(t) = -U(t)^T$, one obtains $\dot{B}(t)|_{t=0} = [U(0), A] \in T_A \mathcal{N}_{\lambda(Y)}$ for $B(t) = \Theta(t)A\Theta(t)^T$. □

THEOREM 7. *Let $N = \mathrm{diag}(n_1, \ldots, n_n)$ be a real diagonal matrix with pairwise distinct diagonals, i.e. $n_i \neq n_j$, $i \neq j$, and let $A(0) = A_0$ be a matrix with eigenvalues $\lambda_i = \sigma_i + i\omega_i$, $i = 1 \ldots n$, which have pairwise distinct reals parts except for complex conjugate pairs. Then 2^m out of $n!$ equilibrium points are locally asymptotically stable, where m is the number of conjugate complex pairs of eigenvalues. All the other*

equilibrium points are unstable. The locally asymptotically stable equilibrium points are (strict) global maxima of the function $\mathrm{trace}(NA)$ *on* $\mathcal{N}_{\lambda(A_0)}$, *i.e. if* $A(\infty)$ *is locally asymptotically stable, then* $\mathrm{trace}(NA(\infty)) = n_1\sigma_{\pi(1)} + \ldots + n_n\sigma_{\pi(n)} \geq \mathrm{trace}(NA)$ *for all* $A \in \mathcal{N}_{\lambda(A_0)}$.

Proof. The claims in the theorem are proved in 2 steps. Step 1: An equilibrium point $A(\infty)$ which does not maximize $\mathrm{trace}(NA(\infty))$ on $\mathcal{N}_{\lambda(A_0))}$ is unstable. Step 2: An equilibrium point $A(\infty)$ which maximizes $\mathrm{trace}(NA(\infty))$ on $\mathcal{N}_{\lambda(A_0)}$ is locally asymptotically stable. There exist 2^m locally asymptotically stable equilibrium points.

Step 1: It is shown that if an equilibrium point $A(\infty)$ on $\mathcal{N}_{\lambda(A_0)}$ does not maximize $\mathrm{trace}(NA(\infty)) = n_1\sigma_{\pi(1)} + \ldots + n_n\sigma_{\pi(n)}$, then $A(\infty)$ is not a local maximum of the function

$$(3.23) \qquad W(A) = \frac{1}{2}\mathrm{trace}(N(A + A^T))$$

on $\mathcal{N}_{\lambda(A_0)}$. Since $\dot{W}$ is positive *definite* in a neighborhood on $\mathcal{N}_{\lambda(A_0)}$, because the equilibrium points are isolated, it follows by Chetaev's instability theorem (Lemma 3) that $A(\infty)$ is unstable on $\mathcal{N}_{\lambda(A_0)}$ and thus by Lemma 4 unstable.

It is shown now that the gradient of $W : \mathcal{N}_{\lambda(A_0)} \to \mathbb{R}$ vanishes at $A(\infty)$ and that the Hessian is negative semidefinite only if $A(\infty)$ maximizes $\mathrm{trace}(NA(\infty))$.

The directional derivative of W is given by

$$(3.24) \qquad dW(A) = \frac{1}{2}\mathrm{trace}(N(A' + A'^T))$$

with $A' \in T_{A(\infty)}\mathcal{N}_{\lambda(A_0)}$, i.e.

$$(3.25) \qquad dW(A) = \mathrm{trace}(N[U, A])$$

with $U = -U^T$. Because of $A(\infty) = D + S$, D diagonal, S skewsymmetric, $[N, D] = 0$, and $[N, S]$ symmetric, dW vanishes at $A(\infty)$:

$$(3.26) \qquad dW(A(\infty)) = \mathrm{trace}(N[U, D + S]) = 0.$$

The quadratic form associated to the Hessian of W is given by

$$(3.27) \qquad d^2W(A) = \mathrm{trace}(N[U, A']) = \mathrm{trace}(N[U, [U, A]]).$$

Hence, one obtains $d^2W(A(\infty)) = \mathrm{trace}(N[U, [U, D + S]])$.

$$(3.28) \qquad d^2W(A(\infty)) = \mathrm{trace}(N([U, [U, D]]) = \mathrm{trace}([N, U][U, D]).$$

Notice that $[N, U]$ is symmetric and by direct calculation one obtains $([N, U])_{ij} = u_{ij}(n_i - n_j)$, $i \leq j$. Analogous, $([U, D])_{ij} = -u_{ij}(d_i - d_j)$, $i \leq j$. Therefore,

$$(3.29) \qquad d^2W(A(\infty)) = -\sum_{i<j}(n_i - n_j)(d_i - d_j)u_{ij}^2.$$

Thus, in order for (3.29) to be negative semidefinite,

$$(3.30) \qquad (n_i - n_j)(d_i - d_j) \geq 0$$

must hold. Condition (3.30) is called similarly ordered and implies that trace(ND), $D = \mathrm{diag}(d_1, \ldots, d_n) = \mathrm{diag}(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(n)})$ is maximized [16], see also [6–8]. For example, in case $n_1 > n_2 > \ldots > n_n$, $d_1 \geq d_2 \geq \ldots \geq d_n$ follows immediately from (3.30). In other words, if (3.30) is not satisfied, i.e. $W(A(\infty))$ is not a maximum, then $A(\infty)$ is unstable.

Step 2: It is shown that all equilibrium points $A(\infty)$ which maximize trace $(NA(\infty))(= \text{trace}(ND))$ on $\mathcal{N}_{\lambda(A_0)}$ are locally asymptotically stable. Notice first that the maxima on $\mathcal{N}_{\lambda(A_0)}$ are indeed equilibrium points. Suppose $A_1^o(\infty), \ldots, A_s^o(\infty)$, $s \leq n!$ are distinct equilibrium points which all satisfy (3.30) and consequently $W(A_1^o(\infty)) = \ldots = W(A_s^o(\infty))$. Now, let $\Gamma \in \mathcal{N}_{\lambda(A_0)}$, $\Gamma \neq A_i^o(\infty)$, $i \in \{1 \ldots s\}$, be a real normal matrix i.e. $\Gamma = U^* \Lambda_\pi U$, $U^* U = I$, Λ_π diagonal. Then the eigenvalues of $\Gamma^T + \Gamma = 2U^* \Re\{\Lambda_\pi\} U$ are $2\Re\{\Lambda_\pi\}$. Therefore, for $\Gamma^T + \Gamma$ not diagonal, it follows that

$$(3.31) \quad W(\Gamma) = \frac{1}{2}\text{trace}(N(\Gamma^T + \Gamma)) < \frac{1}{2}\text{trace}(N(A_i^o(\infty)^T + A_i^o(\infty))) = W(A_i^o(\infty)),$$

because W is maximized only if $\Gamma^T + \Gamma$ is diagonal [31] (see also (3.19), [6, 8]). On the other hand, if $\Gamma^T + \Gamma$ is diagonal, then Γ must coincide with an equilibrium point because $\Gamma^T + \Gamma$ diagonal and $\Gamma \in \mathcal{N}_{\lambda(A_0)}$ implies that Γ has the same structure (eq. (3.3)) as an equilibrium point $A_i(\infty)$, $i \in \{1, \ldots, n!\}$ (same arguments as in the proof of Theorem 5 - see (3.18) and below). Hence, the maxima of (3.23) on $\mathcal{N}_{\lambda(A_0)}$ are $A_1^o(\infty), \ldots, A_s^o(\infty)$.

Consequently, there exists a small neighborhood $\mathcal{U} \subset \mathcal{N}_{\lambda(A_0)}$ of $A_i^o(\infty)$ such that $\Gamma^T + \Gamma$ is not diagonal for any $\Gamma \in \mathcal{U} \setminus \{A_i^o(\infty)\}$, and hence $W(\Gamma) < W(A_i^o(\infty))$ for all $\Gamma \in \mathcal{U} \setminus \{A_i^o(\infty)\}$. Therefore, W has a strict maximum at $A_i^o(\infty)$ (i.e. $W(A) - W(A_i^o(\infty))$ is negative definite in the neighborhood $\mathcal{U}$) and since $\dot{W}$ is positive definite on $\mathcal{U}$, it follows that $A_i^o(\infty)$ is an asymptotically stable equilibrium point on $\mathcal{N}_{\lambda(A_0)}$ and by Lemma 4 (use V given by (3.4) where $\mathcal{E} = \mathcal{N}_{\lambda(A_0)}$ and the domain $\mathcal{D}$ of V as well as the ambient space of $\mathcal{E}$ is considered to be the A_0-isospectral manifold), an asymptotically stable equilibrium point on the A_0-isospectral manifold.

Finally, to count the number of stable equilibrium points, consider (3.3). Since for an asymptotically stable equilibrium point the diagonal of $A(\infty)$ is fixed due to (3.30), the number of asymptotically stable equilibrium points is given by the number of equilibrium points generated by switching the sign of the imaginary parts ($\pm\omega_{\pi(i)}$). It can be easily seen that the number of possible configurations is 2^m, where m is the number of conjugate complex pairs of eigenvalues.

$\square$

REMARK 8. *Notice that in Theorem 7 locally asymptotically stability of an equilibrium point refers to an locally asymptotically stable equilibrium point on the A_0-isospectral manifold where (1.2) flows. Notice also that if $A(0)$ has a real spectrum, then there exists only one stable equilibrium point, analogous to the double-bracket equation [6–8]. However, if $A(0)$ has complex conjugate eigenvalues, then there exist several stable equilibrium points. For example, for a $3{\times}3$ matrix $A(0)$ with eigenvalues $a \pm bi, c,\ c > a$ and $N = \mathrm{diag}(1,2,3)$ two of six equilibrium points, i.e.*

$$(3.32) \qquad A_1^o(\infty) = \begin{bmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & c \end{bmatrix}, \ A_2^o(\infty) = \begin{bmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & c \end{bmatrix},$$

are asymptotically stable since $\mathrm{trace}(NA_1^o) = \mathrm{trace}(NA_2^o)$ *is maximal in these two cases.*

Summarizing, Theorem 7 allows to interpret (1.2) as a dynamical system which solves the following optimization problem:

$$(3.33) \qquad \begin{aligned} T^o &= \arg \max_{T \in \mathrm{GL}_n(\mathbb{R})} \ \mathrm{trace}(NT^{-1}A_0T) \\ &\quad s.t. \ \left[(T^{-1}A_0T)^T, T^{-1}A_0T\right] = 0 \end{aligned}$$

with $A(0) = A_0$ as in Theorem 7 and $A(\infty) = (T^o)^{-1}A_0T^o$ asymptotically stable.

3.3. A Lie algebraic setting for (1.2)**.** In the following, a Lie algebraic generalization of (1.2) is derived. In order to give a Lie algebraic interpretation of the flow described by (1.2), first we will restrict to the case of semisimple Lie algebras [1, 21, 23]. By definition, a semisimple Lie algebra over a field of characteristic zero, like the real or complex fields, is a Lie algebra with the property that its Killing form $\beta(A, B) = \mathrm{trace}(\mathrm{ad}_A\mathrm{ad}_B)$ is nondegenerate.

Recall also that for a semisimple Lie algebra $\mathfrak{g}$ the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, $A \mapsto \mathrm{ad}_A(.) = [A, .]$ is faithful, so it enables to realize the abstract Lie algebra $\mathfrak{g}$ as a concrete Lie algebra of matrices, once a basis for $\mathfrak{g}$ has been chosen.

DEFINITION 9. *Consider a semisimple Lie algebra $\mathfrak{g}$ over the real field. A decomposition of $\mathfrak{g}$ into direct sum of subspaces*

$$(3.34) \qquad\qquad\qquad \mathfrak{g} = \mathfrak{t} \oplus \mathfrak{p},$$

such that

$$(3.35) \qquad\qquad [\mathfrak{t}, \mathfrak{t}] \subset \mathfrak{t}, \quad [\mathfrak{t}, \mathfrak{p}] \subset \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{t}$$

and such that the Killing form

$$(3.36) \qquad \beta(A, B) = \mathrm{trace}(\mathrm{ad}_A\mathrm{ad}_B) = \begin{cases} \text{negative definite on } \mathfrak{t} \\ \text{positive definite on } \mathfrak{p} \end{cases}$$

is called Cartan decomposition. The projections from $\mathfrak{g}$ to $\mathfrak{p}$ and $\mathfrak{t}$ are denoted by $\pi_{\mathfrak{p}}$ and $\pi_{\mathfrak{t}}$, respectively.

Notice that the Killing form is non-degenerate on $\mathfrak{g}$ and $\mathfrak{t}$ is a subalgebra of $\mathfrak{g}$, while $\mathfrak{p}$ is not. Moreover, $\mathfrak{t}$ and $\mathfrak{p}$ are mutually orthogonal with respect to the Killing form.

To elucidate the relationship with the flow (1.2), consider the specific example of $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{R})$, $n \geq 2$, i.e. the semisimple Lie algebra of real $n \times n$ matrices with zero trace. In this case, it is easy to see that the Cartan decomposition is given by the set $\mathfrak{t} = \mathfrak{so}_n(\mathbb{R})$ of skewsymmetric matrices and the set $\mathfrak{p}$ of symmetric matrices with zero trace. Moreover, the commutation relations among $\mathfrak{t}$ and $\mathfrak{p}$ are nothing else then a restatement of the well-known commutation relations between symmetric and skewsymmetric matrices, i.e. $[\text{skew}, \text{sym}] = \text{sym}$, $[\text{sym}, \text{sym}] = \text{skew}$, $[\text{skew}, \text{skew}] = \text{skew}$. Since these commutation relations are fundamental in the analysis of the flow (1.2), it is no surprise that the setup of a Cartan decomposition for a real semisimple Lie algebra is the right framework in order to generalize (1.2).

In a first step, it is necessary to clarify what is an appropriate analog of a transposition of a matrix in a semisimple Lie algebra. Consider matrices A and B in $\mathfrak{sl}_n(\mathbb{R})$ and consider the transposition map $(.)^T$. It is immediate to see that $[A, B]^T = [B^T, A^T]$, so T is an anti-automorphism of real matrix Lie algebras of order 2, since $((A)^T)^T = A$.

For a real semisimple Lie algebra $\mathfrak{g}$ one can define the following map:

LEMMA 10. *Given a real semisimple Lie algebra $\mathfrak{g}$ with a Cartan decomposition $\mathfrak{t} \oplus \mathfrak{p}$. Then the map $\tau : \mathfrak{g} \to \mathfrak{g}$ given by*

$$(3.37) \qquad \tau(A) = \pi_{\mathfrak{p}}(A) - \pi_{\mathfrak{t}}(A)$$

satisfies $\tau([A, B]) = [\tau(B), \tau(A)]$, i.e. τ is an anti-automorphism of $\mathfrak{g}$ of order 2.

Notice that τ is essentially a Cartan involution. Observe also that $\mathfrak{t}$ and $\mathfrak{p}$ are eigenspaces for τ, with eigenvalues -1 and 1 respectively. Using τ one can endow $\mathfrak{g}$ with an Euclidean scalar product, which is later used to define Lyapunov functions in a proper way:

LEMMA 11. *Given a real semisimple Lie algebra $\mathfrak{g}$ with a Cartan decomposition $\mathfrak{t} \oplus \mathfrak{p}$ and let τ be given by (3.37). Then the pairing $(.,.) : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ given by*

$$(3.38) \qquad (A, B) = \text{trace}(\text{ad}_{\tau(A)}\text{ad}_B)$$

is an Euclidean scalar product.

Before discussing how the operator τ is a generalization of the transpose operator $(.)^T$ acting on matrices, an intrinsic definition of the transposition is given: Given a linear map $L : V \to W$ between to Euclidean scalar product vector spaces $(V, < .,. >_V)$ and $(W, < .,. >_W)$. The transpose $L^T : W \to V$ is defined as the unique linear

map such that $< L^T(w), v >_V = < w, L(v) >_W$ for any $v \in V$, $w \in W$. Therefore, to define the transposition of a vector space endomorphism, it is necessary to endow the vector space with a scalar product (unless one introduces the dual of the vector space).

LEMMA 12. *Let $\mathfrak{g}$ be a real semisimple Lie algebra with a Cartan decomposition $\mathfrak{t} \oplus \mathfrak{p}$ and a scalar product (3.38). Then, the following diagram is commutative:*

(3.39)

$$
\begin{array}{ccc}
\mathfrak{g} & \xrightarrow{\ \tau(.)\ } & \mathfrak{g} \\
{\scriptstyle \mathrm{ad}}\Big\uparrow & & {\scriptstyle \mathrm{ad}}\Big\uparrow \\
\mathfrak{gl}(\mathfrak{g}) & \xrightarrow{\ (.)^T\ } & \mathfrak{gl}(\mathfrak{g})
\end{array}
$$

i.e. for any $A \in \mathfrak{g}$, $\mathrm{ad}_{\tau(A)} = (\mathrm{ad}_A)^T$, where the transposition is defined considering $\mathfrak{g}$ equipped with (3.38). Moreover, ad_A is symmetric if and only if $A \in \mathfrak{p}$; it is skewsymmetric if and only if $A \in \mathfrak{t}$.

The preceding lemmas can be found in standard books about Lie theory (see e.g. [1, 21, 23]). Using the results established above, one can generalize the isospectral flow studied in the previous section. Consider the following vector field on $\mathfrak{g}$:

(3.40)
$$\dot{\mathrm{ad}}_A = \mathrm{ad}_{[U(A),A]},$$

where $A \in \mathfrak{g}$, $U : \mathfrak{g} \to \mathfrak{g}$. Notice that $\mathrm{ad}_{[U(A),A]} = [\mathrm{ad}_{U(A)}, \mathrm{ad}_A]$. Notice also, since $\mathfrak{g}$ is a Lie algebra, the flow described by equation (3.40) will evolve in $\mathfrak{g}$, whenever $U(A) \in \mathfrak{g}$. Observe that (3.40) represents an evolution on $\mathfrak{g}$, but it is a matrix differential equation like (2.1) due to the adjoint representation. One of the most important properties of (3.40) is that it preserves the spectrum of any solution $\mathrm{ad}_A = \mathrm{ad}_{A(t)}$ (see Theorem 1). However, in the following an intrinsic form of (3.40) is considered, namely:

(3.41)
$$\dot{A} = [U(A), A],$$

where $A \in \mathfrak{g}$, $U : \mathfrak{g} \to \mathfrak{g}$, $\mathfrak{g}$ is a real semisimple Lie algebra with a Cartan decomposition $\mathfrak{t} \oplus \mathfrak{p}$ and a scalar product (3.38). By intrinsic is meant that while (2.1) is a matrix differential equation, (3.41) is a differential equation in a finite-dimensional vector space over $\mathbb{R}$ (Lie algebra) and the elements $A \in \mathfrak{g}$ are vectors and not linear operators per se. However, the *isospectrality* property of (3.41) is intrinsic, in the sense that any realization of (3.41) as a matrix differential equation, for example through the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}_n(\mathbb{R})$, i.e. (3.40), is isospectral. Notice also that (3.41) can be rewritten as

(3.42)
$$\pi_{\mathfrak{p}}(\dot{A}) = [\pi_{\mathfrak{t}}(U(A)), \pi_{\mathfrak{p}}(A)] + [\pi_{\mathfrak{p}}(U(A)), \pi_{\mathfrak{t}}(A)]$$
$$\pi_{\mathfrak{t}}(\dot{A}) = [\pi_{\mathfrak{t}}(U(A)), \pi_{\mathfrak{t}}(A)] + [\pi_{\mathfrak{p}}(U(A)), \pi_{\mathfrak{p}}(A)]$$

with $A = \pi_{\mathfrak{p}}(A) + \pi_{\mathfrak{t}}(A)$, $U(A) = \pi_{\mathfrak{p}}(U(A)) + \pi_{\mathfrak{t}}(U(A))$. This will help in the following to simplify the calculations in the upcoming proofs for Theorem 13, 14.

Recall now that (1.2) can be considered as feedback system given by the control system (2.1) and the feedback (2.3). An important point of the feedback (2.3) is, that the two feedback components $[N, A^T + A]$ and $[A^T, A]$ are orthogonal, in the sense that the first component is skewsymmetric while the second one is symmetric. This is important (but not necessary) in order to establish the convergence behavior stated in Theorem 5. In order to emphasize this fact, the generalization of Theorem 5 is carried out in two steps, first by specifying the symmetric component of the feedback only (Theorem 13), and then by specifying also the skewsymmetric component (Theorem 14). In particular, in a first step, the feedback for (3.41) is chosen to be

$$(3.43) \qquad\qquad U(A) = \pi_{\mathfrak{t}}(U(A)) + \nu[\tau(A), A],$$

where ν is a positive constant. Notice that

$$(3.44) \qquad\qquad \pi_{\mathfrak{p}}(U(A)) = \nu[\tau(A), A] = 2\nu[\pi_{\mathfrak{p}}(A), \pi_{\mathfrak{t}}(A)] \in \mathfrak{p}.$$

It is shown in Theorem 13 that this choice of feedback implies that the solution of (3.41),(3.43) will converge into the set of normal elements, i.e. $[\tau(A), A] = 0$, independently of the choice of $\pi_{\mathfrak{t}}(U(A))$. In a second step, the $\pi_{\mathfrak{t}}$-component of the feedback (3.43) is specified as

$$(3.45) \qquad\qquad \pi_{\mathfrak{t}}(U(A)) = [N, \pi_{\mathfrak{p}}(A)],$$

where N is a fixed element belonging to $\mathfrak{p}$. With this feedback, an analogous convergence result as derived in Theorem 5 is established in Theorem 14.

THEOREM 13. *For any initial condition $A(0) = A_0 \in \mathfrak{g}$, the solution $A = A(t)$ of (3.41),(3.43) converges to the set of normal elements, i.e.*

$$(3.46) \qquad\qquad \lim_{t \to \infty} [\tau(A(t)), A(t)] = 0.$$

Proof. Consider the following semidefinite function on $\mathfrak{g}$:

$$(3.47) \qquad\qquad V(A) = (\pi_{\mathfrak{t}}(A), \pi_{\mathfrak{t}}(A)).$$

First, it is shown that V is monotonically decreasing under the flow of (3.41), (3.43), as long as the element $A = A(t)$ is not normal. Taking the time derivative of (3.47) along $A = A(t)$, one gets

$$(3.48) \qquad\qquad \dot{V}(A) = 2(\pi_{\mathfrak{t}}(\dot{A}), \pi_{\mathfrak{t}}(A)).$$

Now, using (3.42) and (3.44), it follows that

$$(3.49) \qquad\qquad \dot{V}(A) = -4\nu\big([\pi_{\mathfrak{p}}(A), \pi_{\mathfrak{t}}(A)], [\pi_{\mathfrak{p}}(A), \pi_{\mathfrak{t}}(A)]\big) < 0.$$

This inequality is true as long as A is not normal, i.e. $[\tau(A), A] = 2[\pi_{\mathfrak{p}}(A), \pi_{\mathfrak{t}}(A)] \neq 0$, due to the fact that $(.,.)$ is a (positive definite) scalar product on $\mathfrak{g}$. Therefore, since the function (3.47) is bounded from below by zero, all bounded solutions of (3.41),(3.43) converges to the set of normal elements (use Lemma 2), i.e.

$$(3.50) \qquad \lim_{t\to\infty}[\tau(A(t)), A(t)] = 0.$$

Finally, in order to prove that the limit (3.50) is well-defined for any initial condition, even though the function V is only semidefinite, one can reason as follows. Suppose $\big(A(t), A(t)\big) \to \infty$, for $t \to T_{\sup}$, where $T_{\sup}$ is the supremum of the maximal interval of existence of a solution for (3.41),(3.43), that is suppose that the solution $A = A(t)$ will blow up in finite time. On the other hand, it is clear from what has been proved so far that V stays bounded as $t \to T_{\sup}$. Now, rewrite V as

$$(3.51) \qquad V(A) = (\tau(A) - A, \tau(A) - A)$$

and this is equal to

$$(3.52) \qquad \begin{aligned} V(A) &= \operatorname{trace}(\operatorname{ad}_{\tau(\tau(A)-A)}\operatorname{ad}_{\tau(A)-A}) \\ &= 2(A, A) - \operatorname{trace}((\operatorname{ad}_A)^2) - \operatorname{trace}((\operatorname{ad}_{\tau(A)})^2). \end{aligned}$$

Observe that $\operatorname{trace}((\operatorname{ad}_A)^2)$ is constant, since the time evolution of ad_A is isospectral, and the same is true for the term $\operatorname{trace}((\operatorname{ad}_{\tau(A)})^2)$, because $\operatorname{ad}_{\tau(A)} = (\operatorname{ad}_A)^T$ by Lemma 12. Therefore, the proper function (A, A) has to remain bounded and thus the solution exists for any future time. $\qquad\square$

THEOREM 14. *If the spectrum of ad_N is distinct and the spectrum of ad_{A_0} has pairwise distinct real parts, then the solution $A = A(t)$ converges to an equilibrium point $A(\infty)$ with*

$$(3.53) \qquad [\tau(A(\infty)), A(\infty)] = 0,$$

$$(3.54) \qquad [N, \pi_{\mathfrak{p}}(A(\infty))] = 0$$

and $\lambda(ad_{A_0}) = \lambda(ad_{A(\infty)})$.

Proof. Notice that (3.53) and (3.54) define an equilibrium point. Since Theorem 13 implies (3.53), it remains to show (3.54), i.e.

$$(3.55) \qquad \lim_{t\to\infty}[N, \pi_{\mathfrak{p}}(A(t))] = 0.$$

To show this, consider the function

$$W(A) = (\eta, \pi_{\mathfrak{p}}(A))$$

and its time derivative along the flow described by (3.41):

$$(3.56) \qquad \dot{W}(A) = (N, \pi_{\mathfrak{p}}(\dot{A})).$$

Now, using (3.42), (3.43), and (3.45), it follows that

$$(3.57) \qquad \dot{W}(A) = ([N, \pi_{\mathfrak{p}}(A)], [N, \pi_{\mathfrak{p}}(A)]) + 2\nu([N, \pi_{\mathfrak{t}}(A)], [\pi_{\mathfrak{t}}(A), \pi_{\mathfrak{p}}(A)]).$$

From the Theorem 13 follows that any solution $A = A(t)$ exists for all $t \geq 0$ and it is bounded (in the norm induced by the scalar product introduced in $\mathfrak{g}$). Moreover, it has been shown in Theorem 13 that any solution converges to the invariant set of normal elements $\mathcal{N} = \{A \in \mathfrak{g} : [\tau(A), A] = 2[\pi_{\mathfrak{t}}(A), \pi_{\mathfrak{p}}(A)] = 0\}$, hence the positive limit set $\Omega(A_0)$ lies in $\mathcal{N}$. Furthermore, for solutions $A = A(t)$ in $\mathcal{N}$, it follows from (3.57) that

$$(3.58) \qquad \dot{W}(A) = ([N, \pi_{\mathfrak{p}}(A)], [N, \pi_{\mathfrak{p}}(A)]) > 0,$$

whenever $[N, \pi_{\mathfrak{p}}(A)] \neq 0$. Finally, using the same arguments as in the proof of Theorem 5 (see equation (3.9) to equation (3.14)), it can be shown that the set of normal elements $\{A \in \mathfrak{g} : [\tau(A), A] = 2[\pi_{\mathfrak{t}}(A), \pi_{\mathfrak{p}}(A)] = 0\}$ is invariant and that any solution $A = A(t)$ satisfies

$$(3.59) \qquad \lim_{t \to \infty} [N, \pi_{\mathfrak{p}}(A(t))] = 0.$$

Notice that the equations (3.53) and (3.54) have a finite number of solutions because of the facts the a semisimple Lie algebra has a faithful adjoint representation and that any real Lie algebra $\mathfrak{g}$ can be realized as a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{R})$ (Ado's Theorem, [21], Ch. VI). Furthermore, notice that the assumption that ad_N must be diagonal is not necessary, because it can be always reduced to the diagonal case. $\square$

REMARK 15. *In the setup of a real semisimple Lie algebra with general elements N and A_0, the element $A(\infty)$ in Theorem 14 cannot be further specified. But by making further assumptions, sharper statements are possible, like in Theorem 5, which can be considered as a special case of Theorem 13 and 14 with $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{R})$.*

REMARK 16. *Obviously, Theorem 5 applies to general matrices with nonzero trace (A_0 is an element in $\mathfrak{gl}_n(\mathbb{R})$ but not in $\mathfrak{sl}_n(\mathbb{R})$). Since $\mathfrak{gl}_n(\mathbb{R})$ is not semisimple, it is natural to ask to what extent is it possible to generalize the previous results to non-semisimple Lie algebras? One basic idea is to shift the trace, for example by $S(A_0) = A_0 - n^{-1}\mathrm{trace}(A_0)I$. Then $S(A_0) \in \mathfrak{sl}_n(\mathbb{R})$ and $A(\infty) = S(A(\infty)) + n^{-1}\mathrm{trace}(A_0)I$. This Ansatz can be used to establish analogous results for non-semisimple Lie algebras. First, observe that if $\mathfrak{g}$ is semisimple, then $\mathrm{ad} : \mathfrak{g} \hookrightarrow \mathfrak{sl}(\mathfrak{g})$, because $\mathrm{trace}(\mathrm{ad}_A) = 0$ for any $A \in \mathfrak{g}$ as is immediate to see. Therefore, any semisimple Lie algebra is isomorphic to a Lie subalgebra of $\mathfrak{sl}_n$, for some n. On the other hand, any real Lie algebra $\mathfrak{g}$ can be realized as a Lie algebra of matrices (Ado's Theorem, [21], Ch. VI). Let $\mathfrak{g}$ by any real Lie algebra, and consider a faithful representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$, where V is a real vector space. Observe that the trace map gives a fibration: $\mathrm{trace} : \mathfrak{gl}(V) \to \mathbb{R}$, sending a matrix $\rho(A)$ to the value of its trace. Moreover, $\mathfrak{sl}(V)$ is simply the inverse*

image of 0 and it is a hyperplane in $\mathfrak{gl}(V)$. Instead, $L_a := \text{trace}^{-1}(a)$ is not a Lie algebra, it is a hyperplane and it is a translate of $\mathfrak{sl}_n$ (any matrix A with trace a will translate $\mathfrak{sl}_n$ to $\text{trace}^{-1}(a)$). Lets call s_{-a} the translation sending L_a to $L_0 = \mathfrak{sl}(V)$. Now the equation

$$(3.60) \qquad \dot{\rho}(A) = [\rho(U(A)), \rho(A)]$$

is isospectral, so if A_0 is the initial condition, then the corresponding flow evolves in $L_{\text{trace}(A_0)}$. Using the translation $s_{-\text{trace}(A_0)}$ the flow can be also translated to L_0 which is a semisimple Lie algebra and where the previous construction applies. Notice that in the case of a non-semisimple Lie algebra $\mathfrak{g}$, more data are required: the choice of a faithful representation ρ and the choice of the initial condition that will determine the translation to L_0.

Summarizing, Theorem 13 and 14 establish the convergence behavior of the flow described by

$$(3.61) \qquad \begin{aligned} \pi_{\mathfrak{p}}(\dot{A}) &= [[N, \pi_{\mathfrak{p}}(A)], \pi_{\mathfrak{p}}(A)] + 2\nu[[\pi_{\mathfrak{p}}(A), \pi_{\mathfrak{t}}(A)], \pi_{\mathfrak{t}}(A)] \\ \pi_{\mathfrak{t}}(\dot{A}) &= [[N, \pi_{\mathfrak{p}}(A)], \pi_{\mathfrak{t}}(A)] + 2\nu[[\pi_{\mathfrak{p}}(A), \pi_{\mathfrak{t}}(A)], \pi_{\mathfrak{p}}(A)], \end{aligned}$$

$N \in \mathfrak{p}$, for an abstract Lie algebra $\mathfrak{g}$ endowed with a Cartan decomposition $\mathfrak{t} \oplus \mathfrak{p}$. Notice that for $[\pi_{\mathfrak{p}}(A), \pi_{\mathfrak{t}}(A)] = 0$, $\pi_{\mathfrak{p}}(\dot{A}) = [[N, \pi_{\mathfrak{p}}(A)], \pi_{\mathfrak{p}}(A)]$, which is the double-bracket equation. Moreover, the results in this subsection provide a framework for several variations of flows of this type.

4. Applications.

4.1. Sorting Eigenvalues and Roots of Polynomials. Consider the polynomial

$$(4.1) \qquad p(x) = x^5 - 9x^4 + 41x^3 - 103x^2 + 128x - 78$$

with the corresponding companion matrix

$$A(0) = \begin{bmatrix} 9 & -41 & 103 & -128 & 78 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The roots of the polynomial p (eigenvalues of $A(0)$) are $3, 1 \pm i, 2 \pm 3i$. Using Matlab, numerical integration (ode15s) of (1.2) with $\nu = 1$ and $N = \text{diag}(1, 2, 3, 4, 5)$ yields:

$$A(5) = \begin{bmatrix} 1.00 & -1.00 & 0.00 & 0.00 & 0.00 \\ 1.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.99 & -2.99 & 0.00 \\ 0.00 & 0.00 & 2.99 & 1.99 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 3.00 \end{bmatrix},$$

and with $N = \text{diag}(1, 3, 5, 4, 2)$:

$$
A(5) = \begin{bmatrix}
1.00 & 0.00 & 0.00 & 0.00 & 1.00 \\
0.00 & 1.99 & 0.00 & 2.99 & 0.00 \\
0.00 & 0.00 & 3.00 & 0.00 & 0.00 \\
0.00 & -2.99 & 0.00 & 1.99 & 0.00 \\
-1.00 & 0.00 & 0.00 & 0.00 & 1.00
\end{bmatrix}.
$$

In both cases, one can observe the sorting property established in Theorem 7, and in both cases $\text{trace}(NA(5)) = 32$, which is the maximal value.

4.2. Polynomial Spectral Factorization. In the following polynomial spectral factorization problem, a sorting property of (1.2) is utilized which is not exactly captured in Theorem 7. However, since it may be of interest in certain applications, it is presented here and may be considered as a pointer for further investigations of (1.2). Polynomial spectral factorization has many applications in control and estimation theory (e.g. systems identification). The basic task is to factorize a polynomial with spectral constraints (separation of roots). In particular, consider an even monic polynomial with real coefficients and with no purely imaginary roots:

$$(4.2) \qquad p(x) = x^{2n} + a_{2n-2}x^{2n-2} + \ldots + a_2 x^2 + a_0.$$

Then the spectral factorization of p is a decomposition of the form

$$(4.3) \qquad p(x) = (-1)^n q(x) q(-x),$$

where the so-called spectral factor

$$(4.4) \qquad q(x) = x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0$$

has roots in the open left half plane only. In order to obtain q, one can utilize the dynamical system (1.2) in the following way. If

$$(4.5) \qquad N = \begin{bmatrix} c_1 I_n & 0 \\ 0 & c_2 I_n \end{bmatrix},$$

$c_1 < c_2$, is defined in this block diagonal form and if $A(0)$ is the companion matrix defined by (4.2), then

$$(4.6) \qquad A(\infty) = \begin{bmatrix} A_1(\infty) & 0 \\ 0 & A_2(\infty) \end{bmatrix}.$$

Due to the sorting property of (1.2), the trace of $A_1(\infty)$ is minimized. In particular, from (3.2) follows that

$$(4.7) \qquad A(\infty) = \begin{bmatrix} -P_1 & 0 \\ 0 & P_2 \end{bmatrix} + \begin{bmatrix} S_1 & S_2 \\ -S_2^T & S_3 \end{bmatrix}$$

with P_1, P_2 symmetric and S_1, S_3 skewsymmetric. Analogous arguments along the lines of the proof of Theorem 7 (Lemma 6) (diagonal matrices become now block-diagonal matrices), lead to the condition (instead of condition (3.29)) that trace $(Y_2^T Y_2 P_2 + Y_2 Y_2^T P_1) \geq 0$ in order that (4.7) is a stable equilibrium point (see condition (3.28) where $D = \operatorname{diag}(-P_1, P_2)$ and Y_2 is the (1,2)-block of Y). Since $A(\infty)$ is normal, the eigenvalues of $-P_1, P_2$ are the real parts of the eigenvalues of A_0. Assume now P_1 has a negative eigenvalue σ with the corresponding eigenvector u, then also P_2 has the negative eigenvalue σ with the corresponding eigenvector v, because the real parts are symmetric distributed w.r.t. the imaginary axis. Therefore, with $Y_2 = uv^T$, one would get $\operatorname{trace}(Y_2 P_2 Y_2^T + Y_2^T P_1 Y_2) < 0$. Hence, both P_1 and P_2 in (4.7) must be positive definite in order that (4.7) is a stable equilibrium point. Moreover, from (3.1) follows that $-P_1 S_2 - S_2 P_2 = 0$ or equivalently $(-P_2^{-1} \otimes P_1 - I)vec(S_2) = 0$ which implies that S_2 must zero, since $-P_2^{-1} \otimes P_1$ is negative definite. This means a stable equilibrium point $A(\infty)$ has indeed the form (4.6) and the trace of $A_1(\infty)$ is the sum of the n smallest real parts of the eigenvalues of $A(0)$. Thus, the spectrum of the normal matrices $A_1(\infty)$ and $A_2(\infty)$ lie in the open left half plane and in the open right half plane respectively. Hence, since $p(x) = \det(xI_{2n} - A(0)) = \det(xI_{2n} - A(\infty)) = \det(xI_n - A_1(\infty))\det(xI_n - A_2(\infty))$, one obtains $q(x) = \det(xI_n - A_1(\infty))$, $(-1)^n q(-x) = \det(xI_n - A_2(\infty))$. For example, for a polynomial $p(x) = x^6 - 2x^4 - 5x^2 - 6$ with

$$(4.8) \qquad A(0) = \begin{bmatrix} 0 & 2 & 0 & 5 & 0 & 6 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

one obtains

$$(4.9) \quad A_1(4) = \begin{bmatrix} -1.36 & 0.01 & -0.97 \\ 0.97 & -0.70 & -0.52 \\ -0.11 & 1.10 & -0.7 \end{bmatrix}, A_2(4) = \begin{bmatrix} 1.41 & -0.53 & 0.78 \\ 0.69 & 0.44 & -0.78 \\ 0.64 & 0.89 & 0.96 \end{bmatrix}$$

and thus the spectral factor is $q(x) = \det(xI_3 - A_1(4)) = x^3 + 2.82x^2 + 2.97x + 2.45$.

5. Summary. In this paper, dynamical systems of the form (1.2) has been analyzed and generalized. Firstly, a sorting property has been established analogous to the double-bracket equation. In particular, the sorting property is a consequence of the fact that the function $\operatorname{trace}(NA)$ is maximized by (1.2) under the constraint of A being normal. Secondly, a generalization of (1.2) in a Lie algebraic setting has been established. In particular, introducing a transposition operator and an Euclidian scalar product on a semisimple Lie algebra with a Cartan decomposition, (1.2)

has been generalized to a flow in a Lie algebra that converges to a certain canonical element. The use of a Cartan decomposition clearly reveals the underlying idea in (1.2), namely the orthogonality between the symmetric and skewsymmetric feedback components. Moreover, it allows to design other type of flows, which might be useful in various contexts.

There are several open questions left for future research. For example, due to the use of the Cartan decomposition, it is worthwhile to carry out a Lie group interpretation (symmetric spaces) of the results established here. Other open questions are the study of the convergence behavior of (1.2) if the equilibria are not isolated, the discretization of (1.2), a generalization to the infinite-dimensional case, the structure preserving properties of (1.2) for special eigenvalue problems, as well as further questions concerning properties of (1.2), e.g. the evolution of $G = G(t) \in SL_n(\mathbb{R})$ defined by $\dot{G} = UG$, $U = U(GA_0G^{-1}) \in \mathfrak{sl}_n(\mathbb{R})$ given by (2.3), $G(0) = I^2$. In particular, it is easy to see that $G(\infty)$ delivers the corresponding eigenvectors for A_0.

REFERENCES

[1] H. ABBASPOUR AND M. MOSKOWITZ. *Basic Lie Theory.* World Scientific Publishing, 2007.

[2] P.A. ABSIL, R. MAHONY, AND R. SEPULCHRE. *Optimization Algorithms on Matrix Manifolds.* Princeton University Press, 2007.

[3] M.A. BELABBAS. *Hamiltonian Systems for Computation.* PhD thesis, Harvard University, 2006.

[4] A.M. BLOCH. *Steepest descent, linear programming and Hamiltonian flows.* In: J.C. Lagarias and M.J. Todd, editors, Mathematical developments arising from linear programming, volume 114, pages 77–88. Contemporary Mathematics, A.M.S., 1990.

[5] A.M. BLOCH, R.W. BROCKETT, AND T.S. RATIU. *Completely integrable gradient flows.* Communications in Mathematical Physics, 147(1992), pp. 57–74.

[6] R.W. BROCKETT. *Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems.* In: Proc. of the 27th IEEE Conference on Decision and Control (CDC), Austin, USA, pages 799–803, 1988.

[7] R.W. BROCKETT. *Least squares matching problems.* Linear Algebra and its Applications, 124(1989), pp. 761–777.

[8] R.W. BROCKETT. *Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems.* Linear Algebra and its Applications, 146(1991), pp. 76–91.

[9] R.W. BROCKETT. *Differential geometry and the design of gradient algorithms.* Proceedings of Symposia in Pure Mathematics, 54(1993), pp. 69–92.

[10] M.T. CHU AND K.R. DRIESSEL. *The projected gradient method for least squares matrix approximations with spectral constraints.* SIAM Journal of Numerical Analysis, 27(990), pp. 1050–1060.

---

[2] Notice that the relationship between (2.1) with $A(0) = A_0$ and $\dot{G} = U(GA_0G^{-1})G$ with $G(0) = I$ is given by $A(t) = G(t)A_0G(t)^{-1}$, see e.g. [25, 28, 32].

[11] M.T. Chu and G.H. Golub. *Inverse Eigenvalue Problems: Theory, Algorithms, and Applications*. Oxford Science Publications, 2004.

[12] P. Deift, T. Nanda, and C. Tomei. *Differential equations for the symmetric eigenvalue problem*. SIAM Journal of Numerical Analysis, 20(1983), pp. 1–22.

[13] C. Ebenbauer. *A dynamical system that computes eigenvalues and diagonalizes matrices with a real spectrum*. In: Proc. of the 46th IEEE Conference on Decision and Control (CDC), New Orleans, USA, pages 1704–1709, 2007.

[14] P. Fernández and R.V. Solé. *The role of computation in complex regulatory networks*. In: E. V. Koonin, Y.I. Wolf, and G.P. Karev, editors, Power Laws, Scale-Free Networks and Genome Biology, pages 206–225. Molecular Biology Intelligence Unit, Springer, 2007.

[15] W. Hahn. *Stability of Motion*. Springer, 1967.

[16] G. Hardy, J.E. Littlewood, and G. Polya. *Inequalities*. Cambridge University Press, 2nd edition, 1952.

[17] U. Helmke and J.B. Moore. *Optimization and Dynamical Systems*. Springer, 2nd edition, 1996.

[18] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985 (Reprint 1999).

[19] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[20] A. Iggidr, B. Kalitine, and R. Outbib. *Semidefinite Lyapunov functions stability and stabilization*. Mathematics of Control, Signals, and Systems (MCSS), 9(1996), pp. 95–106.

[21] N. Jacobson. *Lie algebras*. Dover Publications, 1979.

[22] H.K. Khalil. *Nonlinear Systems*. Prentice Hall, 3rd edition, 2002.

[23] A.W. Knapp. *Lie Groups Beyond an Introduction*. Birkhäuser, 2nd edition, 2005.

[24] J.P. LaSalle. *The Stability of Dynamical Systems*. SIAM Press, 1976.

[25] P.D. Lax. *Linear Algebra and its Applications*. Wiley, 2nd edition, 2007.

[26] W. Maass, P. Joshi, and E. Sontag. *Computational aspects of feedback in neuronal circuits*. Computational Biology, 3:e165(2007), pp. 1–20.

[27] B.J. MacLennan. *A review of analog computation*, 2008. Encyclopedia of Complexity and System Science.

[28] J. Moser and E.J. Zehnder. *Notes on dynamical systems*. American Mathematical Society, 2005.

[29] T. Nanda. *Differential equations and the QR flow*. SIAM J. on Num. Analysis, 22(1985), pp. 310–321.

[30] H.T. Siegelmann. *Neuronal Networks and Analog Computation. Beyond the Turing Limit*. Birkhäuser, 1999.

[31] J. von Neumann. *Some matrix-inequalities and metrization of matrix spaces*. Tomsk Univ. Rev., 1(1937), pp. 286–300.

[32] D.S. Watkins and L. Elsner. *Self-similar flows*. Linear Algebra and its Applications, 100(1988), pp. 213–242.