Construction of a contraction metric by meshless collocation

Article (Accepted Version)

Giesl, Peter and Wendland, Holger (2019) Construction of a contraction metric by meshless collocation. Discrete and Continuous Dynamical Systems - Series B, 24 (8). pp. 3843-3863. ISSN 1531-3492

Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

doi:10.3934/xx.xx.xx.xx

pp. **X–XX**

CONSTRUCTION OF A CONTRACTION METRIC BY MESHLESS COLLOCATION

Peter Giesl*

Department of Mathematics
University of Sussex
Falmer, BN1 9QH, UK

Holger Wendland

Department of Mathematics
University of Bayreuth
95440 Bayreuth, Germany

(Communicated by the associate editor name)

Abstract. A contraction metric for an autonomous ordinary differential equation is a Riemannian metric such that the distance between adjacent solutions contracts over time. A contraction metric can be used to determine the basin of attraction of an equilibrium and it is robust to small perturbations of the system, including those varying the position of the equilibrium.

The contraction metric is described by a matrix-valued function $M(x)$ such that $M(x)$ is positive definite and $F(M)(x)$ is negative definite, where F denotes a certain first-order differential operator. In this paper, we show existence, uniqueness and continuous dependence on the right-hand side of the matrix-valued partial differential equation $F(M)(x) = -C(x)$. We then use a construction method based on meshless collocation, developed in the companion paper [12], to approximate the solution of the matrix-valued PDE. In this paper, we justify error estimates showing that the approximate solution itself is a contraction metric. The method is applied to several examples.

1. **Introduction.** Many important problems from applications require the analysis of autonomous ordinary differential equations (ODE). The long-term behaviour of solutions can often be determined by attractors and their basins of attraction.

In this paper, we are interested in the existence, uniqueness and exponential stability of an equilibrium, as well as the determination of its basin of attraction. To avoid the direct numerical approximation of many solutions which is costly and requires estimates to be exact, other methods have been developed. If the equilibrium is known, then Lyapunov functions [21] are one way of analysing the stability of the equilibrium as well as its basin of attraction. There is a vast literature on Lyapunov functions, see, e.g., [29, 14] as well as the review [16] for converse theorems proving the existence of Lyapunov functions and the review [11] on computational methods. Other methods include density functions [25], which show that almost all solutions are attracted by the origin, and Koopman operators [22], which are linear operators

2010 *Mathematics Subject Classification.* Primary: 37B25, 65N35; Secondary: 37M99, 65N15.
Key words and phrases. Contraction metric, basin of attraction, matrix-valued partial differential equation, meshless collocation.

* Corresponding author.

(observables) on a function space, such that their spectral properties give insight into the stability properties of the dynamical system.

A different way of studying stability and the basin of attraction, which does not require any knowledge about the equilibrium and which is also robust with respect to perturbations of the ODE, uses contraction metrics. Contraction analysis can be used to study the distance between trajectories, without reference to an attractor, establishing (exponential) attraction of adjacent trajectories, see [17, 15] and also [11, Section 2.10]; it can be generalised to the study of a Finsler-Lyapunov function [5].

If the distance between the trajectory through x and any adjacent trajectory is contracted over time, then solutions converge to an equilibrium. If the attractor is, e.g., a periodic orbit, then the distance to adjacent trajectories in the tangential direction to the trajectories cannot contract.

Only few converse theorems for contraction metrics have been obtained, establishing the existence of a contraction metric, see [8] for some references. A constructive converse theorem, providing algorithms for the explicit construction of a contraction metric, is given in [2] for the global stability of an equilibrium in polynomial systems, using Linear Matrix Inequalities (LMI) and sums of squares (SOS). In contrast to our method which requires a compact set, [2] shows global stability. On the other hand, our method can deal with general smooth right-hand sides, while [2] assumes the right-hand side to be polynomial. Moreover, since the SOS condition is not equivalent to the positivity of a matrix, but just a sufficient condition, one cannot prove a converse theorem, while we will establish such a result in this paper. An algorithm to construct a continuous piecewise affine (CPA) contraction metric for periodic orbits in time-periodic systems using semi-definite optimization has been proposed in [9].

In this paper, we will introduce a method to construct a contraction metric for an equilibrium using meshless collocation. The (Riemannian) contraction metric will be expressed by a matrix-valued function $M\colon \mathbb{R}^n \to \mathbb{S}^{n\times n}$, where $\mathbb{S}^{n\times n}$ denotes the symmetric $n\times n$ matrices with real entries, such that $M(x)$ is positive definite for every x. It defines a (point-dependent) scalar product on $\mathbb{R}^n$ by $\langle v,w\rangle_{M(x)} = v^\top M(x)w$. For M to be a contraction metric, we require the distance between adjacent solutions to decrease with respect to such a contraction metric. This can be expressed by the negative definiteness of $F(M)(x) = Df(x)^\top M(x) + M(x)Df(x) + M'(x)$, where $(M'(x))_{ij} = \nabla M_{ij}(x)f(x)$ denotes the orbital derivative.

To construct a contraction metric we approximate the matrix-valued solution $M(x)$ of the PDE

$$F(M)(x) = Df(x)^\top M(x) + M(x)Df(x) + M'(x) = -C,$$

where $C \in \mathbb{S}^{n\times n}$ is a given positive definite matrix. After fixing a finite number of collocation points, the approximation $S(x)$ is obtained as the optimal recovery of M based on the information in the collocation points in a reproducing kernel Hilbert space (RKHS), more precisely a Sobolev space. In practice, the approximation $S(x)$ is obtained by solving a system of linear equations.

In the accompanying paper [12], we derive the numerical framework for approximating solutions to general matrix-valued PDEs, including error estimates. We also consider the specific type of differential operator F for contraction metrics, described above. In this paper, we will derive explicit formulas for our specific case.

The outline of this paper is as follows: in Section 2 we will prove existence, uniqueness and continuous dependence on the right-hand side of the equation

$$F(M)(x) = M(x)Df(x) + Df(x)^\top M(x) + M'(x) \quad = \quad -C(x).$$

In Section 3 we will discuss the method to construct a contraction metric by solving a matrix-valued PDE using meshless collocation by adapting the general method from [12] to our case. We will show error estimates and establish that the approximation itself is a contraction metric if the collocation points are sufficiently dense. Finally, we apply the method to several examples in Section 4, including a perturbed van der Pol system with reversed time and a three-dimensional example, before we conclude in Section 5.

2. Contraction metric.

We consider the autonomous ODE

$$\dot{x} \quad = \quad f(x) \tag{1}$$

where $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$; further assumptions on the smoothness of f will be made later. The solution $x(t)$ with initial condition $x(0) = \xi$ is denoted by $x(t) =: S_t\xi$ and is assumed to exist for all $t \geq 0$.

We are interested in the existence, uniqueness and exponential stability of an equilibrium, as well as the determination of its basin of attraction. An equilibrium is a point $x_0 \in \mathbb{R}^n$ such that $f(x_0) = 0$. The basin of attraction of an asymptotically stable equilibrium is defined by $A(x_0) = \{x \in \mathbb{R}^n \mid \lim_{t\to\infty} S_t x = x_0\}$.

Existence and uniqueness of an exponentially stable equilibrium, as well as information about its basin of attraction can be obtained from a Riemannian contraction metric. There are many results in this direction, dating back to the middle of the 20th century [19, 17, 18, 20]. For example, there are results available on the rate of attraction, generalisations to manifolds or relaxations of the assumptions on, e.g. the smoothness of M. Here, we cite a theorem from [8].

Theorem 2.1. *Let $\emptyset \neq G \subset \mathbb{R}^n$ be a compact, connected and positively invariant set and M be a Riemannian contraction metric in G, i.e.*

- *$M \in C^1(G, \mathbb{S}^{n\times n})$, such that $M(x)$ is positive definite for all $x \in G$.*
- *$F(M)(x) := Df(x)^\top M(x) + M(x)Df(x) + M'(x)$ is negative definite for all $x \in G$, where $(M'(x))_{ij} = \nabla M_{ij}(x)f(x)$ denotes the orbital derivative.*

Then there exists one and only one equilibrium x_0 in G; x_0 is exponentially stable and $G \subset A(x_0)$.

To apply the theorem, we need to constructively find such a contraction metric as well as a positively invariant set. A positively invariant set can, e.g., be found through a level set of a Lyapunov-like function, on which the function is strictly decreasing along orbits as discussed at the end of Section 5. In this paper, we focus on the problem of construction a contraction metric.

In [8], the existence of a solution of the linear first-order PDE

$$F(M)(x) := M(x)Df(x) + Df(x)^\top M(x) + M'(x) \quad = \quad -C \tag{2}$$

for all $x \in A(x_0)$ was established, where $C \in \mathbb{S}^{n\times n}$ is a given positive definite matrix. On the other hand, in [12] a method for solving such linear matrix-valued PDEs was introduced. This method uses the optimal recovery S of M in a RKHS and provides an error estimate on $\|F(M) - F(S)\|$ depending on the fill distance of the collocation points.

In this section we will establish uniqueness of solutions of (2). Moreover, we will prove existence and uniqueness of the slightly more general equation

$$F(M)(x) = M(x)Df(x) + Df(x)^\top M(x) + M'(x) \quad = \quad -C(x).$$

This will enable us to prove an estimate of the form

$$\|M_1(x) - M_2(x)\|_{L_\infty(K;\mathbb{S}^{n\times n})} \le c\|F(M_1) - F(M_2)\|_{L_\infty(K;\mathbb{S}^{n\times n})},$$

where K is a compact, positively invariant subset of $A(x_0)$. In particular, we can conclude in Section 3 that the meshless collocation method will construct a contraction metric if the collocation points are sufficiently dense.

2.1. **Existence.** The following existence theorem is a generalisation of [8, Theorem 4.4], where $C(x)$ is the constant matrix C. The proof follows the same strategy with some modifications. Note that, in addition to the existence of a solution of (3), we show that $M(x)$ is symmetric and positive definite, and has a particular integral form, which will be used in Theorem 2.4.

Theorem 2.2. *Consider the dynamical system given by $\dot{x} = f(x)$, $f \in C^s(\mathbb{R}^n, \mathbb{R}^n)$, $s \ge 2$, and assume that x_0 is an exponentially stable equilibrium with basin of attraction $A(x_0)$. Let $C \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$, such that $C(x)$ is a positive definite matrix for all $x \in A(x_0)$.*

Then there exists a function $M \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$, such that $M(x)$ is symmetric and positive definite for all $x \in A(x_0)$, and

$$Df(x)^\top M(x) + M(x)Df(x) + M'(x) = -C(x) \text{ for all } x \in A(x_0). \tag{3}$$

M is of the form $M(x) = \int_0^\infty \phi(\tau, 0; x)^\top C(S_\tau x)\phi(\tau, 0; x)\, d\tau$, where ϕ is the principal fundamental matrix solution of $\dot{y} = Df(S_t x)y$ at initial time 0.

Proof. For the proof we choose the vector norm $\|x\|^2 = \|x\|_2^2 = \sum_{i=1}^n x_i^2$ and the induced matrix norm, defined by $\|A\| = \|A\|_2 = \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}$. Note that the matrix norm is sub-multiplicative and satisfies $\|A\| = \|A^\top\|$.

Step 1: Definition of M

We consider the linear, non-autonomous ODE

$$\dot{y} = Df(S_t x)y.$$

As $Df(S_t x)$ is defined and continuous for all $x \in A(x_0)$ and $t \ge 0$, the principal fundamental matrix solution of the initial value problem with initial time t_0 exists and we denote it by

$$\phi(t, t_0; x).$$

Note that for fixed x there exists a $\theta_0 > 0$ such that $S_t x$, $Df(S_t x)$ and thus also $\phi(t, t_0; x)$ are defined for all $t, t_0 \ge -\theta_0$, and $\phi(t, t_0; x)$ is C^{s-1} with respect to x, t and t_0.

By the Chapman-Kolmogorov identities, cf. e.g. [4], p. 151, we have

$$\frac{d}{dt}\phi(t, t_0; x) \quad = \quad Df(S_t x)\phi(t, t_0; x), \tag{4}$$

$$\frac{d}{dt_0}\phi(t, t_0; x) \quad = \quad -\phi(t, t_0; x)Df(S_{t_0} x), \tag{5}$$

$$\phi(t_0, t_0; x) \quad = \quad I, \tag{6}$$

$$\phi(t, 0; S_\theta x) \quad = \quad \phi(t + \theta, \theta; x). \tag{7}$$

for all $t, t + \theta \ge -\theta_0$.

We define the function

$$M(x) = \int_0^\infty \phi(\tau,0;x)^\top C(S_\tau x)\phi(\tau,0;x)\,d\tau. \tag{8}$$

We will show that $M(x)$ is well defined for all $x \in A(x_0)$, symmetric and positive definite and also satisfies the equation (3) in Step 2 and that it is C^{s-1} in Step 3.

Step 2: Matrix equation

Define

$$g_T(\theta,x) = \int_\theta^{T+\theta} \phi(\tau,\theta;x)^\top C(S_\tau x)\phi(\tau,\theta;x)\,d\tau. \tag{9}$$

We have for all $\theta \geq -\theta_0$ by a change of variables and (7)

$$g_T(\theta,x) = \int_0^T \phi(\tau+\theta,\theta;x)^\top C(S_{\tau+\theta}x)\phi(\tau+\theta,\theta;x)\,d\tau \tag{10}$$

$$= \int_0^T \phi(\tau,0;S_\theta x)^\top C(S_{\tau+\theta}x)\phi(\tau,0;S_\theta x)\,d\tau. \tag{11}$$

We will show that $g_T(\theta,x)$ converges pointwise and $\frac{d}{d\theta}g_T(\theta,x)$ converges uniformly for $|\theta| \leq \theta_0$ as $T \to \infty$ so that $\frac{d}{d\theta}\lim_{T\to\infty} g_T(\theta,x) = \lim_{T\to\infty}\frac{d}{d\theta}g_T(\theta,x)$.

By the exponential stability of x_0 there is a positively invariant, compact neighbourhood U of x_0 and $K,\mu > 0$ such that $\|S_t x - x_0\| \leq Ke^{-\mu t}$ holds for all $x \in U$ and $t \geq 0$.

Fix $x \in A(x_0)$. Since Df is locally Lipschitz-continuous at x_0, there is $d > 0$ such that

$$\|Df(S_{t+\theta}x) - Df(x_0)\| \leq de^{-\mu t} \tag{12}$$

for all $t \geq 0$ and all $|\theta| \leq \theta_0$. To show (12), we use the Lipschitz-continuity and the exponential decay for all $t \geq T_0$, where T_0 is so large that $S_{t-\theta_0}x \in U$ for all $t \geq T_0$. Then we choose the constant d so large that (12) also holds for all $t \in [0,T]$ and $|\theta| \leq \theta_0$.

[8, Lemma A.2], applied to $A(t) = Df(S_t S_\theta x)$ and $A = Df(x_0)$, gives

$$\|\phi(t,0;S_\theta x)\| \leq ce^{-\rho t} \tag{13}$$

for all $|\theta| \leq \theta_0$ and $t \geq 0$.

Recall that $S_{t-\theta}x \in U$ for all $t \geq T_0$. Hence, $\bigcup_{t=0}^\infty \{S_{t-\theta_0}x\} \subset \bigcup_{t=0}^{T_0}\{S_{t-\theta_0}x\} \cup U$, which shows that the set $S = \overline{\bigcup_{t=-\theta_0}^\infty \{S_t x\}}$ is compact and, since $C(\cdot)$ is continuous, there exists $C^* > 0$ such that

$$\|C(S_t x)\| \leq C^* \tag{14}$$

for all $t \geq -\theta_0$.

Thus we have, with a new constant c, by (13) and (14)

$$\|\phi(\tau,0;S_\theta x)^\top C(S_{\tau+\theta}x)\phi(\tau,0;S_\theta x)\| \leq ce^{-2\rho\tau} \tag{15}$$

for all $|\theta| \leq \theta_0$ and all $\tau \geq 0$. The right-hand side is integrable over $\tau \in [0,\infty)$. Hence, by Lebesgue's dominated convergence theorem, the function $g_T(\theta,x)$, see (11), converges pointwise for $T \to \infty$. This shows that $M(x)$ is well definied, and it is clear from (8) that $M(x)$ is symmetric and positive definite.

Also, using (9), (6) and (5), we have, similar to (9) to (11)

$$\begin{aligned}\frac{d}{d\theta}g_T(\theta,x) &= \phi(T+\theta,\theta;x)^\top C(S_{T+\theta}x)\phi(T+\theta,\theta;x)-C(S_\theta x)\\ &\quad -Df(S_\theta x)^\top\int_\theta^{T+\theta}\phi(\tau,\theta;x)^\top C(S_\tau x)\phi(\tau,\theta;x)\,d\tau\\ &\quad -\int_\theta^{T+\theta}\phi(\tau,\theta;x)^\top C(S_\tau x)\phi(\tau,\theta;x)\,d\tau\ Df(S_\theta x)\\ &= \phi(T+\theta,\theta;x)^\top C(S_{T+\theta}x)\phi(T+\theta,\theta;x)-C(S_\theta x)\\ &\quad -Df(S_\theta x)^\top\int_0^{T}\phi(\tau+\theta,\theta;x)^\top C(S_{\tau+\theta} x)\phi(\tau+\theta,\theta;x)\,d\tau\\ &\quad -\int_0^{T}\phi(\tau+\theta,\theta;x)^\top C(S_{\tau+\theta} x)\phi(\tau+\theta,\theta;x)\,d\tau\ Df(S_\theta x)\\ &= \phi(T,0;S_\theta x)^\top C(S_{T+\theta}x)\phi(T,0;S_\theta x)-C(S_\theta x)\\ &\quad -Df(S_\theta x)^\top\int_0^{T}\phi(\tau,0;S_\theta x)^\top C(S_{\tau+\theta} x)\phi(\tau,0;S_\theta x)\,d\tau\\ &\quad -\int_0^{T}\phi(\tau,0;S_\theta x)^\top C(S_{\tau+\theta} x)\phi(\tau,0;S_\theta x)\,d\tau\ Df(S_\theta x)\end{aligned}$$

by (7). By (15), the right-hand side converges uniformly for $|\theta|\le\theta_0$ as $T\to\infty$. Hence, for $|\theta|\le\theta_0$ we can exchange limit and derivative, obtaining

$$\begin{aligned}&\frac{d}{d\theta}\lim_{T\to\infty}g_T(\theta,x)\\ &= \lim_{T\to\infty}\frac{d}{d\theta}g_T(\theta,x)\\ &= -C(S_\theta x)-Df(S_\theta x)^\top\int_0^\infty\phi(\tau,0;S_\theta x)^\top C(S_{\tau+\theta}x)\phi(\tau,0;S_\theta x)\,d\tau\\ &\quad -\int_0^\infty\phi(\tau,0;S_\theta x)^\top C(S_{\tau+\theta}x)\phi(\tau,0;S_\theta x)\,d\tau\ Df(S_\theta x).\end{aligned}\tag{16}$$

Altogether, we thus have

$$\begin{aligned}M'(x) &= \frac{d}{d\theta}M(S_\theta x)\Big|_{\theta=0}\\ &= \frac{d}{d\theta}\lim_{T\to\infty}\left[\int_0^T\phi(\tau,0;S_\theta x)^\top C(S_\tau S_\theta x)\phi(\tau,0;S_\theta x)\,d\tau\right]\Bigg|_{\theta=0}\\ &= \frac{d}{d\theta}\lim_{T\to\infty}\left[\int_0^T\phi(\tau+\theta,\theta;x)^\top C(S_{\tau+\theta}x)\phi(\tau+\theta,\theta;x)\,d\tau\right]\Bigg|_{\theta=0} \quad\text{by (7)}\\ &= \frac{d}{d\theta}\lim_{T\to\infty}g_T(\theta,x)\Big|_{\theta=0} \quad\text{by (10)}\\ &= -C(x)-Df(x)^\top M(x)-M(x)Df(x)\end{aligned}$$

by (16) and (8). This shows the matrix equation (3).

Step 3: Smoothness of M

Let $-\nu < 0$ be the largest real part of all eigenvalues of $Df(x_0)$ and let $\epsilon = \frac{\nu}{2}$. By Step 1 of [8, Theorem 4.1] there is an invertible matrix $T \in \mathbb{R}^{n\times n}$ such that

$$\max_{\|w\|=1} w^\top T^{-1} Df(x_0) T w \;\le\; -\nu + \frac{\epsilon}{2}. \tag{17}$$

Since $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ we can choose a positively invariant, compact neighborhood U of x_0 so small that

$$\|T^{-1}[Df(x) - Df(x_0)]T\| \;\le\; \frac{\epsilon}{2} \tag{18}$$

holds for all $x \in U$.

Let $\nu' = \min\left(\frac{\nu}{4}, \rho\right)$, where ρ was defined in (13). The proof of

$$\|T^{-1}\partial_x^\alpha \phi(t,0;x)\| \;\le\; C_\alpha e^{-\nu' t} \tag{19}$$

for all $\alpha \in \mathbb{N}_0^n$ with $|\alpha| := \sum_{i=1}^n |\alpha_i| \le s-1$, $x \in U$ and $t \ge 0$ is as in the proof of [8, Theorem 4.4].

Next, we show that $\int_0^T \partial_x^\alpha(\phi(\tau,0;x)^\top C(S_\tau x)\phi(\tau,0;x))\,d\tau$ converges uniformly with respect to x as $T \to \infty$ for $1 \le |\alpha| \le s-1$.

First, note that there is a constant c such that

$$\|\partial_x^\alpha C(S_\tau x)\| \;\le\; c \tag{20}$$

for all $0 \le |\alpha| \le s-1$ and all $x \in U$, where U is the previously defined compact and positively invariant neighborhood of x_0. This follows from the chain rule, since C is C^{s-1}, $S_\tau x$ is C^s with respect to x, and U is compact and positively invariant.

Then we can proceed as in the proof of [8, Theorem 4.4] to show that

$$\int_0^T \|\partial_x^\alpha(\phi(\tau,0;x)^\top C(S_\tau x)\phi(\tau,0;x))\|\,d\tau \;\le\; \int_0^T \tilde{c}e^{-2\nu'\tau}\,d\tau$$

for all $x \in U$, $T \ge 0$ and $|\alpha| \le s-1$. For $x \in A(\Omega)$ we can choose a bounded, open neighborhood of x with $\overline{O} \subset A(\Omega)$. Then there is a $T_0 \ge 0$ such that $S_{T_0+t}\overline{O} \subset U$ for all $t \ge 0$. Hence, $\int_0^T \partial_x^\alpha(\phi(\tau,0;x)^\top C(S_\tau x)\phi(\tau,0;x))\,d\tau$ converges uniformly in x as $T \to \infty$. This proves that $M \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$. □

2.2. **Uniqueness and continuous dependence.** In the next theorem we will prove the uniqueness of solutions of (21). Note that, as usual, a solution is a function $M \in C(A(x_0), \mathbb{R}^{n\times n})$, such that the orbital derivative $M'(x)$ exists for all $x \in A(x_0)$ and (21) holds for all $x \in A(x_0)$.

Note that (21) is considered without an initial condition, it will turn out that this is not needed; for details see the proof of the following theorem.

Theorem 2.3. *Let $f \in C^s(\mathbb{R}^n, \mathbb{R}^n)$, $s \ge 2$. Let x_0 be an exponentially stable equilibrium of $\dot{x} = f(x)$ with basin of attraction $A(x_0)$. Let $C \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$ such that $C(x)$ is a positive definite matrix for all $x \in A(x_0)$.*

Then the matrix equation

$$Df(x)^\top M(x) + M(x)Df(x) + M'(x) \;=\; -C(x) \text{ for all } x \in A(x_0) \tag{21}$$

has a unique solution.

In particular, $M \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$, $M(x)$ is positive definite for all $x \in A(x_0)$ and M is of the form $M(x) = \int_0^\infty \phi(\tau,0;x)^\top C(S_\tau x)\phi(\tau,0;x)\,d\tau$, where $\phi(\tau,0;x)$ is the principal fundamental matrix solution of $\dot{y} = Df(S_t x)y$ at initial time 0.

Proof. In Theorem 2.2 it was shown that there exists a solution to the matrix equation with the properties described in the lemma. Hence, it is enough to show that the matrix equation has at most one solution.

Assume that $M_1, M_2 \colon A(x_0) \to \mathbb{R}^{n\times n}$ are two solutions of the matrix differential equation (21). Fix $x \in A(x_0)$. Along the solution $S_t x$, (21) becomes

$$Df(S_tx)^\top M_i(S_tx) + M_i(S_tx)Df(S_tx) + \frac{d}{dt}M_i(S_tx) = -C(S_tx). \tag{22}$$

Regarding $M_i(S_tx)$, $i = 1, 2$ as a vector-valued function in $\mathbb{R}^{(n^2)}$, the equation (22) is a linear ODE with respect to t for M_i with continuous coefficients, and thus the corresponding initial value problem at time $t = 0$ with $M_i(x) = M_i^0$ has a unique solution, which exists for all $t \geq 0$, since $Df(S_tx)$ is defined and continuous for all $t \geq 0$; note that $x \in A(x_0)$. However, we are not given an initial value for M_i.

Denote by $\phi(t, t_0; x)$ the principal fundamental matrix solution of $\dot y = Df(S_tx)y$ with initial time t_0, see (4) to (7). We have for $i = 1, 2$ with (22) and (4)

$$\begin{aligned}
&\frac{d}{dt}\left(\phi(t,0;x)^\top M_i(S_tx)\phi(t,0;x)\right)\\
&= \phi(t,0;x)^\top \left[Df(S_tx)^\top M_i(S_tx) + \frac{d}{dt}M_i(S_tx) + M_i(S_tx)Df(S_tx)\right]\phi(t,0;x)\\
&= -\phi(t,0;x)^\top C(S_tx)\phi(t,0;x)
\end{aligned}$$

and hence

$$\frac{d}{dt}\left(\phi(t,0;x)^\top [M_1(S_tx) - M_2(S_tx)]\phi(t,0;x)\right) = 0.$$

This implies

$$\begin{aligned}
0 &= \int_0^t \frac{d}{d\tau}\left[\phi(\tau,0;x)^\top [M_1(S_\tau x) - M_2(S_\tau x)]\phi(\tau,0;x)\right] d\tau\\
&= \phi(t,0;x)^\top [M_1(S_tx) - M_2(S_tx)]\phi(t,0;x) - [M_1(x) - M_2(x)].
\end{aligned}$$

Note that all terms depend continuously on t, since $M_i(\cdot)$ and the solution S_tx of the ODE do. Letting $t \to \infty$ we have with $S_tx \to x_0$, since $x \in A(x_0)$, the continuity of M_i and (13)

$$[M_1(x) - M_2(x)] = 0.$$

This shows $M_1(x) = M_2(x)$ for an arbitrary $x \in A(x_0)$ and thus the uniqueness. □

In the next theorem we will prove the continuous dependence of solutions on the right-hand side. The proof uses the uniqueness in Theorem 2.3, in particular the expression $M(x) = \int_0^\infty \phi(\tau,0;x)^\top C(S_\tau x)\phi(\tau,0;x)\, d\tau$ for M.

This will enable us in Section 3 to extend the existing error estimates on $\|F(M) - F(S)\|$, where S is the approximation of M by meshless collocation, to an error estimate on $\|M - S\|$.

Theorem 2.4. *Let $f \in C^s(\mathbb{R}^n, \mathbb{R}^n)$, $s \geq 2$. Let x_0 be an exponentially stable equilibrium of $\dot x = f(x)$ with basin of attraction $A(x_0)$. Let $C_i \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$, $i = 1, 2$, such that $C_i(x)$ is a positive definite matrix for all $x \in A(x_0)$.*

Let $M_i \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$ be the unique solution (see Theorem 2.3) of the matrix equation

$$Df(x)^\top M_i(x) + M_i(x)Df(x) + M_i'(x) = -C_i(x) \text{ for all } x \in A(x_0) \tag{23}$$

where $i = 1, 2$.

Let $K \subset A(x_0)$ *be a compact set.*

Then there is a constant c*, independent of* M_i *and* C_i *such that*

$$\|M_1 - M_2\|_{L_\infty(K;\mathbb{S}^{n\times n})} \quad \leq \quad c\|C_1 - C_2\|_{L_\infty(\overline{\gamma^+(K)};\mathbb{S}^{n\times n})} \tag{24}$$

where $\gamma^+(K) = \bigcup_{t\geq 0} S_t K$.

Remark 1. Note that for a positively invariant and compact set K we have $\overline{\gamma^+(K)} = K$.

Proof. For the proof we choose again the induced matrix norm $\|\cdot\| = \|\cdot\|_2$. Note that the matrix norm is sub-multiplicative and satisfies $\|A\| = \|A^\top\|$. The result follows for a general matrix norm with a different constant.

By Theorems 2.2 and 2.3, the unique solutions satisfy

$$M_i(x) = \int_0^\infty \phi(\tau, 0; x)^\top C_i(S_\tau x)\phi(\tau, 0; x)\, d\tau.$$

Hence, for fixed $x \in K$ we have

$$\begin{aligned}
\|M_1(x) - M_2(x)\| &= \left\| \int_0^\infty \phi(\tau, 0; x)^\top (C_1(S_\tau x) - C_2(S_\tau x))\phi(\tau, 0; x)\, d\tau \right\| \\
&\leq \int_0^\infty \|\phi(\tau, 0; x)\|^2 \|C_1(S_\tau x) - C_2(S_\tau x)\|\, d\tau \\
&\leq \|C_1 - C_2\|_{L_\infty(\overline{\gamma^+(K)};\mathbb{S}^{n\times n})} \int_0^\infty \|\phi(\tau, 0; x)\|^2\, d\tau.
\end{aligned}$$

We will now show that there are uniform constants ρ and c_1 such that

$$\|\phi(t, 0; x)\| \quad \leq \quad c_1 e^{-\rho t} \tag{25}$$

for all $x \in K$ and all $t \geq 0$.

By the exponential stability of x_0 there is a positively invariant, compact neighbourhood U of x_0 such that $\|S_t x - x_0\| \leq c_0 e^{-\mu t}$ holds for all $x \in U$ and $t \geq 0$; we can choose U so small that Df satisfies a Lipschitz condition on U.

Since x_0 attracts the compact set K uniformly there is a time $T^* > 0$ such that $S_t K \subset U$ for all $t \geq T^*$. Since Df is Lipschitz-continuous in U we can conclude

$$\|Df(S_t x) - Df(x_0)\| \leq L\|S_t x - x_0\| \leq d e^{-\mu t}$$

for all $t \geq 0$ and $x \in K$; this follows directly for all $t \geq T^*$, but holds also for all $t \geq 0$ when choosing d appropriately as $\bigcup_{\tau=0}^{T^*} S_\tau K$ is a compact set.

Now, for any $x \in K$, we can apply [8, Lemma A.2] to $A(t) = Df(S_t x)$ and $A = Df(x_0)$ to obtain (25).

Thus we have for all $x \in K$, using (25)

$$\begin{aligned}
\|M_1(x) - M_2(x)\| &\leq \|C_1 - C_2\|_{L_\infty(\overline{\gamma^+(K)};\mathbb{S}^{n\times n})} \int_0^\infty c_1^2 e^{-2\rho t}\, d\tau \\
&\leq \frac{c_1^2}{2\rho} \|C_1 - C_2\|_{L_\infty(\overline{\gamma^+(K)};\mathbb{S}^{n\times n})}.
\end{aligned}$$

□

3. **Meshless collocation.** In this section we want to solve the matrix-valued equation

$$F(M)(x) := Df(x)^\top M(x) + M(x)Df(x) + M'(x) = -C. \tag{26}$$

We use meshless collocation to approximate the solution of the above PDE by a matrix-valued function S, see [12]. The error estimates will establish that the approximation S itself is a contraction metric, as $S(x)$ will be positive definite and $F(S)(x)$ negative definite if the collocation points are sufficiently dense. We continue to assume that x_0 is an exponentially stable equilibrium of (1) with basin of attraction $A(x_0)$.

We summarise the methodology and establish error estimates in Section 3.1 before we compute the specific formulas for our case in Section 3.2.

3.1. **Matrix-valued collocation and error estimates.** Meshless collocation, in particular by Radial Basis Functions, is used to approximate multivariate functions and approximately solve Partial Differential Equations [24, 3, 26]. For a general introduction to meshless collocation and reproducing kernel Hilbert spaces (RKHS) see [28] for real-valued functions. We will follow [12] here to introduce the method for matrix-valued functions.

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain with Lipschitz boundary and denote by $H^\sigma(\Omega)$ the Sobolev space of order $\sigma > n/2$, where the weak derivatives are measured in the $L_2(\Omega)$-norm. The assumption $\sigma > n/2$ ensures $H^\sigma(\Omega) \subseteq C(\Omega)$ by the Sobolev embedding theorem.

We are interested in RKHS of matrix-valued functions. Let us first introduce RKHS of functions with values in a general Hilbert space W.

Definition 3.1. Let W be a real Hilbert space and denote the linear space of all linear and bounded operators $L\colon W \to W$ by $\mathcal{L}(W)$.

The Hilbert space $\mathcal{H}(\Omega; W)$ of functions $g\colon \Omega \to W$ is called a *reproducing kernel Hilbert space (RKHS)* if there is a function $\Phi : \Omega \times \Omega \to \mathcal{L}(W)$ with

1. $\Phi(\cdot, x)\alpha \in \mathcal{H}(\Omega; W)$ for all $x \in \Omega$ and all $\alpha \in W$.
2. $\langle g(x), \alpha\rangle_W = \langle g, \Phi(\cdot, x)\alpha\rangle_{\mathcal{H}(\Omega;W)}$ for all $g \in \mathcal{H}(\Omega; W)$, all $x \in \Omega$ and all $\alpha \in W$.

The function Φ is called the *reproducing kernel* of $\mathcal{H}(\Omega; W)$.

Before we choose W to be the space of symmetric matrices, let us look at the classical RKHS of real-valued functions. Here, we choose $W = \mathbb{R}$ with the usual inner product. $\mathcal{H}(\Omega; \mathbb{R})$ consists of real-valued functions and each element of $\mathcal{L}(\mathbb{R})$ can be represented by $Lx = \ell x$, $\ell \in \mathbb{R}$, i.e. $\mathcal{L}(\mathbb{R})$ can be identified with $\mathbb{R}$. The first condition $\Phi(\cdot, x)\alpha \in \mathcal{H}(\Omega; \mathbb{R})$ for all $\alpha \in \mathbb{R}$ is equivalent to $\Phi(\cdot, x) \in \mathcal{H}(\Omega; \mathbb{R})$, and the second condition is equivalent to $g(x) = \langle g, \Phi(\cdot, x)\rangle_{\mathcal{H}(\Omega;\mathbb{R})}$.

We will now consider $W = \mathbb{S}^{n\times n}$ to be the Hilbert space of all symmetric $n \times n$ matrices with inner product

$$\langle \alpha, \beta\rangle_W \quad = \quad \sum_{i,j=1}^{n} \alpha_{ij}\beta_{ij}, \qquad \alpha = (\alpha_{ij}), \beta = (\beta_{ij}). \tag{27}$$

We define $E^s_{\mu\mu}$ to be the matrix with value 1 at position (μ, μ) and value zero everywhere else. For $\mu < \nu$, we define $E^s_{\mu\nu}$ to be the matrix with value $1/\sqrt{2}$ at positions (μ, ν) and (ν, μ) and value zero everywhere else. It is easy to see that $\{E^s_{\mu\nu} : 1 \le \mu \le \nu \le n\}$ is an orthonormal basis of $\mathbb{S}^{n\times n}$. We also define $E_{\mu\nu} \in \mathbb{R}^{n\times n}$ to be the matrix with value 1 at position (μ, ν) and value zero everywhere else.

The matrix-valued Sobolev space $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ consists of all symmetric matrix-valued functions M having each component M_{ij} in $H^\sigma(\Omega)$. $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ is a RKHS with inner product given by

$$\langle M,S\rangle_{H^\sigma(\Omega;\mathbb{S}^{n\times n})} := \sum_{i,j=1}^n \langle M_{ij},S_{ij}\rangle_{H^\sigma(\Omega)}.$$

A kernel Φ is a mapping $\Phi:\Omega\times\Omega\to\mathcal{L}(\mathbb{S}^{n\times n})$ and can be represented by a tensor of order 4, i.e. we will write

$$\Phi = (\Phi_{ij\mu\nu})$$

and define its action on $\alpha\in\mathbb{S}^{n\times n}$ by

$$(\Phi(x,y)\alpha)_{ij} \quad = \quad \sum_{\mu,\nu=1}^n \Phi(x,y)_{ij\mu\nu}\alpha_{\mu\nu}. \tag{28}$$

If $\phi:\Omega\times\Omega\to\mathbb{R}$ is a reproducing kernel of $H^\sigma(\Omega)$, then

$$\Phi(x,y)_{ij\mu\nu} := \phi(x,y)\delta_{i\mu}\delta_{j\nu} \tag{29}$$

for $x,y\in\Omega$ and $1\le i,j,\mu,\nu\le n$ is a reproducing kernel of $H^\sigma(\Omega;\mathbb{S}^{n\times n})$, see [12, Lemma 3.2].

Note that $\phi(x,y)=\psi_{\ell,k}(c\|x-y\|)$, $c>0$, given by a Wendland function $\psi_{\ell,k}$, see [27], with $\ell=\lfloor\frac{n}{2}\rfloor+k+1$, $k\in\mathbb{N}$, is a reproducing kernel of the Sobolev space $H^\sigma(\Omega;\mathbb{R})$, $\sigma=k+\frac{n+1}{2}$ with equivalent norm.

Let $X=\{x_1,\ldots,x_N\}\subseteq\Omega\subseteq A(x_0)$ be pairwise distinct points. We define the fill distance by

$$h_{X,\Omega} \quad = \quad \sup_{x\in\Omega}\min_{x_i\in X}\|x-x_i\|_2. \tag{30}$$

We define the linear functionals $\lambda_k^{(i,j)}:H^\sigma(\Omega;\mathbb{S}^{n\times n})\to\mathbb{R}$ by

$$\begin{aligned}\lambda_k^{(i,j)}(M) &= e_i^\top F(M)(x_k)e_j\\ &= e_i^\top\left[Df(x_k)^\top M(x_k)+M(x_k)Df(x_k)+\nabla M(x_k)\cdot f(x_k)\right]e_j\\ &=: e_i^\top F_k(M)e_j\end{aligned} \tag{31}$$

for $x_k\in\Omega$, $1\le k\le N$ and $1\le i\le j\le n$. Here, e_i denotes the ith unit vector in $\mathbb{R}^n$ and $f\in C^s(\mathbb{R}^n,\mathbb{R}^n)$, where $s=\sigma+1$.

Let $C=(C_{ij})_{i,j=1,\ldots,n}\in\mathbb{S}^{n\times n}$ be positive definite. We seek to compute the solution S of the following optimal recovery problem: S satisfies $\lambda_k^{(i,j)}(S)=-C_{ij}, 1\le k\le N, 1\le i\le j\le n$ and minimises the norm $\|S\|_{H^\sigma(\Omega;\mathbb{S}^{n\times n})}$. It turns out that the solution can be computed by solving a system of $N\frac{n(n+1)}{2}$ linear equations.

Theorem 3.2 (see [12], Theorem 5.2)**.** *Let $\sigma>n/2+1$, $s=\sigma+1$ and let $\Phi:\Omega\times\Omega\to\mathcal{L}(\mathbb{S}^{n\times n})$ be a reproducing kernel of $H^\sigma(\Omega;\mathbb{S}^{n\times n})$. Let $X=\{x_1,\ldots,x_N\}\subseteq\Omega\subseteq A(x_0)$ be pairwise distinct points and let $\lambda_k^{(i,j)}\in H^\sigma(\Omega;\mathbb{S}^{n\times n})^*$, $1\le k\le N$, $1\le i\le j\le n$ be defined by (31).*

Then there is a unique function $S\in H^\sigma(\Omega;\mathbb{S}^{n\times n})$ solving

$$\min\left\{\|S\|_{H^\sigma(\Omega;\mathbb{S}^{n\times n})}:\lambda_k^{(i,j)}(S)=-C_{ij},1\le i\le j\le n,1\le k\le N\right\},$$

where $C=(C_{ij})_{i,j=1,\dots,n}$ is a symmetric, positive definite matrix. It has the form

$$\begin{aligned} S(x) &= \sum_{k=1}^N \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} \sum_{1\le \mu\le\nu\le n} \lambda_k^{(i,j)}(\Phi(\cdot,x)E^s_{\mu\nu})E^s_{\mu\nu} \\ &= \sum_{k=1}^N \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} \Bigg[\sum_{\mu=1}^n F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij} E_{\mu\mu} \\ &\quad + \frac{1}{2} \sum_{\mu,\nu=1,\mu\neq\nu}^n [F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} + F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}] E_{\mu\nu} \Bigg] \end{aligned} \tag{32}$$

where the coefficients $\gamma_k=(\gamma_k^{(i,j)})_{1\le i\le j\le n}$ are determined by $\lambda_\ell^{(i,j)}(S)=-C_{ij}$ for $1\le i\le j\le n$ and $1\le \ell\le N$.

If the kernel Φ is given by (29) then we also have the alternative expression

$$S(x) = \sum_{k=1}^N \sum_{i,j=1}^n \beta_k^{(i,j)} \sum_{\mu,\nu=1}^n (F_k(\Phi(\cdot,x))_{\cdot,\cdot,\mu,\nu})_{ij} E_{\mu\nu} \tag{33}$$

where the symmetric matrices $\beta_k\in\mathbb{S}^{n\times n}$ are defined by $\beta_k^{(j,i)}=\beta_k^{(i,j)}=\frac{1}{2}\gamma_k^{(i,j)}$ if $i\neq j$ and $\beta_k^{(i,i)}=\gamma_k^{(i,i)}$.

Note that the assumptions that x_0 is an exponentially stable equilibrium of $\dot{x}=f(x)$ and $X\subseteq A(x_0)$ imply that the conditions of [12, Theorem 4.6] are satisfied.

We have the following error estimate, see [12, Theorem 5.3], which makes use of Theorem 2.4. This error estimate, both on $\|F(M)-F(S)\|$ and $\|M-S\|$, shows that $F(S)$ is negative definite and S is positive definite, if the fill distance of the collocation points is small enough. Hence, we can conclude that the approximation S is itself a contraction metric.

Theorem 3.3. *Let $f\in C^s(\mathbb{R}^n;\mathbb{R}^n)$, $\mathbb{N}\ni s>n/2+2$ and set $\sigma=s-1$. Let x_0 be an exponentially stable equilibrium of $\dot{x}=f(x)$ with basin of attraction $A(x_0)$. Let $C\in\mathbb{S}^{n\times n}$ be a positive definite (constant) matrix and let $M\in C^\sigma(A(x_0),\mathbb{S}^{n\times n})$ be the solution of (2). Let $K\subseteq\Omega\subseteq A(x_0)$ be a positively invariant and compact set, where Ω is open with Lipschitz boundary. Finally, let S be the optimal recovery from Theorem 3.2. Then, we have the error estimate*

$$\|M-S\|_{L_\infty(K;\mathbb{S}^{n\times n})} \le c_1\|F(M)-F(S)\|_{L_\infty(\Omega;\mathbb{S}^{n\times n})} \le c_2 h_{X,\Omega}^{\sigma-1-n/2}\|M\|_{H^\sigma(\Omega;\mathbb{S}^{n\times n})}.$$

for all $X\subseteq\Omega$ with sufficiently small $h_{X,\Omega}$, see (30) for the definition of the fill distance. The constants c_1,c_2 do not depend on the collocation points X.

In particular, S itself is a contraction metric provided $h_{X,\Omega}$ is sufficiently small.

3.2. **Explicit formulas for the calculations.** To derive explicit formulas for the specific operators F_k, let us choose a radially symmetric kernel of the form $\phi(x,y)=\psi_0(\|x-y\|_2)$ and denote $\psi_{d+1}(r)=\frac{d\psi_d(r)/dr}{r}$ for $d=0,1$. We assume that ψ_1 and ψ_2 can be continuously extended up to $r=0$; this is, e.g. the case for (smooth enough) Wendland functions. Note that the kernel Φ is by (29) of the form

$$\Phi(\cdot,x)_{ij\mu\nu} = \psi_0(\|\cdot-x\|_2)\delta_{i\mu}\delta_{j\nu}. \tag{34}$$

Thus, for the linear operators F_k, see (31), we have

$$\begin{aligned}
(F_k(M))_{ij} &= \sum_{p=1}^{n} Df_{pi}(x_k)M_{pj}(x_k) + \sum_{p=1}^{n} M_{ip}(x_k)Df_{pj}(x_k) \\
&\quad + \sum_{p=1}^{n} \partial_p M_{ij}(x_k) f_p(x_k) \\
(F_k(\Phi(\cdot,x))_{\cdot,\cdot,\mu,\nu})_{ij} &= \sum_{p=1}^{n} \psi_0(\|x_k - x\|_2) Df_{pi}(x_k)\delta_{p\mu}\delta_{j\nu} \\
&\quad + \sum_{p=1}^{n} \psi_0(\|x_k - x\|_2)\delta_{i\mu}\delta_{p\nu} Df_{pj}(x_k) \\
&\quad + \sum_{p=1}^{n} \psi_1(\|x_k - x\|_2)(x_k - x)_p f_p(x_k)\delta_{i\mu}\delta_{j\nu} \\
&= \psi_0(\|x_k - x\|_2) Df_{\mu i}(x_k)\delta_{j\nu} + \psi_0(\|x_k - x\|_2)\delta_{i\mu} Df_{\nu j}(x_k) \\
&\quad + \psi_1(\|x_k - x\|_2)\langle x_k - x, f(x_k)\rangle \delta_{i\mu}\delta_{j\nu},
\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the standard scalar product in $\mathbb{R}^n$.

Now we can compute $S(x)$, using (33) of Theorem 3.2. We have

$$\begin{aligned}
S(x) &= \sum_{k=1}^{N} \sum_{i,j=1}^{n} \beta_k^{(i,j)} \sum_{\mu,\nu=1}^{n} (F_k(\Phi(\cdot,x))_{\cdot,\cdot,\mu,\nu})_{ij} E_{\mu\nu} \\
&= \sum_{k=1}^{N} \Bigg[\sum_{i,\mu,\nu=1}^{n} \beta_k^{(i,\nu)} \psi_0(\|x_k - x\|_2) Df_{\mu i}(x_k) E_{\mu\nu} \\
&\quad + \sum_{j,\mu,\nu=1}^{n} \beta_k^{(\mu,j)} \psi_0(\|x_k - x\|_2) Df_{\nu j}(x_k) E_{\mu\nu} \\
&\quad + \sum_{\mu,\nu=1}^{n} \beta_k^{(\mu,\nu)} \psi_1(\|x_k - x\|_2)\langle x_k - x, f(x_k)\rangle E_{\mu\nu} \Bigg] \\
&= \sum_{k=1}^{N} \Bigg[\psi_0(\|x_k - x\|_2) \left[Df(x_k)\beta_k + \beta_k Df(x_k)^\top \right] \\
&\quad + \psi_1(\|x_k - x\|_2)\langle x_k - x, f(x_k)\rangle \beta_k \Bigg]. \qquad (35)
\end{aligned}$$

Hence,

$$\begin{aligned} F(S(x)) &= \sum_{k=1}^N \psi_0(\|x_k-x\|_2)\big[Df(x)^\top Df(x_k)\beta_k + Df(x)^\top\beta_k Df(x_k)^\top \\ &\quad + Df(x_k)\beta_k Df(x) + \beta_k Df(x_k)^\top Df(x)\big] \\ &\quad + \sum_{k=1}^N \psi_1(\|x_k-x\|_2)\langle x_k-x, f(x_k)\rangle\left[Df(x)^\top\beta_k + \beta_k Df(x)\right] \\ &\quad + \sum_{k=1}^N \psi_1(\|x_k-x\|_2)\langle x-x_k, f(x)\rangle\left[Df(x_k)\beta_k + \beta_k Df(x_k)^\top\right] \\ &\quad - \sum_{k=1}^N \psi_1(\|x_k-x\|_2)\langle f(x), f(x_k)\rangle\beta_k \\ &\quad + \sum_{k=1}^N \psi_2(\|x_k-x\|_2)\langle x_k-x, f(x_k)\rangle\langle x-x_k, f(x)\rangle\beta_k. \end{aligned} \tag{36}$$

Observe that $F(S(x))$ is a symmetric matrix if all β_k, $k=1,\dots,N$ are symmetric.

After establishing the formulas for S and $F(S)$, let us now consider the linear system for the coefficients γ and β, respectively.

Let us first calculate the coefficients $b_{k,\ell,i,j,\mu,\nu}$ for $1\le k,\ell\le N$, $1\le i,j,\mu,\nu\le n$ such that

$$(F(S)(x_\ell))_{i,j} = \sum_{k=1}^N\sum_{\mu,\nu=1}^n b_{k,\ell,i,j,\mu,\nu}\beta_k^{(\mu,\nu)}. \tag{37}$$

By (36) we have

$$\begin{aligned} b_{k,\ell,i,j,\mu,\nu} &= \psi_0(\|x_k-x_\ell\|_2)\Bigg[\sum_{p=1}^n Df_{pi}(x_\ell)Df_{p\mu}(x_k)\delta_{\nu j} + Df_{\mu i}(x_\ell)Df_{j\nu}(x_k) \\ &\quad + Df_{i\mu}(x_k)Df_{\nu j}(x_\ell) + \delta_{i\mu}\sum_{p=1}^n Df_{p\nu}(x_k)Df_{pj}(x_\ell)\Bigg] \\ &\quad + \psi_1(\|x_k-x_\ell\|_2)\langle x_k-x_\ell, f(x_k)\rangle\left[Df_{\mu i}(x_\ell)\delta_{\nu j} + \delta_{i\mu}Df_{\nu j}(x_\ell)\right] \\ &\quad + \psi_1(\|x_k-x_\ell\|_2)\langle x_\ell-x_k, f(x_\ell)\rangle\left[Df_{i\mu}(x_k)\delta_{\nu j} + \delta_{i\mu}Df_{j\nu}(x_k)\right] \\ &\quad - \psi_1(\|x_k-x_\ell\|_2)\langle f(x_\ell), f(x_k)\rangle\delta_{i\mu}\delta_{j\nu} \\ &\quad + \psi_2(\|x_k-x_\ell\|_2))\langle x_k-x_\ell, f(x_k)\rangle)\langle x_\ell-x_k, f(x_\ell)\rangle\delta_{i\mu}\delta_{j\nu}. \end{aligned} \tag{38}$$

It is now easy to see that

$$b_{k,\ell,i,j,\mu,\nu} = b_{\ell,k,\mu,\nu,i,j} \text{ and} \tag{39}$$

$$b_{k,\ell,i,j,\mu,\nu} = b_{k,\ell,j,i,\nu,\mu}. \tag{40}$$

We, however, will consider a smaller linear system with unknowns $\gamma_k^{(\mu,\nu)}$ and coefficient matrix $c_{k,\ell,i,j,\mu,\nu}$

$$\begin{aligned}\sum_{k=1}^{N}\sum_{1\le\mu\le\nu\le n} c_{k,\ell,i,j,\mu,\nu}\gamma_k^{(\mu,\nu)} &= (F(S)(x_\ell))_{i,j}\\ &= \lambda_\ell^{(i,j)}(S)\\ &= -C_{ij}\end{aligned}\tag{41}$$

for $1\le\ell\le N$, $1\le i\le j\le n$. While the system (37) is of size Nn^2, the coefficient matrix $c_{k,\ell,i,j,\mu,\nu}$ is symmetric (see below) and of size $N\frac{n(n+1)}{2}$.

Let us express the $c_{k,\ell,i,j,\mu,\nu}$ in terms of the previously calculated $b_{k,\ell,i,j,\mu,\nu}$. Noting that

$$\sum_{k=1}^{N}\sum_{\mu,\nu=1}^{n} b_{k,\ell,i,j,\mu,\nu}\beta_k^{(\mu,\nu)} = \sum_{k=1}^{N}\sum_{1\le\mu\le\nu\le n} c_{k,\ell,i,j,\mu,\nu}\gamma_k^{(\mu,\nu)},\tag{42}$$

$\gamma_k^{(\mu,\mu)}=\beta_k^{(\mu,\mu)}$ and $\frac{1}{2}\gamma_k^{(\mu,\nu)}=\beta_k^{(\mu,\nu)}=\beta_k^{(\nu,\mu)}$ for $\mu\neq\nu$, we have

$$\begin{aligned}\sum_{k=1}^{N}\sum_{\mu,\nu=1}^{n} b_{k,\ell,i,j,\mu,\nu}\beta_k^{(\mu,\nu)} &= \sum_{k=1}^{N}\sum_{\mu=1}^{n} b_{k,\ell,i,j,\mu,\mu}\beta_k^{(\mu,\mu)}\\ &\quad+\sum_{k=1}^{N}\sum_{1\le\mu<\nu\le n}(b_{k,\ell,i,j,\mu,\nu}\beta_k^{(\mu,\nu)}+b_{k,\ell,i,j,\nu,\mu}\beta_k^{(\nu,\mu)})\\ &= \sum_{k=1}^{N}\sum_{\mu=1}^{n} b_{k,\ell,i,j,\mu,\mu}\gamma_k^{(\mu,\mu)}\\ &\quad+\sum_{k=1}^{N}\sum_{1\le\mu<\nu\le n}\frac{1}{2}(b_{k,\ell,i,j,\mu,\nu}+b_{k,\ell,i,j,\nu,\mu})\gamma_k^{(\mu,\nu)}.\end{aligned}\tag{43}$$

Comparing (43) to (42) gives, using (40)

$$\begin{aligned}c_{k,\ell,i,i,\mu,\mu} &= b_{k,\ell,i,i,\mu,\mu}\\ c_{k,\ell,i,i,\mu,\nu} &= \frac{1}{2}\left(b_{k,\ell,i,i,\mu,\nu}+b_{k,\ell,i,i,\nu,\mu}\right)\\ c_{k,\ell,i,j,\mu,\mu} &= b_{k,\ell,i,j,\mu,\mu}=\frac{1}{2}\left(b_{k,\ell,i,j,\mu,\mu}+b_{k,\ell,j,i,\mu,\mu}\right)\\ c_{k,\ell,i,j,\mu,\nu} &= \frac{1}{2}\left(b_{k,\ell,i,j,\mu,\nu}+b_{k,\ell,i,j,\nu,\mu}\right)\\ &= \frac{1}{4}\left(b_{k,\ell,i,j,\mu,\nu}+b_{k,\ell,j,i,\nu,\mu}+b_{k,\ell,i,j,\nu,\mu}+b_{k,\ell,j,i,\mu,\nu}\right)\end{aligned}\tag{44}$$

where we assume $\mu<\nu$ and $i<j$. The matrix $c_{k,\ell,i,j,\mu,\nu}$ is symmetric due to (39).

Summarising, for the computations we calculate the coefficients $c_{k,\ell,i,j,\mu,\nu}$ using (44) and (38). Then we determine $\gamma_k^{(\mu,\nu)}$ by solving (41) and compute $\beta_k\in\mathbb{S}^{n\times n}$ from γ_k; recall that $\beta_k^{(j,i)}=\beta_k^{(i,j)}=\frac{1}{2}\gamma_k^{(i,j)}$ if $i\neq j$ and $\beta_k^{(i,i)}=\gamma_k^{(i,i)}$. $S(x)$ and $F(S)(x)$ are then given by (35) and (36).

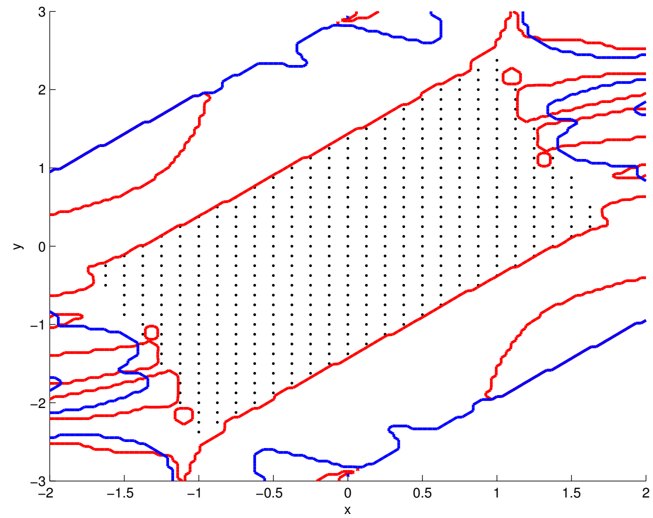

FIGURE 1. System (45) with $\epsilon = 0$. The collocation points used for the approximation together with the boundaries of the areas where $\mathrm{sign}(\mathrm{tr}\, F(S)(x, y)) - \mathrm{sign}(\det F(S)(x, y)) = -2$ (red) and $\mathrm{sign}(\mathrm{tr}\, S(x, y)) + \mathrm{sign}(\det S(x, y)) = 2$ (blue). Blue and red lines are lines where one of the requirements of a contraction metric is violated. The constructed metric is thus a valid contraction metric where the collocation points are placed, but not beyond the first red or blue line.

4. **Examples.** When applying the method to examples, we choose the symmetric and positive definite matrix C on the right-hand side of (26) to be $C = I$. Note that due to (41), choosing $C = aI$ with $a \in \mathbb{R}^+$ would result in multiplying the matrix entries of the solution $\gamma_k^{(\mu,\nu)}$ also by a and thus in the same regions where $S(x)$ and $F(S)(x)$ are positive/negative definite, respectively.

4.1. **Perturbed van der Pol.** One of the advantages of a contraction metric compared to a Lyapunov function is that a contraction metric remains a contraction metric for a perturbed system, even if the perturbation varies the position of the equilibrium.

As an example, we consider the following system with parameter ϵ

$$\left\{ \begin{array}{rcl} \dot{x} &=& -y + \epsilon \\ \dot{y} &=& x - (3 + \epsilon)(1 - x^2)y \end{array} \right. \tag{45}$$

and denote the right-hand side by $f_\epsilon(x, y)$.

In [12, Section 6.2], the system with $\epsilon = 0$, which is the classical van der Pol equation with reversed time, has been considered. It has an exponentially stable equilibrium at the origin with basin of attraction bounded by an unstable periodic orbit.

In [12] the contraction metric was constructed by approximately solving $F(M) = -I$ where $f = f_0$ is given by the right-hand side of (45) with $\epsilon = 0$. We used

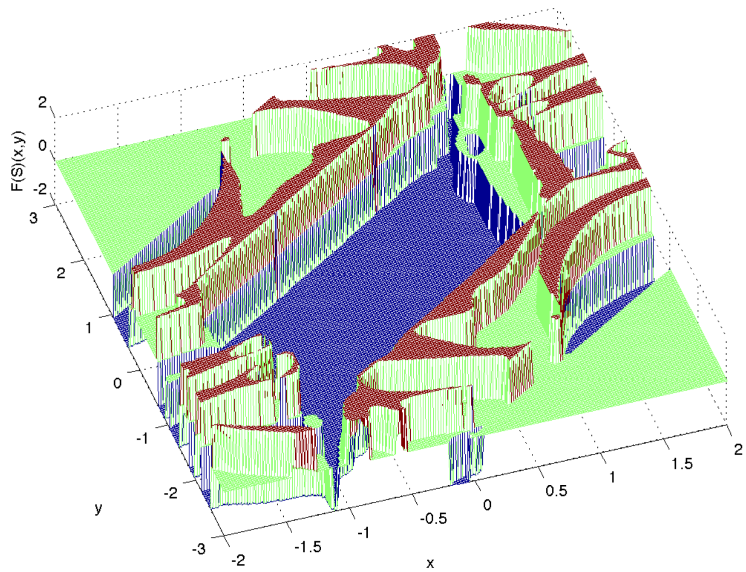

FIGURE 2. $\text{sign}(\text{tr}\, F_\epsilon(S)(x,y)) - \text{sign}(\det F_\epsilon(S)(x,y))$. If this function is -2, then $F_\epsilon(S)(x,y)$ is negative definite, which is one of the requirements for S to be a contraction metric for the system with $\epsilon = 0.1$.

the $N = 501$ collocation points $X = 0.125 \cdot \mathbb{Z}^2 \cap \{(x,y) \in \mathbb{R}^2 \mid x - 1.5 < y < 1.5 + x, -3x - 5.5 < y < -3x + 5.5\}$, and as each collocation point requires 3 variables of a symmetric 2×2 matrix, we need to solve a linear system with a 1503×1503 matrix. The kernel given by Wendland's function $\psi_{6,4}(r) = (1 - cr)_+^{10}(2145(cr)^4 + 2250(cr)^3 + 1050(cr)^2 + 250cr + 25)$ with $c = 0.9$ was used with corresponding RKHS H^σ with $\sigma = 4 + \frac{2+1}{2} = 5.5$; here $x_+ = x$ for $x \geq 0$ and $x_+ = 0$ for $x < 0$.

We need to check that the constructed matrix-valued function $S(x)$ is positive definite and $F(S)(x)$ is negative definite, where $F(S)(x) = Df(x)^\top S(x) + S(x)Df(x) + S'(x)$. To verify that a 2×2 matrix A is positive (negative) definite we check that $\text{tr}(A)$ is positive (negative) and $\det(A)$ is positive ($-\det(A)$ is negative). Figure 1 summarises the results by displaying the collocation points and the set where $\text{sign}(\text{tr}\, F(S)(x)) - \text{sign}(\det F(S)(x)) = -2$ (red), bounding the area where $F(S)(x)$ is negative definite, as well as set where $\text{sign}(\text{tr}\, S(x)) + \text{sign}(\det S(x)) = +2$ (blue), respectively, bounding the area where $S(x)$ is positive definite. If either of these two conditions is violated, the function $S(x)$ is not a contraction metric.

To illustrate the fact that the constructed Riemannian metric is also valid for a perturbed system, we will now take the metric S above, constructed for the system with $\epsilon = 0$, and show that, in a large area, S is still a contraction metric for the perturbed system f_ϵ with $\epsilon = 0.1$. Note that the origin is no equilibrium in the perturbed system.

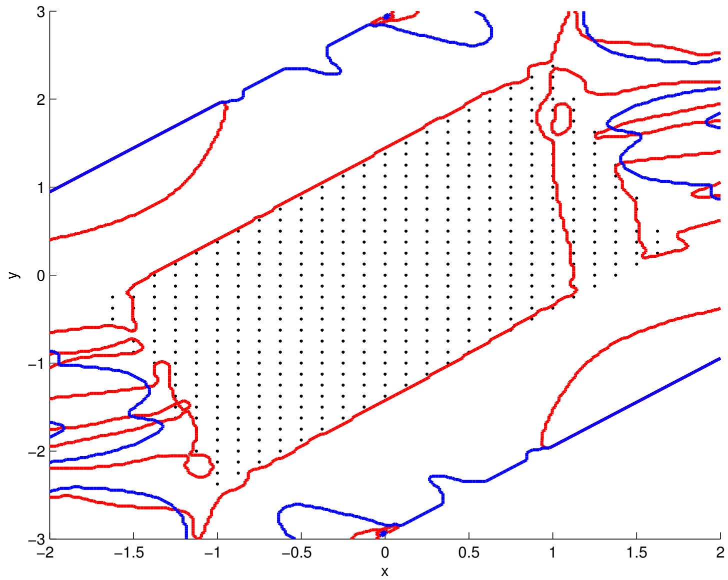

FIGURE 3. The collocation points used for the approximation with f_0 together with the areas where $\mathrm{sign}(\mathrm{tr}\, F_\epsilon(S)(x,y)) - \mathrm{sign}(\det F_\epsilon(S)(x,y)) > -2$ (red) and $\mathrm{sign}(\mathrm{tr}\, S(x,y)) + \mathrm{sign}(\det S(x,y)) < 2$ (blue). Blue and red lines are lines where one of the requirements of a contraction metric is violated. Hence, there are collocation points, where the constructed metric is not a contraction metric, since it was computed using a different system, namely with $\epsilon = 0$.

The constructed matrix-valued function $S(x)$ is the same as above, so positive definite in the same area. It remains to check where $F_\epsilon(S)(x)$ is negative definite, where $F_\epsilon(S)(x) = Df_\epsilon(x)^\top S(x) + S(x)Df_\epsilon(x) + \nabla S(x) \cdot f_\epsilon(x)$, see (36), i.e.

$$\begin{aligned} F_\epsilon(S(x)) \quad = \quad & \sum_{k=1}^N \psi_0(\|x_k - x\|)\big[Df_\epsilon(x)^\top Df(x_k)\beta_k + Df_\epsilon(x)^\top \beta_k Df(x_k)^\top \\ & + Df(x_k)\beta_k Df_\epsilon(x) + \beta_k Df(x_k)^\top Df_\epsilon(x)\big] \\ & + \sum_{k=1}^N \psi_1(\|x_k - x\|)\langle x_k - x, f(x_k)\rangle \left[Df_\epsilon(x)^\top \beta_k + \beta_k Df_\epsilon(x)\right] \\ & + \sum_{k=1}^N \psi_1(\|x_k - x\|)\langle x - x_k, f_\epsilon(x)\rangle \left[Df(x_k)\beta_k + \beta_k Df(x_k)^\top\right] \\ & - \sum_{k=1}^N \psi_1(\|x_k - x\|)\langle f_\epsilon(x), f(x_k)\rangle \beta_k \\ & + \sum_{k=1}^N \psi_2(\|x_k - x\|))\langle x_k - x, f(x_k)\rangle)\langle x - x_k, f_\epsilon(x_\ell)\rangle \beta_k. \end{aligned}$$

Figure 2 shows $\mathrm{sign}(\mathrm{tr}\, F_\epsilon(S)(x)) - \mathrm{sign}(\det F_\epsilon(S)(x))$, and Figure 3 summarises the results by displaying the collocation points and the areas where $S(x)$ is positive definite (blue) and $F_\epsilon(S)(x)$ is negative definite (red). Note that the area where $S(x)$ is a contraction metric does not contain all collocation points any more, but still covers a large part including the new equilibrium $(0.2848, 0.1)$.

4.2. **Three-dimensional example.** We consider the three-dimensional example

$$\begin{aligned} \dot{x} &= x(x^2+y^2-1) - y(z^2+1) \\ \dot{y} &= y(x^2+y^2-1) + x(z^2+1) \\ \dot{z} &= 10z(z^2-1) \end{aligned}$$

which was discussed in [6, Example 6.4] and has an exponentially stable equilibrium at the origin with basin of attraction $A(0,0,0) = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2 < 1, |z| < 1\}$.

We have used the collocation points $X = 0.13 \cdot \mathbb{Z}^3 \cap [-0.65, 0.65]^3$ with $N = 1331$ points, and as each collocation point requires 6 variables of a symmetric 3×3 matrix, we solve a linear system with a 7986×7986 matrix. We have used the kernel given by Wendland's function $\psi_{6,4}(r) = (1-cr)_+^{10}(2145(cr)^4 + 2250(cr)^3 + 1050(cr)^2 + 250cr + 25)$ with $c = 0.9$. The RKHS in this case is H^σ with $\sigma = 4 + \frac{3+1}{2} = 6$.

We need to check that the constructed matrix-valued function $S(x)$ is positive definite and $F(S)(x)$ is negative definite, where $F(S) = Df(x)^\top S(x) + S(x)Df(x) + S'(x)$. To check that a 3×3 matrix A is positive/negative definite we check the signs of the leading principal minors.

Figure 4 summarises the results by displaying the collocation points and the areas where $F(S)$ is negative definite. Note that S is positive definite in the whole area displayed.

5. **Conclusion and outlook.** In this paper we have introduced a method to construct a contraction metric for the determination of the basin of attraction of an equilibrium. The contraction metric is characterised as matrix-valued solution of a linear first-order PDE. We have shown uniqueness of solutions for this PDE as well as continuous dependence on the right-hand side. The construction method is based on approximately solving this linear, first-order PDE for a matrix-valued function by meshless collocation. This is done by choosing collocation points and solving a system of linear equations.

We have established error estimates which prove that the approximation itself is a contraction metric, provided that the collocation points are dense enough. Thus, we have proven a constructive converse theorem.

Note that the error estimates require that the collocation points are placed in the basin of attraction which is unknown. In practice, however, the numerical method using meshless collocation will always give a result, even if the collocation points are placed outside the basin of attraction, but we cannot guarantee that $S(x)$ will be positive definite and $F(S)(x)$ negative definite.

Moreover, to determine a subset of the basin of attraction, we need to find a compact and positively invariant set. This can, e.g., be achieved by determining a level set of a Lyapunov-like function, such that the orbital derivative of the function is non-positive at the level set. The computation of Lyapunov(-like) functions can be achieved by various methods, see [11]. Some methods, such as SOS [23, 1] or the revised CPA [10] method, construct a Lyapunov function which has a negative

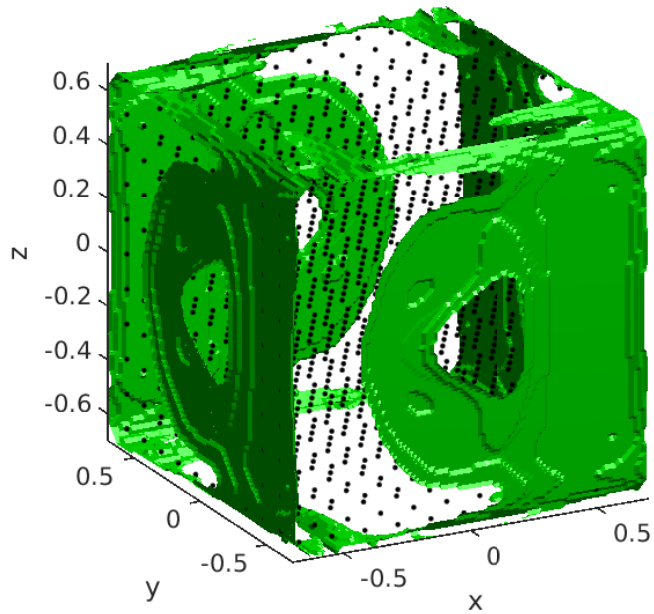

FIGURE 4. The collocation points used for the approximation together with the areas where $F(S)$ is not negative definite (green). Note that S is positive definite in the whole area displayed. Hence, the constructed metric is a contraction metric inside the cube bounded by the green areas.

orbital derivative also near the equilibrium, but even methods which fail to have negative orbital derivative near the equilibrium such as the classical CPA [13] or RBF [6] methods can be employed. The reason is that if K is a compact sublevel set of the Lyapunov(-like) function which covers the area with non-negative derivative, then K is a compact and positively invariant set. After that, we can construct a contraction metric in K, which shows that there is a unique exponentially stable equilibrium in K and K is a subset of its basin of attraction.

Hence, a combination of Lyapunov function and a contraction metric is a natural next step, see [7] for a similar idea in the context of periodic orbits. As the computation of a Lyapunov function, a scalar-valued function, is computationally less demanding than computing a matrix-valued function, one should start with computing a Lyapunov function.

The advantage of such a combined method not only solves the problem of finding a compact and positively invariant set K, but also requires the more demanding computation of a contraction metric only in a relatively small set K, thus reducing computation time. Compared to just computing a Lyapunov function, the combination with a contraction metric is robust to perturbations, as both the positive invariance of the set K and the contraction metric remain valid for a perturbed system.

REFERENCES

[1] J. Anderson and A. Papachristodoulou. Advances in computational Lyapunov analysis using sum-of-squares programming. *Discrete Contin. Dyn. Syst. Ser. B*, 8:2361–2381, 2015.

[2] E. Aylward, P. Parrilo, and J.-J. Slotine. Stability and robustness analysis of nonlinear systems via contraction metrics and SOS programming. *Automatica*, 44:2163–2170, 2008.

[3] M. Buhmann. *Radial basis functions: theory and implementations*, volume 12 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2003.

[4] C. Chicone. *Ordinary Differential Equations with Applications*. Springer, Texts in Applied Mathematics 34, 1999.

[5] F. Forni and R. Sepulchre. A differential Lyapunov framework for Contraction Analysis. *IEEE Trans. Automat. Control*, 59(3):614–628, 2014.

[6] P. Giesl. *Construction of Global Lyapunov Functions Using Radial Basis Functions*. Lecture Notes in Math. 1904, Springer, 2007.

[7] P. Giesl. On the determination of the basin of attraction of periodic orbits in three- and higher-dimensional systems. *J. Math. Anal. Appl.*, 354(2):606–618, 2009.

[8] P. Giesl. Converse theorems on contraction metrics for an equilibrium. *J. Math. Anal. Appl.*, 424(2):1380–1403, 2015.

[9] P. Giesl and S. Hafstein. Construction of a CPA contraction metric for periodic orbits using semidefinite optimization. *Nonlinear Anal.*, 86:114–134, 2013.

[10] P. Giesl and S. Hafstein. Revised CPA method to compute Lyapunov functions for nonlinear systems. *J. Math. Anal. Appl.*, 410:292–306, 2014.

[11] P. Giesl and S. Hafstein. Review on computational methods for Lyapunov functions. *Discrete Contin. Dyn. Syst. Ser. B*, 20(8):2291–2331, 2015.

[12] P. Giesl and H. Wendland. Kernel-based discretisation for solving matrix-valued PDEs. submitted.

[13] S. Hafstein. *An Algorithm for Constructing Lyapunov Functions*, volume 8 of *Electronic Journal of Differential Equations*. Texas State University - San Marcos, Department of Mathematics, San Marcos, TX, 2007. available from: http://ejde.math.txstate.edu/.

[14] W. Hahn. *Theory and application of Liapunov's direct method*. English edition prepared by Siegfried H. Lehnigk; translation by Hans H. Losenthien and Siegfried H. Lehnigk. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

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

[16] Ch. M. Kellett. Classical converse theorems in Lyapunov's second method. *Discrete Contin. Dyn. Syst. Ser. B*, 20(8):2333–2360, 2015.

[17] N. Krasovskiĭ. *Problems of the Theory of Stability of Motion*. Mir, Moskow, 1959. English translation by Stanford University Press, 1963.

[18] G. Leonov, I. Burkin, and A. Shepelyavyi. *Frequency Methods in Oscillation Theory*. Ser. Math. and its Appl.: Vol. 357, Kluwer, 1996.

[19] D. Lewis. Metric properties of differential equations. *Amer. J. Math.*, 71:294–312, 1949.

[20] W. Lohmiller and J.-J. Slotine. On Contraction Analysis for Non-linear Systems. *Automatica*, 34:683–696, 1998.

[21] A. M. Lyapunov. The general problem of the stability of motion. *Internat. J. Control*, 55(3):521–790, 1992. Translated by A. T. Fuller from Édouard Davaux's French translation (1907) of the 1892 Russian original.

[22] I. Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynam.*, 41(1-3):309–325, 2005.

[23] P. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimiziation*. PhD thesis, California Institute of Technology Pasadena, 2000.

[24] M. J. D. Powell. The theory of radial basis function approximation in 1990. In *Advances in numerical analysis, Vol. II (Lancaster, 1990)*, Oxford Sci. Publ., pages 105–210. Oxford Univ. Press, New York, 1992.

[25] A. Rantzer. A dual to Lyapunov's stability theorem. *Systems Control Lett.*, 42(3):161–168, 2001.

[26] R. Schaback and H. Wendland. Kernel techniques: from machine learning to meshless methods. *Acta Numer.*, 15:543–639, 2006.

[27] H. Wendland. Error estimates for interpolation by compactly supported radial basis functions of minimal degree. *J. Approx. Theory*, 93(2):258–272, 1998.

[28] H. Wendland. *Scattered data approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005.

[29] V. I. Zubov. *The methods of A. M. Lyapunov and their applications*. Izdat. Leningrad. Univ., Moscow, 1957.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: p.a.giesl@sussex.ac.uk
E-mail address: holger.wendland@uni-bayreuth.de