Kernel-based discretisation for solving matrix-valued PDEs

Article (Accepted Version)

KERNEL-BASED DISCRETISATION FOR SOLVING MATRIX-VALUED PDES

PETER GIESL* AND HOLGER WENDLAND†

Abstract. In this paper, we discuss the numerical solution of certain matrix-valued partial differential equations. Such PDEs arise, for example, when constructing a Riemannian contraction metric for a dynamical system given by an autonomous ODE. We develop and analyse a new meshfree discretisation scheme using kernel-based approximation spaces. However, since these approximation spaces have now to be matrix-valued, the kernels we need to use are fourth order tensors. We will review and extend recent results on even more general reproducing kernel Hilbert spaces. We will then apply this general theory to solve a matrix-valued PDE and derive error estimates for the approximate solution. The paper ends with applications to typical examples from dynamical systems.

Keywords. Meshfree Methods, Radial Basis Functions, Autonomous Systems, Contraction Metric.

AMS subject classifications. 65N35, 65N15, 37B25, 37M99

1. Introduction. Kernel-based discretisation methods provide an extremely flexible, general framework to approximate the solution to even rather unconventional problems (see for example [7, 5, 47, 13, 15, 39]). They are meshfree methods, requiring only a discrete data set for discretising the underlying domain. Since the kernel can be chosen problem dependent, it is very easy to construct in particular smooth approximation spaces and high order methods.

Kernel-based methods have extensively been used for solving partial differential equations (see for example [17, 28, 14, 46]). They have been used in the context of dynamical systems for constructing Lyapunov functions ([19, 24]) and they also play a key role in learning theory ([10, 11, 35, 40, 43, 41]) and high-dimensional integration (see for example [12]) and many other areas.

Our main motivation for extending these methods to solving matrix-valued PDEs is the following application from the theory of dynamical systems. We consider the autonomous ODE

$$\dot{x} = f(x) \tag{1.1}$$

where $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. 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 $t \geq 0$. A set $G \subseteq \mathbb{R}^n$ is called positively invariant if $S_tG \subseteq G$ 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 $f(x_0) = 0$ and its basin of attraction is defined by $A(x_0) = \{x \in \mathbb{R}^n \mid \lim_{t\to\infty} S_tx = x_0\}$.

If the equilibrium is known, then Lyapunov functions are one way of analysing the basin of attraction of the equilibrium as well as its basin of attraction, see the recent survey article [23] for constructing such Lyapunov functions. 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.

*Department of Mathematics, University of Sussex, Falmer BN1 9QH, United Kingdom, p.a.giesl@sussex.ac.uk

†Applied and Numerical Analysis, Department of Mathematics, University of Bayreuth, 95440 Bayreuth, Germany, holger.wendland@uni-bayreuth.de

A Riemannian contraction metric is a matrix-valued function $M\colon \mathbb{R}^n \to \mathbb{R}^{n\times n}$, such that $M(x)$ is symmetric and positive definite for every x. It defines a (point-dependent) scalar product on $\mathbb{R}^n$ by $\langle v,w\rangle_{M(x)} = v^T M(x) w$. For M to be a contraction metric, we require the distance between adjacent solutions of (1.1) to decrease with respect to such a contraction metric. This can be expressed by the negative definiteness of

$$F(M)(x) := Df(x)^T M(x) + M(x) Df(x) + M'(x); \tag{1.2}$$

see Theorem 1.1 below. Here, Df is the matrix of first-order deriviatives of f and M' denotes the so-called orbital derivative, i.e. it is component-wise defined to be $(M'(x))_{ij} = \nabla M(x)_{ij} \cdot f(x)$. The existence of a contraction metric in a certain set G gives information about the basin of attraction of a unique equilibrium in G.

THEOREM 1.1 ([20]). *Let $\emptyset \neq G \subseteq \mathbb{R}^n$ be a compact, connected and positively invariant set and M be a Riemannian contraction metric in G, i.e.*

- *$M \in C^1(G, \mathbb{R}^{n\times n})$, such that $M(x)$ is symmetric and positive definite for all $x \in G$.*
- *$F(M)(x)$ is negative definite for all $x \in G$.*

Then there exists one and only one equilibrium x_0 in G; x_0 is exponentially stable and G is a subset of the basin of attraction $A(x_0)$.

The difficulty of this approach is to constructively find such a contraction metric. In [20], a contraction metric is characterised as the solution of a first-order PDE of the form $F(M)(x) = -C$ for all $x \in A(x_0)$, where $C \in \mathbb{R}^{n\times n}$ is a given constant, symmetric and positive definite matrix.

As we do not know $A(x_0)$ in advance, we thus seek to reconstruct the matrix-valued function $M : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^{n\times n}$ from the matrix-valued PDE

$$F(M)(x) = -C, \qquad x \in \Omega \subseteq \mathbb{R}^n, \tag{1.3}$$

where $\Omega \subseteq \mathbb{R}^n$ is a given, sufficiently large domain. We then need to ensure that the solution M is also symmetric and positive definite.

In the accompanying paper [25], we will prove the theoretical results required in the dynamical system context. In this paper, however, we will concentrate on deriving the numerical framework for discretising even more general PDEs of the form

$$F(M)(x) = -C(x), \qquad x \in \Omega, \tag{1.4}$$

where F is not necessarily of the form (1.2) but can be a rather general differential operator which maps matrix-valued Sobolev functions of order σ to matrix-valued Sobolev functions of order τ and C is a smooth, not necessarily constant matrix-valued function.

Other applications for matrix-valued valued PDEs arise, e.g., in image processing, in particular magnetic resonance imaging in the medical sciences [8]. While many models rely on nonlinear PDEs [9], in [44] linear matrix-valued diffusion techniques are compared to nonlinear improvements. For a study of linear matrix-valued PDEs from a theoretical point of view see [34].

The paper is organised as follows. In Section 2 we will review and extend results on optimal recovery in general reproducing kernel Hilbert spaces, going far beyond the usual definition. In Section 3 we will employ these general results in the concrete situation of reproducing kernel Hilbert spaces of matrix-valued functions which are also Sobolev spaces. In Section 4 we will derive error estimates for the optimal recovery processes of solutions to (1.4). Section 5 then deals with the application to the above mentioned problem to construct a contraction metric for an autonomous system by solving (1.3). The final section gives numerical examples.

2. Optimal Recovery in Reproducing Kernel Hilbert Spaces. Reproducing kernel Hilbert Spaces (RKHS) have first been introduced to describe real-valued functions $f\colon \Omega \to \mathbb{R}$ on a domain $\Omega \subseteq \mathbb{R}^d$ (see for example [2]). They require a kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ with the reproduction property $f(x) = \langle f, \Phi(\cdot, x)\rangle_{\mathcal{H}}$ for $f \in \mathcal{H}$, $x \in \Omega$ where $\mathcal{H}$ denotes a Hilbert space of functions $f : \Omega \to \mathbb{R}$. Later, so-called matrix-valued kernels $\Phi\colon \Omega \times \Omega \to \mathbb{R}^{n\times n}$ with the reproduction property $f(x)^T\alpha = \langle f, \Phi(\cdot, x)\alpha\rangle_{\mathcal{H}}$, have been introduced to recover vector-valued functions $f\colon \Omega \to \mathbb{R}^n$ where $\mathcal{H}$ denotes a Hilbert space of functions $\Omega \to \mathbb{R}^n$ and $\alpha \in \mathbb{R}^n$ is an arbitrary vector (see for example [1, 4, 18, 33, 36, 48]).

In this paper, we are interested in reproducing kernel Hilbert spaces of matrix-valued functions. While it is possible to describe such Hilbert spaces using vector-valued functions, it is, in particular when it comes to the consideration of subspaces, much cleaner to take a broader point of view and employ a more general approach, which we will shortly describe now. More details and applications in learning theory can, for example, be found in [35] and the literature therein.

Let W be a real Hilbert space and denote the linear space of all linear and bounded operators $L : W \to W$ by $\mathcal{L}(W)$. For any $L \in \mathcal{L}(W)$, we will denote the adjoint operator by $L^* \in \mathcal{L}(W)$. Let $\Omega \subseteq \mathbb{R}^d$ be a given domain and let $\mathcal{H}(\Omega; W)$ be a Hilbert space of W-valued functions $f : \Omega \to W$.

Definition 2.1. *The Hilbert space $\mathcal{H}(\Omega; W)$ 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. *$\langle f(x), \alpha\rangle_W = \langle f, \Phi(\cdot, x)\alpha\rangle_{\mathcal{H}}$ for all $f \in \mathcal{H}(\Omega; W)$, all $x \in \Omega$ and all $\alpha \in W$.*

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

The following results are proven as in the real-valued case, see [35] for details.

Lemma 2.2.

1. *The reproducing kernel Φ of a Hilbert space $\mathcal{H}(\Omega; W)$ is uniquely determined.*
2. *The reproducing kernel satisfies $\Phi(x, y)^* = \Phi(y, x)$ for all $x, y \in \Omega$.*
3. *The reproducing kernel is positive semi-definite, i.e. it satisfies*

$$\sum_{i,j=1}^{N} \langle \alpha_i, \Phi(x_i, x_j)\alpha_j\rangle_W \geq 0$$

for all $x_1, \dots, x_N \in \Omega$ and all $\alpha_1, \dots, \alpha_N \in W$.

If the functions $\Phi(\cdot, x_j)\alpha_j$ are linearly independent, the kernel is even positive definite in the sense of the following definition.

Definition 2.3. *A kernel $\Phi : \Omega \times \Omega \to \mathcal{L}(W)$ which satisfies $\Phi(x, y)^* = \Phi(y, x)$ for all $x, y \in \Omega$ is called* positive definite *if for all $N \in \mathbb{N}$, for all $x_1, \dots, x_N \in \Omega$, pairwise distinct, and for all $\alpha_1, \dots, \alpha_N \in W$, not all of them zero, we have*

$$\sum_{i,j=1}^{N} \langle \alpha_i, \Phi(x_i, x_j)\alpha_j\rangle_W > 0.$$

As usual in the theory of reproducing kernel Hilbert spaces, it is also possible to start with a kernel and to build its Hilbert space from scratch. This is done as follows. Suppose we have a positive definite kernel $\Phi : \Omega \times \Omega \to \mathcal{L}(W)$ as in Definition 2.3. Then, we can form the space

$$\mathcal{F}_\Phi(\Omega; W) = \operatorname{span}\{\Phi(\cdot, x)\alpha : x \in \Omega, \alpha \in W\}$$

and equip this space with an inner product defined by

$$\langle \Phi(\cdot, x)\alpha, \Phi(\cdot, y)\beta\rangle_\Phi := \langle \Phi(x, y)\beta, \alpha\rangle_W.$$

The closure of $\mathcal{F}_\Phi(\Omega; W)$ with respect to the norm induced by this inner product is then the corresponding Hilbert space $\mathcal{H}(\Omega; W)$ for which Φ is the reproducing kernel.

Within this general framework, we now want to discuss the more general concept of optimal recovery. Hence, let $\mathcal{H}(\Omega; W)$ be our reproducing kernel Hilbert space with reproducing kernel $\Phi : \Omega \times \Omega \to \mathcal{L}(W)$. As usual, we denote the dual of $\mathcal{H}(\Omega; W)$ by $\mathcal{H}(\Omega; W)^*$.

Definition 2.4. *Given N linearly independent functionals $\lambda_1, \dots, \lambda_N \in \mathcal{H}(\Omega; W)^*$ and N values $f_1 = \lambda_1(f), \dots, f_N = \lambda_N(f) \in \mathbb{R}$ generated by an element $f \in \mathcal{H}(\Omega; W)$. The optimal recovery of f based on this information is defined to be the element $s^* \in \mathcal{H}(\Omega; W)$ which solves*

$$\min\{\|s\|_{\mathcal{H}} : s \in \mathcal{H}(\Omega; W) \text{ with } \lambda_j(s) = f_j, 1 \le j \le N\}.$$

The solution to this minimisation problem is well-known and follows directly from standard Hilbert space theory; it works in any Hilbert space, not only in reproducing kernel Hilbert spaces. We quote the following result from [47, Theorem 16.1]:

Theorem 2.5. *Let H be a Hilbert space. Let $\lambda_1, \dots, \lambda_N \in H^*$ be linearly independent linear functionals with Riesz representers $v_1, \dots, v_N \in H$. Then the element $s^* \in H$ which solves*

$$\min\{\|s\|_H : s \in H \text{ with } \lambda_j(s) = f_j, 1 \le j \le N\}$$

is given by

$$s^* = \sum_{k=1}^N \beta_k v_k,$$

where the coefficients $\beta_k \in \mathbb{R}$ are determined by the generalised interpolation conditions $\lambda_i(s^*) = f_i$, $1 \le i \le N$, which lead to the linear system $A_\Lambda \beta = f$ with the positive definite matrix $A_\Lambda = (a_{ik})$ having entries $a_{ik} = \lambda_i(v_k) = \langle v_k, v_i\rangle_H$.

If we want to to apply this general result to our specific situation $H = \mathcal{H}(\Omega; W)$ then we need to know the Riesz representers of the functionals $\lambda \in \mathcal{H}(\Omega; W)^*$. In the case of a separable Hilbert space W the Riesz representers are given as stated in the next Proposition.

Proposition 2.6. *Assume that the Hilbert space W is separable and that $\{\alpha_j\}_{j\in J}$ is an orthonormal basis of W. Then, the Riesz representer of a functional $\lambda \in \mathcal{H}(\Omega; W)^*$ is given by*

$$v_\lambda(x) = \sum_{j\in J} \lambda(\Phi(\cdot, x)\alpha_j)\alpha_j, \qquad x \in \Omega.$$

Proof. Since $v_\lambda(x) \in W$ for every $x \in \Omega$ and since $\{\alpha_j\}_{j\in J}$ is an orthonormal basis of W, we can expand $v_\lambda(x)$ within this basis using its Fourier representation

$$v_\lambda(x) = \sum_{j\in J} \langle v_\lambda(x), \alpha_j\rangle_W \alpha_j.$$

The result then follows immediately from the reproducing kernel property:

$$\langle v_\lambda(x), \alpha_j\rangle_W = \langle v_\lambda, \Phi(\cdot, x)\alpha_j\rangle_{\mathcal{H}} = \langle \Phi(\cdot, x)\alpha_j, v_\lambda\rangle_{\mathcal{H}} = \lambda(\Phi(\cdot, x)\alpha_j).$$

□

Thus, the optimal recovery problem can be recast as a linear system. From now on, we will write $\lambda^y(\Phi(y, x)\alpha)$ to indicate that the functional λ acts on the variable y of the kernel.

COROLLARY 2.7. *Assume that $\{\alpha_j\}_{j\in J}$ is an orthonormal basis of W. The solution of the minimisation problem of Theorem 2.5 is given by*

$$s^* = \sum_{k=1}^{N} \beta_k \sum_{j\in J} \lambda_k^y(\Phi(y, \cdot)\alpha_j)\alpha_j,$$

and the coefficients $\beta_k \in \mathbb{R}$ are determined by

$$\sum_{k=1}^{N} \lambda_i^x \left[\lambda_k^y \sum_{j\in J} \left(\Phi(y, x)\alpha_j\right)\alpha_j \right] \beta_k = f_i, \qquad 1 \le i \le N.$$

3. Matrix-Valued Theory. After establishing the general theory, we will, in this section, consider special cases to which we will apply the main result of the previous section stated in Corollary 2.7.

To be more precise, we will choose W to be the space $\mathbb{R}^{n\times n}$ of real-valued $n \times n$ matrices or its subspace $\mathbb{S}^{n\times n}$ of symmetric matrices. Moreover, we will consider specific RKHS spaces, namely matrix-valued Sobolev spaces $H^\sigma(\Omega; \mathbb{S}^{n\times n})$, where the kernel is built from the kernel of the corresponding real-valued Sobolev space. The next section is then devoted to specific functionals and an error analysis.

We start this section by setting $W = \mathbb{R}^{n\times n}$ or $W = \mathbb{S}^{n\times n}$, the space of all symmetric $n \times n$ matrices. On W we define the following inner product to make it a Hilbert space.

$$\langle \alpha, \beta\rangle_W = \sum_{i,j=1}^{n} \alpha_{ij}\beta_{ij} = \mathrm{tr}(\alpha\beta^T), \qquad \alpha = (\alpha_{ij}), \beta = (\beta_{ij}). \tag{3.1}$$

According to the general theory of the last section, a kernel Φ is now a mapping $\Phi : \Omega \times \Omega \to \mathcal{L}(\mathbb{R}^{n\times n})$ and can be represented by a tensor of order 4. To this end, we will write $\Phi = (\Phi_{ijk\ell})$ and define its action on $\alpha \in \mathbb{R}^{n\times n}$ by

$$(\Phi(x, y)\alpha)_{ij} = \sum_{k,\ell=1}^{n} \Phi(x, y)_{ijk\ell}\alpha_{k\ell}. \tag{3.2}$$

By the second statement of Lemma 2.2, a necessary requirement for the kernel is the adjoint condition $\langle \Phi(x, y)\alpha, \beta\rangle_W = \langle \alpha, \Phi(y, x)\beta\rangle_W$, which means in the given situation

$$\sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n}\Phi(x,y)_{ijk\ell}\alpha_{k\ell}\beta_{ij} = \sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n}\Phi(y,x)_{ijk\ell}\alpha_{ij}\beta_{k\ell}$$
$$= \sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n}\Phi(y,x)_{k\ell ij}\alpha_{k\ell}\beta_{ij}.$$

Hence, we require our tensor kernel to satisfy

$$\Phi(x,y)_{ijk\ell} = \Phi(y,x)_{k\ell ij}. \tag{3.3}$$

This will motivate the choice of a kernel in (3.6) later on. The kernel Φ is positive definite, see Definition 2.3, if

$$\sum_{\mu,\nu=1}^{N}\langle\alpha^{(\nu)},\Phi(x_\nu,x_\mu)\alpha^{(\mu)}\rangle_W = \sum_{\mu,\nu=1}^{N}\sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n}\Phi(x_\nu,x_\mu)_{ijk\ell}\alpha_{ij}^{(\nu)}\alpha_{k\ell}^{(\mu)} \geq 0 \tag{3.4}$$

and the sum is positive if not all of the $\alpha^{(\nu)}$ are zero. The associated reproducing kernel Hilbert space $\mathcal{H}(\Omega;W) = \mathcal{H}(\Omega;\mathbb{R}^{n\times n})$ consists of matrix-valued functions.

Finally, for a given functional $\lambda \in \mathcal{H}(\Omega;\mathbb{R}^{n\times n})^*$, we can write its Riesz representer as follows. Let $E_{\mu\nu} \in \mathbb{R}^{n\times n}$ be the matrix with value 1 at position (μ,ν) and value zero everywhere else. Then, $\{E_{\mu\nu} : 1 \leq \mu,\nu \leq n\}$ is an orthonormal basis of $W = \mathbb{R}^{n\times n}$ and the Riesz representer of λ hence becomes, by Proposition 2.6,

$$v_\lambda(x) = \sum_{\mu,\nu=1}^{n}\lambda(\Phi(\cdot,x)E_{\mu\nu})E_{\mu\nu}, \qquad x \in \Omega.$$

In the case of symmetric matrices, we can proceed quite similarly. However, we need to consider a different orthonormal basis, namely $\{E_{\mu\nu}^s : 1 \leq \mu \leq \nu \leq n\}$. We define $E_{\mu\mu}^s$ to be the matrix with value 1 at position (μ,μ) and value zero everywhere else. For $\mu < \nu$, we define $E_{\mu\nu}^s$ 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 $\{E_{\mu\nu}^s : 1 \leq \mu \leq \nu \leq n\}$ is an orthonormal basis of $W = \mathbb{S}^{n\times n}$.

For a given functional $\lambda \in \mathcal{H}(\Omega;\mathbb{S}^{n\times n})^*$, the Riesz representer of λ is, by Proposition 2.6, hence given by

$$v_\lambda(x) = \sum_{1\leq\mu\leq\nu\leq n}\lambda(\Phi(\cdot,x)E_{\mu\nu}^s)E_{\mu\nu}^s, \qquad x \in \Omega. \tag{3.5}$$

In the following, we will be concerned with specific functionals defined on specific reproducing kernel Hilbert spaces. We end this section with discussing the spaces. The functionals will be subject of the next section.

Throughout this paper, we will assume that $H^\sigma(\Omega)$ denotes the Sobolev space of order $\sigma > d/2$, where the weak derivatives are measured in the $L_2(\Omega)$-norm. However, σ does not necessarily have to be an integer and the space can then be defined, for example, by interpolation. We will always assume that $\sigma > d/2$ such that the Sobolev embedding theorem yields $H^\sigma(\Omega) \subseteq C(\Omega)$ which particularly means that $H^\sigma(\Omega)$ has a reproducing kernel. The kernel is uniquely determined by the inner product. However, it is possible to define equivalent norms on $H^\sigma(\Omega)$ using other inner products. This then leads to other reproducing kernels. Examples of such kernels comprise the Sobolev (or Matérn) kernels and Wendland's radial basis functions (see [13, 45, 38]). We will also assume that $\Omega \subseteq \mathbb{R}^d$ is a bounded domain with a boundary which is at least Lipschitz continuous.

Definition 3.1. *Let $\Omega \subseteq \mathbb{R}^d$ and $\sigma > d/2$ be given. Then, the matrix-valued Sobolev space $H^\sigma(\Omega;\mathbb{R}^{n\times n})$ consists of all matrix-valued functions M having each component M_{ij} in $H^\sigma(\Omega)$. Similarly, the Sobolev space $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ consists of all symmetric matrix-valued functions M having each component M_{ij} in $H^\sigma(\Omega)$.*

$H^\sigma(\Omega;\mathbb{R}^{n\times n})$ and $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ are Hilbert spaces with inner product given by

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

the same inner product can be used for $H^\sigma(\Omega;\mathbb{S}^{n\times n})$. They are also reproducing kernel Hilbert spaces. The next result shows that a reproducing kernel of such a space can simply be given by using a *diagonal* kernel.

Lemma 3.2. *Let $\Omega \subseteq \mathbb{R}^d$ and $\sigma > d/2$ be given. Assume that $\phi : \Omega \times \Omega \to \mathbb{R}$ is a reproducing kernel of $H^\sigma(\Omega)$. Then, $H^\sigma(\Omega;\mathbb{R}^{n\times n})$ and $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ are also reproducing kernel Hilbert spaces with reproducing kernel Φ defined by*

$$\Phi(x,y)_{ijk\ell} := \phi(x,y)\delta_{ik}\delta_{j\ell} \tag{3.6}$$

for $x, y \in \Omega$ and $1 \le i, j, k, \ell \le n$.

Proof. We have to verify the two defining properties of a reproducing kernel given in Definition 2.1. First of all, we obviously have $\Phi(\cdot, x)\alpha \in H^\sigma(\Omega;\mathbb{R}^{n\times n})$ for all $x \in \Omega$ and all $\alpha \in \mathbb{R}^{n\times n}$ since

$$(\Phi(\cdot,x)\alpha)_{ij} = \sum_{k,\ell=1}^{n} \Phi(\cdot,x)_{ijk\ell}\alpha_{k\ell} = \sum_{k,\ell=1}^{n} \phi(\cdot,x)\delta_{ik}\delta_{j\ell}\alpha_{k\ell} = \phi(\cdot,x)\alpha_{ij}$$

and ϕ is a reproducing kernel of $H^\sigma(\Omega)$. For $H^\sigma(\Omega;\mathbb{S}^{n\times n})$, note that $\Phi(\cdot,x)\alpha$ is symmetric if α is symmetric.

Secondly, we have the reproduction property. If once again $\alpha \in \mathbb{R}^{n\times n}$ and $f \in H^\sigma(\Omega;\mathbb{R}^{n\times n})$ then the computation just made shows

$$\begin{aligned}
\langle f, \Phi(\cdot,x)\alpha\rangle_{H^\sigma(\Omega;\mathbb{R}^{n\times n})} &= \sum_{i,j=1}^{n} \langle f_{ij}, (\Phi(\cdot,x)\alpha)_{ij}\rangle_{H^\sigma(\Omega)} \\
&= \sum_{i,j=1}^{n} \langle f_{ij}, \phi(\cdot,x)\alpha_{ij}\rangle_{H^\sigma(\Omega)} \\
&= \sum_{i,j=1}^{n} \alpha_{ij} f_{ij}(x) = \langle f(x), \alpha\rangle_{\mathbb{R}^{n\times n}},
\end{aligned}$$

using the reproduction property of ϕ in $H^\sigma(\Omega)$. The proof for $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ is the same. ☐

Corollary 3.3. *Let the assumptions of Lemma 3.2 hold with a positive definite kernel $\phi : \Omega \times \Omega \to \mathbb{R}$. Then, also the tensor-valued kernel Φ is positive definite.*

Proof. The kernel is positive definite in the sense of (3.4), since we have

$$\sum_{\mu,\nu=1}^{N}\sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n}\Phi(x_\nu,x_\mu)_{ijk\ell}\alpha_{ij}^{(\nu)}\alpha_{k\ell}^{(\mu)} = \sum_{\mu,\nu=1}^{N}\sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n}\phi(x_\nu,x_\mu)\delta_{ik}\delta_{j\ell}\alpha_{ij}^{(\nu)}\alpha_{k\ell}^{(\mu)}$$
$$= \sum_{i,j=1}^{n}\sum_{\mu,\nu=1}^{N}\phi(x_\nu,x_\mu)\alpha_{ij}^{(\nu)}\alpha_{ij}^{(\mu)} \geq 0$$

and at least one of the inner sums is positive. □

4. Error Analysis of the Reconstruction Process. After having specified the reproducing kernel Hilbert spaces in the last section, we will now analyse the error of the reconstruction process of Theorem 2.5 in this specific setting. To this end, we have to define the relevant functionals on $H^\sigma(\Omega;\mathbb{R}^{n\times n})$ and $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ that we are interested in. Note that using a kernel of the form (3.6) together with point evaluations would simply lead to a component-wise treatment. In such a situation, dealing with each component separately would be more efficient. Here, however, we are interested in the following situation. Suppose $F : H^\sigma(\Omega;\mathbb{R}^{n\times n}) \to H^\tau(\Omega;\mathbb{R}^{n\times n})$ (or $F : H^\sigma(\Omega;\mathbb{S}^{n\times n}) \to H^\tau(\Omega;\mathbb{S}^{n\times n})$) is a linear and bounded map, i.e. there is a constant $C > 0$ such that

$$\|F(M)\|_{H^\tau(\Omega;\mathbb{R}^{n\times n})} \leq C\|M\|_{H^\sigma(\Omega;\mathbb{R}^{n\times n})}, \qquad M \in H^\sigma(\Omega;\mathbb{R}^{n\times n}).$$

Suppose further that $\tau > d/2$ so that $F(M) \in C(\Omega;\mathbb{R}^{n\times n})$ is continuous. Then, we can define functionals of the form

$$\lambda_k^{(i,j)}(M) = e_i^T F(M)(x_k)e_j$$

for $1 \leq i,j \leq n$ (or $1 \leq i \leq j \leq n$ for $\mathbb{S}^{n\times n}$) and $1 \leq k \leq N$, where $X = \{x_1,\dots,x_N\}$ is a given discrete point set in Ω. We will specify the mapping F later on but we can derive a general theory using just these assumptions.

To derive our error estimates, we will follow general ideas from scattered data approximation. In particular, we will measure the error in terms of the so-called fill distance or mesh norm

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

This means that we can derive the classical error estimates based upon sampling inequalities also in this case. We will require the following result (see [37]).

LEMMA 4.1. *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz continuous boundary. Let $\sigma > d/2$ and let $X = \{x_1,\dots,x_N\} \subseteq \Omega$. If $f \in H^\sigma(\Omega)$ vanishes on X, then there is a constant $C > 0$ independent of X and f such that*

$$\|f\|_{L_\infty(\Omega)} \leq Ch_{X,\Omega}^{\sigma-d/2}\|f\|_{H^\sigma(\Omega)}.$$

We can now use this result component-wise to derive estimates for the matrix-valued set-up. We will do this immediately for the situation we are interested in, which gives our first main result of this paper.

THEOREM 4.2. *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz continuous boundary. Let either $W = \mathbb{R}^{n\times n}$ or $W = \mathbb{S}^{n\times n}$. Let $\sigma, \tau > d/2$ be given and F :*

$H^\sigma(\Omega;W) \to H^\tau(\Omega;W)$ *be linear and bounded. Finally, let* $X = \{x_1,\dots,x_N\} \subseteq \Omega$ *be given and let*

$$\lambda_k^{(i,j)}(M) := e_i^T F(M)(x_k)e_j, \qquad 1 \le k \le N, \quad \begin{cases} 1 \le i,j \le n & \text{if } W = \mathbb{R}^{n\times n}, \\ 1 \le i \le j \le n & \text{if } W = \mathbb{S}^{n\times n}. \end{cases}$$

Then each $\lambda_k^{(i,j)}$ *belongs to the dual of* $H^\sigma(\Omega;W)$.

Let us further assume that the functionals are linearly independent. If S *denotes the optimal recovery of* $M \in H^\sigma(\Omega;W)$ *in the sense of Definition 2.4 using these functionals and a reproducing kernel of* $H^\sigma(\Omega;W)$ *then*

$$\|F(M) - F(S)\|_{L_\infty(\Omega;\mathbb{R}^{n\times n})} \le C h_{X,\Omega}^{\tau-d/2} \|M\|_{H^\sigma(\Omega;\mathbb{R}^{n\times n})},$$

where $\|A\|_{L_\infty(\Omega;\mathbb{R}^{n\times n})} = \max_{i,j=1,\dots,n} \|a_{ij}\|_{L_\infty(\Omega)}$.

Proof. We only consider the case $W = \mathbb{R}^{n\times n}$ as the proof for $W = \mathbb{S}^{n\times n}$ is essentially the same. Obviously, the $\lambda_k^{(i,j)}$ are linear. Because of our assumptions, $F(M)$ is indeed continuous by the Sobolev embedding theorem, i.e. the functionals are well-defined. Furthermore,

$$|\lambda_k^{(i,j)}(M)| \le C\|F(M)\|_{H^\tau(\Omega;\mathbb{R}^{n\times n})} \le C\|M\|_{H^\sigma(\Omega;\mathbb{R}^{n\times n})}, \qquad M \in H^\sigma(\Omega;\mathbb{R}^{n\times n}),$$

by the Sobolev embedding theorem and by the continuity of F. This means that all functionals indeed belong to the dual of $H^\sigma(\Omega;\mathbb{R}^{n\times n})$.

For the error estimate we note that the matrix-valued function $F(M) - F(S) \in H^\tau(\Omega;\mathbb{R}^{n\times n})$ vanishes on the data set X. Hence, we can apply Lemma 4.1 to each component of $F(M) - F(S)$ yielding

$$\begin{aligned} \|F(M) - F(S)\|_{L_\infty(\Omega;\mathbb{R}^{n\times n})} &\le C h_{X,\Omega}^{\tau-d/2} \|F(M-S)\|_{H^\tau(\Omega;\mathbb{R}^{n\times n})} \\ &\le C h_{X,\Omega}^{\tau-d/2} \|M-S\|_{H^\sigma(\Omega;\mathbb{R}^{n\times n})} \\ &\le C h_{X,\Omega}^{\tau-d/2} \|M\|_{H^\sigma(\Omega;\mathbb{R}^{n\times n})}, \end{aligned}$$

using also the continuity of F and the fact that S is the $H^\sigma(\Omega;\mathbb{R}^{n\times n})$ optimal recovery of M. ∎

To show linear independence, we follow the scalar-valued case [24] and define singular points for a general linear differential operator F, mapping matrix-valued functions to matrix-valued functions. We will then apply the rather general result of Theorem 4.2 to a particular class of operators F.

Definition 4.3. *Let* $n,d \in \mathbb{N}$, $\Omega \subseteq \mathbb{R}^d$, $\sigma > m + d/2$ *and* $\tau = \sigma - m$. *Let* $W = \mathbb{R}^{n\times n}$ *or* $W = \mathbb{S}^{n\times n}$. *Let* $F : H^\sigma(\Omega;W) \to H^\tau(\Omega;W)$ *be a differential operator of degree* m *of the form*

$$F(M)(x) = \sum_{|\alpha|\le m} c_\alpha(x)[D^\alpha M(x)]$$

where D^α *is applied component-wise and* $c_\alpha : \Omega \to \mathcal{L}(W)$ *is such that* $x \mapsto c_\alpha(x)[D^\alpha M(x)] \in H^\tau(\Omega;W)$ *for every* $M \in H^\sigma(\Omega;W)$.

We define x *to be a* singular point *of* F *if for all* $|\alpha| \le m$ *the linear map* $c_\alpha(x)$ *is not invertible.*

In the next lemma we will show symmetry properties for F, defined on the symmetric matrices, which will later be needed for explicit calculations.

LEMMA 4.4. *Assume that $F : H^\sigma(\Omega;\mathbb{S}^{n\times n}) \to H^\tau(\Omega;\mathbb{S}^{n\times n})$ is a differential operator as in Definition 4.3, i.e. in particular $c_\alpha(x)(M) \in \mathbb{S}^{n\times n}$ for $M \in \mathbb{S}^{n\times n}$. Assume furthermore that the kernel $\Phi(x,y)_{ijk\ell} = \phi(x,y)\delta_{ik}\delta_{j\ell}$ from (3.6) is used. Then*

$$F(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} = F(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ji}. \tag{4.1}$$

Proof. The linear map $c_\alpha(x)$ can, similar to (3.2), be described by a tensor of order 4, i.e.

$$(c_\alpha(x)(M))_{ij} = \sum_{k,\ell=1}^{n} c_\alpha(x)_{ijk\ell} M_{k\ell}. \tag{4.2}$$

We show that we can assume

$$c_\alpha(x)_{ijk\ell} = c_\alpha(x)_{ij\ell k} \tag{4.3}$$

for all $x \in \Omega$ without loss of generality. Indeed, let c_α be given satisfying (4.2) and define $\tilde{c}_\alpha$ by

$$\tilde{c}_\alpha(x)_{ijk\ell} := \frac{1}{2}\left(c_\alpha(x)_{ijk\ell} + c_\alpha(x)_{ij\ell k}\right).$$

It is clear that $\tilde{c}$ satisfies (4.3) and we also have, using $M \in \mathbb{S}^{n\times n}$,

$$\begin{aligned}
\sum_{k,\ell=1}^{n} \tilde{c}_\alpha(x)_{ijk\ell} M_{k\ell} &= \sum_{k=1}^{n} \tilde{c}_\alpha(x)_{ijkk} M_{kk} + \sum_{1\le k<\ell\le n} \tilde{c}_\alpha(x)_{ijk\ell}[M_{k\ell} + M_{\ell k}] \\
&= \sum_{k=1}^{n} c_\alpha(x)_{ijkk} M_{kk} + 2 \sum_{1\le k<\ell\le n} \tilde{c}_\alpha(x)_{ijk\ell} M_{k\ell} \\
&= \sum_{k=1}^{n} c_\alpha(x)_{ijkk} M_{kk} + \sum_{1\le k<\ell\le n} \left(c_\alpha(x)_{ijk\ell} + c_\alpha(x)_{ij\ell k}\right) M_{k\ell} \\
&= \sum_{k,\ell=1}^{n} c_\alpha(x)_{ijk\ell} M_{k\ell} = (c_\alpha(x)(M))_{ij}.
\end{aligned}$$

For $M \in \mathbb{S}^{n\times n}$ we have $c_\alpha(x)(M) \in \mathbb{S}^{n\times n}$ and hence

$$\begin{aligned}
\sum_{k,\ell=1}^{n} c_\alpha(x)_{ijk\ell} M_{k\ell} = (c_\alpha(x)(M))_{ij} = (c_\alpha(x)(M))_{ji} &= \sum_{k,\ell=1}^{n} c_\alpha(x)_{jik\ell} M_{k\ell} \\
= \sum_{k,\ell=1}^{n} c_\alpha(x)_{jik\ell} M_{\ell k} &= \sum_{k,\ell=1}^{n} c_\alpha(x)_{ji\ell k} M_{k\ell}
\end{aligned}$$

as $M \in \mathbb{S}^{n\times n}$. Choosing $M = E^s_{\mu\nu}$ to be a basis "vector" of $\mathbb{S}^{n\times n}$ shows, using (4.3),

$$\begin{aligned}
\sum_{k,\ell=1}^{n} c_\alpha(x)_{ijk\ell} (E^s_{\mu\nu})_{k\ell} &= \frac{1}{\sqrt{2}}\left[c_\alpha(x)_{ij\mu\nu} + c_\alpha(x)_{ij\nu\mu}\right] = \sqrt{2} c_\alpha(x)_{ij\mu\nu}, \\
\sum_{k,\ell=1}^{n} c_\alpha(x)_{ji\ell k} (E^s_{\mu\nu})_{k\ell} &= \frac{1}{\sqrt{2}}\left[c_\alpha(x)_{ji\nu\mu} + c_\alpha(x)_{ji\mu\nu}\right] = \sqrt{2} c_\alpha(x)_{ji\nu\mu},
\end{aligned}$$

i.e.

$$c_\alpha(x)_{ijk\ell} = c_\alpha(x)_{ji\ell k}. \tag{4.4}$$

For (4.1) note that

$$D^\alpha \Phi(\cdot, x)_{i,j,\mu,\nu} = D^\alpha \phi(\cdot, x)\delta_{i\mu}\delta_{j\nu}$$

so that

$$\begin{aligned}
F(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} &= \sum_{|\alpha|\le m} D^\alpha\phi(\cdot,x)\sum_{k,\ell=1}^n c_\alpha(\cdot)_{ijk\ell}\delta_{k\mu}\delta_{\ell\nu} = \sum_{|\alpha|\le m} D^\alpha\phi(\cdot,x)c_\alpha(\cdot)_{ij\mu\nu}\\
&= \sum_{|\alpha|\le m} D^\alpha\phi(\cdot,x)c_\alpha(\cdot)_{ji\nu\mu} = \sum_{|\alpha|\le m} D^\alpha\phi(\cdot,x)\sum_{k,\ell=1}^n c_\alpha(\cdot)_{jik\ell}\delta_{k\nu}\delta_{\ell\mu}\\
&= F(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ji},
\end{aligned}$$

where we have used (4.4). ☐

Proposition 4.5. *Let $\sigma > m + d/2$ and F be a linear differential operator $F : H^\sigma(\Omega;\mathbb{R}^{n\times n}) \to H^\tau(\Omega;\mathbb{R}^{n\times n})$ ($F : H^\sigma(\Omega;\mathbb{S}^{n\times n}) \to H^\tau(\Omega;\mathbb{S}^{n\times n})$) as in Definition 4.3. Let $X = \{x_1,\dots,x_N\}$ be a set of pairwise distinct points which are not singular points of F. Then the functionals*

$$\lambda_k^{(i,j)}(M) := e_i^T F(M)(x_k)e_j, \qquad 1 \le k \le N, 1 \le i,j \le n \quad (1 \le i \le j \le n).$$

are bounded and linearly independent over $H^\sigma(\Omega;\mathbb{R}^{n\times n})$ ($H^\sigma(\Omega;\mathbb{S}^{n\times n})$).

Proof. The boundedness of the functionals is clear from the assumptions. We will prove the linear independence of the functionals over $H^\sigma(\Omega;\mathbb{S}^{n\times n})$. In Theorem 4.2, we have already seen that the functionals belong to the dual of $H^\sigma(\Omega;\mathbb{S}^{n\times n})$. Now assume that

$$\sum_{k=1}^N \sum_{1\le i\le j\le n} d_k^{(i,j)}\lambda_k^{(i,j)} = 0$$

on $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ with certain coefficients $d_k^{(i,j)}$. We need to show that all $d_k^{(i,j)} = 0$.

To this end, let $g \in C_0^\infty(\mathbb{R}^d;\mathbb{R})$ be a flat bump function, i.e. a nonnegative, compactly supported function with support $B(0,1)$, satisfying $g(x) = 1$ on $B(0,1/2)$.

Fix $1 \le \ell \le N$, as well as $i^*, j^* \in \{1,\dots,n\}$ with $i^* \le j^*$. Since x_ℓ is no singular point of F there exists a minimal $|\beta| \le m$ such that $c_\beta(x_\ell)$ is invertible. The function

$$g_\ell(x) = \frac{1}{\beta!}(x - x_\ell)^\beta g\left(\frac{x - x_\ell}{q_X}\right),$$

where q_X denotes the separation distance of X, then satisfies $D^\alpha g_\ell(x_k) = 0$ for all $|\alpha| \le m$ and $x_k \ne x_\ell$. Moreover, $D^\alpha g_\ell(x_\ell) = 0$ for $\alpha \ne \beta$ and $D^\beta g_\ell(x_\ell) = 1$. Hence, defining the matrix valued function $G \in H^\sigma(\Omega;\mathbb{S}^{n\times n})$ by $G(x) = g_\ell(x)c_\beta(x_\ell)^{-1}E^s_{i^*j^*}$, we have

$$\begin{aligned}
0 &= \sum_{k=1}^{N} \sum_{1\le i\le j\le n} d_k^{(i,j)} \lambda_k^{(i,j)}(G) \\
&= \sum_{k=1}^{N} \sum_{1\le i\le j\le n} d_k^{(i,j)} e_i^T F(G)(x_k) e_j \\
&= \sum_{k=1}^{N} \sum_{|\alpha|\le m} \sum_{1\le i\le j\le n} d_k^{(i,j)} e_i^T c_\alpha(x_k) c_\beta(x_\ell)^{-1} E^s_{i^*j^*} e_j \; D^\alpha g_\ell(x_k) \\
&= \sum_{1\le i\le j\le n} d_\ell^{(i,j)} e_i^T c_\beta(x_\ell) c_\beta(x_\ell)^{-1} E^s_{i^*j^*} e_j \\
&= c_{i^*,j^*} d_\ell^{(i^*,j^*)},
\end{aligned}$$

where $c_{i^*,j^*} = \frac{1}{\sqrt{2}}$ for $i^* \neq j^*$ and $c_{i^*,i^*} = 1$. Since ℓ, i^*, j^* were chosen arbitrarily, this shows the linear independence. $\square$

Now we consider a special type of F, which will later arise in the application within Dynamical Systems.

THEOREM 4.6. *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with Lipschitz continous boundary. Let $\sigma > d/2+1$ and let $V \in H^{\sigma-1}(\Omega;\mathbb{R}^{n\times n})$ and $f \in H^{\sigma-1}(\Omega;\mathbb{R}^n)$. Define $F : H^\sigma(\Omega;\mathbb{S}^{n\times n}) \to H^{\sigma-1}(\Omega;\mathbb{S}^{n\times n})$ by*

$$F(M)(x) := V(x)^T M(x) + M(x)V(x) + M'(x),$$

where $(M'(x))_{ij} = \nabla M_{ij}(x) \cdot f(x)$.

For each $x_0 \in \Omega$ with $f(x_0) = 0$ (equilibrium point), we assume that all eigenvalues of $V(x_0)$ have negative real part (positive real part).

Finally, let $X = \{x_1, \dots, x_N\} \subseteq \Omega$ be a set of pairwise distinct points and let

$$\lambda_k^{(i,j)}(M) := e_i^T F(M)(x_k) e_j, \qquad 1 \le k \le N, \quad 1 \le i \le j \le n.$$

Then, each $\lambda_k^{(i,j)}$ belongs to the dual of $H^\sigma(\Omega;\mathbb{S}^{n\times n})$ and they are linearly independent. If S denotes the optimal recovery of $M \in H^\sigma(\Omega;\mathbb{S}^{n\times n})$ in the sense of Definition 2.4 using these functionals, then

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

Proof. The operator F is a differential operator of degree 1 as in Definition 4.3 with

$$\begin{aligned}
c_0(x)(M) &= V(x)^T M + MV(x), \\
c_{e_i}(x)(M) &= f_i(x) M.
\end{aligned}$$

We have $x \mapsto c_\alpha(x)[D^\alpha M(x)] \in H^{\sigma-1}(\Omega;\mathbb{S}^{n\times n})$ for every $M \in H^\sigma(\Omega;\mathbb{S}^{n\times n})$. To apply Proposition 4.5, we have to show that there are no singular points in Ω.

Case 1: If $f(x) \neq 0$, then there is an $i^* \in \{1, \dots, n\}$ with $f_{i^*}(x) \neq 0$ and hence $c_{e_{i^*}}(x)$ is invertible with $c_{e_{i^*}}(x)^{-1} = \frac{1}{f_i(x)}\mathrm{id}$.

Case 2: If $f(x) = 0$, then by assumption $V(x)$ $(-V(x))$ has eigenvalues with only negative real part. Then the so-called Lyapunov equation

$$V(x)^T M + MV(x) = C \qquad (-C)$$

has a unique solution for every $C \in \mathbb{S}^{n\times n}$, see e.g. [29, Theorem 4.6], i.e. the operator $c_0(x)$ is injective and, because it maps the finite-dimensional space $\mathbb{S}^{n\times n}$ into itself, also bijective.

The rest follows from the previous results, in particular Theorem 4.2 by setting $\tau = \sigma - 1$. ∎

5. Contraction metric. In this section we will apply the previous general results to the ODE problem of constructing a contraction metric mentioned in the introduction. We seek to show existence, uniqueness and exponential stability of an equilibrium and to study its basin of attraction through a contraction metric.

Contraction analysis can be used to study the distance between trajectories, without reference to an attractor, establishing (exponential) attraction of adjacent trajectories, see [30, 26, 32], see also [22, Section 2.10]; it can be generalised to the study of a Finsler-Lyapunov function [16].

If contraction to a trajectory through x occurs with respect to all adjacent trajectories, then solutions converge to an equilibrium. If the attractor is, e.g., a periodic orbit, then contraction cannot occur in the direction tangential to the trajectories. Hence, contraction analysis for periodic orbits assumes contraction only to occur in a suitable $(n-1)$-dimensional subspace of the tangent space. Contraction metrics for periodic orbits have been studied by Borg [6] with the Euclidean metric and Stenström [42] with a general Riemannian metric. Further results using a contraction metric to establish existence, uniqueness, stability and information about the basin of attraction of a periodic orbit have been obtained in [27, 31].

Only few converse theorems for contraction metrics have been obtained, establishing the existence of a contraction metric, see [20] for some references. Constructive converse theorems, providing algorithms for the explicit construction of a contraction metric, are given in [3] for the global stability of an equilibrium in polynomial systems, using Linear Matrix Inequalities (LMI) and sums of squares (SOS). This method is applicable to polynomial systems which are globally stable, i.e. the basin of attraction is the whole phase space; the maximal degree of the polynomial for the contraction metric has to be fixed beforehand and the method is slow if the degree is large, however, it verifies the definiteness of the contraction metric. In contrast, our method is applicable to general systems and can determine compact subsets of the basin of attraction. The definiteness of the constructed metric is guaranteed by error estimates for sufficiently dense collocation points, but as these estimates involve unknown quantities, we need to verify the definiteness directly in applications.

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 [21]; this is a dynamically different problem, but in comparable problems meshfree collocation is more efficient than semi-definite optimisation.

In [20], the existence of a contraction metric for an equilibrium was shown which satisfies $F(M) = -C$, where C is a given constant, symmetric and positive definite matrix. In [25], summarised in the following theorem, we establish existence and uniqueness of solutions of the more general matrix-valued PDE (5.1).

THEOREM 5.1. *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 $x \in A(x_0)$. Then, for $i = 1, 2$ the matrix equation

$$Df(x)^T M_i(x) + M_i(x) Df(x) + M_i'(x) = -C_i(x) \tag{5.1}$$

has a unique solution $M_i \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$.

Let $K \subseteq A(x_0)$ be a compact set. Then there is a constant c, independent of M_i and C_i such that

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

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

Applied to $M_1 = M$ and $M_2 = S$, the optimal recovery of M, the theorem shows that if $\|F(M)(x) - F(S)(x)\| \le \epsilon$ for all $x \in \overline{\gamma^+(K)}$, then $\|M(x) - S(x)\| \le c\epsilon$ for all $x \in K$. In particular, as M is positive definite in K, so is S, if ϵ is small enough. Note that for a positively invariant and compact set K we have $\overline{\gamma^+(K)} = K$.

Let $f \in C^s(\mathbb{R}^n, \mathbb{R}^n)$, $s \ge 2$. In what follows, we will always have $d = n$. Let x_0 be an exponentially stable equilibrium of $\dot{x} = f(x)$ with basin of attraction $A(x_0)$. Then, our strategy for constructing a Riemannian contraction metric is to choose a symmetric and positive definite matrix $C \in \mathbb{S}^{n\times n}$ and to approximate the partial differential equation

$$F(M)(x) := Df(x)^T M(x) + M(x) Df(x) + M'(x) = -C. \tag{5.2}$$

using generalised collocation as described in the previous sections. This can be summarised as follows. We set $W = \mathbb{S}^{n\times n}$ to be the space of all symmetric $n \times n$ matrices with inner product as in (3.1) and we define $\mathcal{H} = H^\sigma(\Omega; W)$ to be the matrix-valued Sobolev space of Definition 3.1 with reproducing kernel $\Phi : \Omega \times \Omega \to \mathcal{L}(W)$ as in (3.6), where $\Omega \subseteq \mathbb{R}^n$ will be chosen appropriately later on. Since the solution of the matrix equation satisfies $M \in C^{s-1}(A(x_0), \mathbb{S}^{n\times n})$, we set $\sigma = s - 1$. We then define the linear functionals $\lambda_k^{(i,j)} : H^\sigma(\Omega; W) \to \mathbb{R}$ by

$$\begin{aligned} \lambda_k^{(i,j)}(M) &= e_i^T \left[Df(x_k)^T M(x_k) + M(x_k) Df(x_k) + M'(x_k) \right] e_j \\ &=: e_i^T F_k(M) e_j \\ &= e_i^T F(M)(x_k) e_j \end{aligned} \tag{5.3}$$

for $x_k \in \Omega$, $1 \le k \le N$ and $1 \le i \le j \le n$. Here, e_i denotes once again the ith unit vector in $\mathbb{R}^n$.

Then, we can compute the solution S of the optimal recovery problem as in Definition 2.4. This gives the following result.

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 $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ be pairwise distinct points and let $\lambda_k^{(i,j)} \in H^\sigma(\Omega; \mathbb{S}^{n\times n})^*$, $1 \le k \le N$ and $1 \le i \le j \le n$ be defined by (5.3) with $V := Df$ satisfying the conditions of Theorem 4.6. Then there is a unique function $S \in H^\sigma(\Omega; \mathbb{S}^{n\times n})$ solving*

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

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

It has the form

$$S(x) = \sum_{k=1}^{N} \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} \sum_{1\le\mu\le\nu\le n} \lambda_k^{(i,j)}(\Phi(\cdot,x)E^s_{\mu\nu})E^s_{\mu\nu}$$

$$= \sum_{k=1}^{N} \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} \Bigg[\sum_{\mu=1}^{n} F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij} E_{\mu\mu}$$

$$+ \frac{1}{2} \sum_{\substack{\mu,\nu=1\\ \mu\ne\nu}}^{n} [F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} + F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}] E_{\mu\nu} \Bigg], \tag{5.4}$$

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* $1 \le i \le j \le n$, $1 \le \ell \le N$.

If the kernel Φ *is given by* (3.6) *then we also have the alternative expression*

$$S(x) = \sum_{k=1}^{N} \sum_{i,j=1}^{n} \beta_k^{(i,j)} \sum_{\mu,\nu=1}^{n} F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} E_{\mu\nu} \tag{5.5}$$

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* $i \ne j$ *and* $\beta_k^{(i,i)} = \gamma_k^{(i,i)}$.

Proof. The first formula follows from Corollary 2.7 as by (3.5), the Riesz representers are given by

$$v_{\lambda_k^{(i,j)}}(x) = \sum_{1\le\mu\le\nu\le n} \lambda_k^{(i,j)}(\Phi(\cdot,x)E^s_{\mu\nu})E^s_{\mu\nu}.$$

By (3.2) we have

$$\left(\Phi(\cdot,x)E^s_{\mu\nu}\right)_{ij} = \sum_{k,\ell=1}^{n} \Phi(\cdot,x)_{ijk\ell}(E^s_{\mu\nu})_{k\ell}.$$

For $\mu = \nu$ we have

$$\lambda_k^{(i,j)}(\Phi(\cdot,x)E^s_{\mu\mu})E^s_{\mu\mu} = F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij} E_{\mu\mu}.$$

For $\mu < \nu$ we have

$$\lambda_k^{(i,j)}(\Phi(\cdot,x)E^s_{\mu\nu})E^s_{\mu\nu} = \frac{1}{\sqrt{2}} \left(F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} + F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}\right) \frac{1}{\sqrt{2}}(E_{\mu\nu} + E_{\nu\mu})$$

$$= \frac{1}{2} \left(F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} + F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}\right)(E_{\mu\nu} + E_{\nu\mu}).$$

Hence, this yields

$$v_{\lambda_k^{(i,j)}}(x) = \sum_{\mu=1}^{n} F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij} E_{\mu\mu}$$

$$+ \frac{1}{2} \sum_{\substack{\mu,\nu=1\\ \mu\ne\nu}}^{n} [F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} + F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}] E_{\mu\nu},$$

which shows (5.4). To show (5.5), note that by (4.1) we have

$$F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} = F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ji}. \tag{5.6}$$

To show (5.5) it suffices to establish

$$\sum_{i,j=1}^n \beta_k^{(i,j)} \sum_{\mu,\nu=1}^n F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} E_{\mu\nu} = \sum_{\mu=1}^n \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij} E_{\mu\mu}$$
$$+ \sum_{\substack{\mu,\nu=1\\ \mu\neq\nu}}^n \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} \frac{1}{2}\left[F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} + F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}\right] E_{\mu\nu}$$

for $1 \le k \le N$. We compare the expressions on both sides above for each $E_{\mu\nu}$. For $\mu = \nu$ we have to show

$$\sum_{i,j=1}^n \beta_k^{(i,j)} F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij} = \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij}.$$

This is true, since for $i = j$ we have $\gamma_k^{(i,i)} = \beta_k^{(i,i)}$ and for $i \neq j$ we have $F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ij} = F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\mu})_{ji}$ by (5.6) and $\frac{1}{2}\gamma_k^{(i,j)} = \beta_k^{(i,j)} = \beta_k^{(j,i)}$.

For $\mu \neq \nu$ we have to show

$$\sum_{i,j=1}^n \beta_k^{(i,j)} F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij}$$
$$= \frac{1}{2} \sum_{1\le i\le j\le n} \gamma_k^{(i,j)} \left[F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ij} + F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}\right].$$

Again, this is shown using (5.6) since for $i = j$ we have $F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ii} = F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ii}$ and $\gamma_k^{(i,i)} = \beta_k^{(i,i)}$, and for $i \neq j$ we have $F_k(\Phi(\cdot,x)_{\cdot,\cdot,\mu,\nu})_{ji} = F_k(\Phi(\cdot,x)_{\cdot,\cdot,\nu,\mu})_{ij}$ and $\frac{1}{2}\gamma_k^{(i,j)} = \beta_k^{(i,j)} = \beta_k^{(j,i)}$. ∎

The error estimate from Theorem 4.2, or more precisely from Theorem 4.6, gives together with Theorem 5.1 our final result.

THEOREM 5.3. *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 (5.2) from Theorem 5.1. 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 5.2. 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 $h_{X,\Omega}$. The constants c_1, c_2 do not depend on the collocation points X.

In particular, S itself is a contraction metric in K in the sense of Theorem 1.1, provided $h_{X,\Omega}$ is sufficiently small.

Proof. The error estimates and the linear independence of the $\lambda_k^{(i,j)}$ follow immediately from Theorem 4.6 with $V(x) = Df(x) \in H^{s-1}(\Omega;\mathbb{R}^{n\times n})$ as well as Theorem

5.1 with $\sigma = s-1$. To see that S itself defines a contraction metric, we have to verify that S is positive definite and $F(S)$ is negative definite. We will do this only for S as the proof for $F(S)$ is almost identical. The main idea here is that the eigenvalues of symmetric matrix depend continuously on the matrix values. To be more precise, since $M(x)$ is positive definite for every $x \in K$ all its eigenvalues $\lambda_j(x)$, $1 \le j \le n$ are positive. If we order them by size, i.e. $0 < \lambda_1(x) \le \lambda_2(x) \le \dots \lambda_n(x)$, then we have for $x, y \in K$,

$$|\lambda_j(x) - \lambda_j(y)| \le \|M(x) - M(y)\|$$

for any natural matrix norm. Since M is continuous, so is each function λ_j. Since K is compact, there is a $\lambda_{\min}$ such that $\lambda_j(x) \ge \lambda_{\min} > 0$ for all $1 \le j \le n$ and all $x \in K$. If we now sort the eigenvalues $\mu_j(x)$ of $S(x)$ in the same way, similar arguments as above show

$$|\lambda_1(x) - \mu_1(x)| \le \|M(x) - S(x)\| \le c_2 h_{X,\Omega}^{\sigma-1-n/2} \|M\|_{H^\sigma(\Omega;\mathbb{S}^{n\times n})}$$

Hence, if we choose $h_{X,\Omega}$ so small that the term on the right-hand side becomes less than $\lambda_{\min}/2$, we see that $\mu_1(x) \ge \lambda_{\min}/2$ for all $x \in K$, i.e. $S(x)$ is also positive definite for all $x \in K$. □

While this result guarantees that $S(x)$ is eventually positive definite for all $x \in K$, it does not provide us with an a priori estimate on how small $h_{X,\Omega}$ actually has to be since we neither know the constant $c_2 > 0$ nor the norm of the unknown function M. Hence, in applications, we have to verify the positive definiteness directly.

6. Examples.

6.1. Linear example. As a first example we consider the linear system

$$\dot{x} = -x + y, \qquad \dot{y} = x - 2y,$$

which was considered in [21] as a time-periodic example. Note that the solution of the matrix equation (5.2) with $C = I$ is constant and can easily be calculated as

$$M(x) = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \tag{6.1}$$

which allows us to analyse the error to the exact solution. Also note that any set of the form $K_c = [-c, c]^2$ with $c > 0$ is positively invariant. We have used grids of the form $X_\alpha = \{(x, y) \in \mathbb{R}^2 : x, y = -1, \dots, -2\alpha, -\alpha, 0, \alpha, 2\alpha, \dots, 1\}$ with $\alpha = 1, \frac{1}{2}, \frac{1}{2^2}, \dots, \frac{1}{2^5}$.

As kernel we have used Wendland's $C^8(\mathbb{R}^2)$ function

$$\phi(r) = (1 - cr)_+^{10}(2145(cr)^4 + 2250(cr)^3 + 1050(cr)^2 + 250cr + 25),$$

where $x_+ = x$ if $x \ge 0$ and $x_+ = 0$ if $x < 0$. ϕ is a reproducing kernel in $H^\sigma(\mathbb{R}^2)$ with $\sigma = 5.5$, see [47]. We have used $c = 0.9$ to balance the trade-off between good approximation and condition number of the collocation matrix; similar results are achieved for other values of c of the same size.

In each case we have calculated the errors

$$\begin{aligned} e_\alpha &= \max_{x \in X_{check}} \|S^\alpha(x) - M(x)\|_{\max} = \max_{x \in X_{check}} \max_{i,j=1,2} |S^\alpha_{ij}(x) - M_{ij}(x)| \\ e^s_\alpha &= \max_{x \in X_{check}} \|F(S^\alpha)(x) - F(M)(x)\|_{\max}, \end{aligned}$$

with $X_{check} = \{(x, y) \in \mathbb{R}^2 : x, y = -1+\frac{1}{2}\alpha_0, \ldots, -\frac{3}{2}\alpha_0, -\frac{1}{2}\alpha_0, \frac{1}{2}\alpha_0, \frac{3}{2}\alpha_0, \ldots, 1-\frac{1}{2}\alpha_0\}$ with $\alpha_0 = \frac{1}{2^6}$. By Theorem 5.3 we expect the errors to behave like

$$\frac{e_{2\alpha}}{e_\alpha} \approx 2^{\sigma-1-n/2} = 2^{3.5}.$$

Table 6.1 shows the above described errors for different α, the expected ratios and the condition numbers of the collocation matrices. The expected approximation order is well-matched by the observed error $F(S) - F(M)$. In the case of $S - M$, the observed error is signficantly better than predicted.

α	e^s_α	$e^s_{2\alpha}/e^s_\alpha$	e_α	$e_{2\alpha}/e_\alpha$	condition number
1/2	2.5724		1.2334		779.3362
1/4	1.2833	2.0045	0.9169	1.3452	2.6230e+3
1/8	0.3516	3.6499	0.0124	73.9435	2.9894e+5
1/16	0.0329	10.6838	5.6040e-4	22.1271	5.1283e+8
1/32	0.0025	13.1918	1.6311e-5	34.3572	9.8693e+11
$2^{3.5}$		11.3137		11.3137	

Table 6.1

Errors for various computation grids together with the error behaviour and the condition number of the collocation matrix.

Next, we have fixed the grid to consist of the $N = 1681$ points $X = \{(x, y) \in \mathbb{R}^2 : x, y = -4, -3.8, -3.6, \ldots, 0, 0.2, \ldots, 4\}$. As each grid point requires 3 variables of a symmetric 2×2 matrix, we solve a linear system with a 5043×5043 matrix; its condition number is 1.6419e+5. We need to check that the constructed matrix-valued function $S(x)$ is positive definite and $F(S)(x)$ is negative definite, where $F(S) = Df(x)^T S(x) + S(x)Df(x) + S'(x)$. To check that a 2×2 matrix A is positive/negative definite it suffices to check that $\operatorname{tr}(A)$ is positive/negative and $\det(A)$ is positive/$-\det(A)$ is negative. For this example, $\operatorname{tr} S(x)$, $\det S(x)$ are positive in the whole area $[-4, 4]^2$, while Figure 6.1, left, shows small areas near the boundary where $F(S)(x)$ is not negative definite, together with the collocation points. Figure 6.1, right, illustrates the metric $S(x)$ by plotting ellipses $x + v$ around x with $(v - x)^T S(x)(v - x)$ =const, showing that $S(x)$ approximates the constant solution (6.1) well.

6.2. Van der Pol. We consider the van der Pol system with reversed time, which has an exponentially stable equilibrium at the origin. Its basin of attraction is bounded by an unstable periodic orbit. The system is given by

$$\dot{x} = -y, \qquad \dot{y} = x - 3(1 - x^2)y,$$

which was, for example, considered in [19] to compute a Lyapunov function. In our computations, we have used $C = I$ and the grid $X = \{(x, y) \in \mathbb{R}^2 \mid x, y = \ldots, -0.25, -0.125, 0, 0.125, \ldots\} \cap \{x - 1.5 < y < 1.5 + x, -3x - 5.5 < y < -3x + 5.5\}$ with $N = 501$ points, and as each grid point requires 3 variables of a symmetric 2×2 matrix, we have solved a linear system with a 1503×1503 matrix; the condition number is 3.0024e+6. We have used the same kernel as in the previous example.

Figure 6.3, left, shows the collocation points and the basin of attraction of the origin, bounded by an unstable periodic orbit as well as the boundaries of the areas where $F(S)(x)$ is negative definite (red) and $S(x)$ is positive definite (blue).

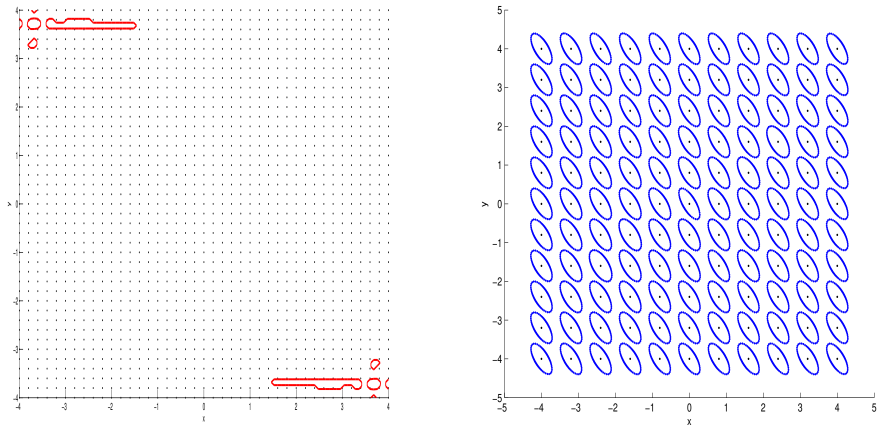

FIG. 6.1. *Left: The collocation points used for the RBF approximation together with the areas where $F(S)(x, y)$ is not negative definite (red). Right: To illustrate the approximation S, around some points x, we have plotted the curve of equal distance with respect to metric $S(x)$, in particular the set $\{x + v \mid (v - x)^T S(x)(v - x) = \text{const}\}$.*

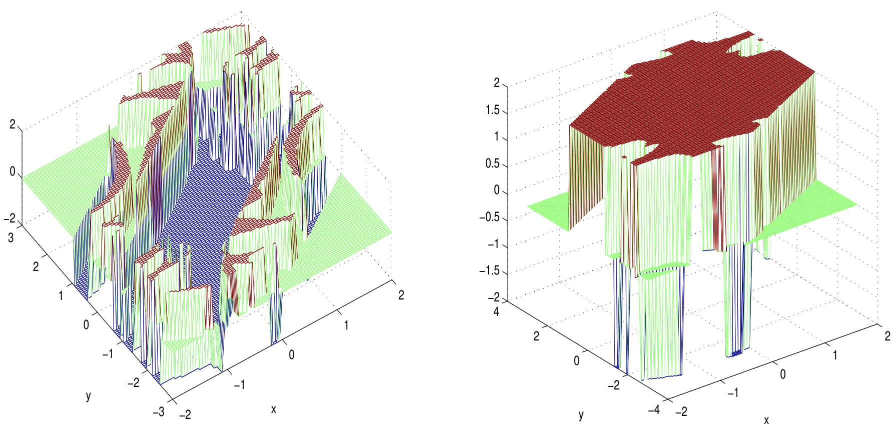

FIG. 6.2. *Left:* $\text{sign}(\text{tr}\, F(S)(x, y)) - \text{sign}(\det F(S)(x, y))$. *If this function is -2, then $F(S)(x, y)$ is negative definite. Right:* $\text{sign}(\text{tr}\, S(x, y)) + \text{sign}(\det S(x, y))$. *If this function is $+2$, then $S(x, y)$ is positive definite.*

In more detail, Figure 6.2 shows $\text{sign}(\text{tr}\, F(S)(x)) - \text{sign}(\det F(S)(x))$, which is -2 in the area where we placed the collocation points (left), as well as $\text{sign}(\text{tr}\, S(x)) + \text{sign}(\det S(x))$ which is $+2$ in the area where we placed the collocation points (right). Figure 6.3, right, shows the point-dependent metric $S(x)$ by plotting ellipses $x + v$ around x with $(v - x)^T S(x)(v - x)$ =const.

REFERENCES

[1] L. AMODEI, *Reproducing kernels of vector-valued function spaces*, in Surface Fitting and Multiresolution Methods, A. L. Méhaut'e, C. Rabut, and L. L. Schumaker, eds., Nashville,

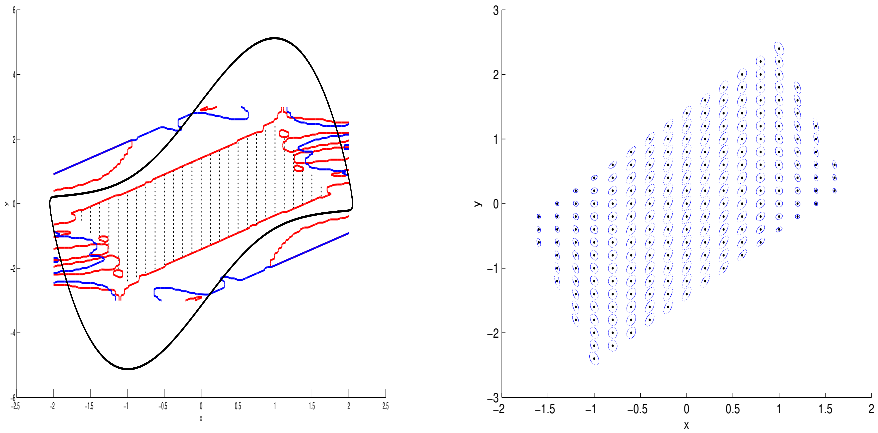

Fig. 6.3. *Left: Collocation points and the boundary of the basin of attraction of the origin, an unstable periodic orbit, together with the boundaries of the areas where* $\operatorname{sign}(\operatorname{tr} F(S)(x,y)) - \operatorname{sign}(\det F(S)(x,y)) = -2$ *(red) and* $\operatorname{sign}(\operatorname{tr} S(x,y)) + \operatorname{sign}(\det S(x,y)) = 2$ *(blue). Right: To illustrate the approximation S, around some points x, we have plotted the set of equal distance with respect to metric $S(x)$, in particular the set* $\{x + v \mid (v-x)^\top S(x)(v-x) = \mathrm{const}\}$.

1997, Vanderbilt University Press, pp. 17–26.

[2] N. Aronszajn, *Theory of reproducing kernels*, Trans. Am. Math. Soc., 68 (1950), pp. 337–404.

[3] E. M. Aylward, P. A. Parrilo, and J.-J. Slotine, *Stability and robustness analysis of nonlinear systems via contraction metrics and SOS programming*, Automatica, 44 (2008), pp. 2163–2170.

[4] M. N. Benbourhim and A. Bouhamidi, *Meshless pseudo-polyharmonic divergence-free and curl-free vector fields approximation*, SIAM J. Math. Anal., 42 (2010), pp. 1218–1245.

[5] A. Berlinet and C. Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Springer, New York, 2004.

[6] G. Borg, *A condition for the existence of orbitally stable solutions of dynamical systems*, vol. 153 of Kungl. Tekn. Högsk. Handl., Elander, 1960.

[7] M. D. Buhmann, *Radial Basis Functions*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2003.

[8] B. Burgeth, S. Didas, L. Florack, and J. Weickert, *A generic approach to diffusion filtering of matrix-fields*, Computing, 81 (2007), pp. 179–197.

[9] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, *Constrained flows of matrix-valued functions: Application to diffusion tensor regularization*, in Computer Vision — ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I, A. Heyden, M. Sparr, Gunnarand Nielsen, and P. Johansen, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2002, pp. 251–265.

[10] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*, Cambridge University Press, Cambridge, 2000.

[11] F. Cucker and S. Smale, *On the mathematical foundation of learning*, Bull. Amer. Math. Soc., 39 (2002), pp. 1–49.

[12] J. Dick, F. Y. Kuo, and I. H. Sloan, *High-dimensional integration: The quasi-Monte Carlo way*, in Acta Numerica, A. Iserles, ed., vol. 22, Cambridge University Press, 2013, pp. 133–288.

[13] G. Fasshauer, *Meshfree Approximation Methods with MATLAB*, World Scientific Publishers, Singapore, 2007.

[14] N. Flyer and B. Fornberg, *Radial basis functions: Developments and applications to planetary scale flows*, Computers and Fluids, 46 (2011), pp. 23–32.

[15] B. Fornberg and N. Flyer, *Solving PDEs with radial basis functions*, in Acta Numerica, A. Iserles, ed., vol. 24, Cambridge University Press, 2015, pp. 215–258.

[16] F. Forni and R. Sepulchre, *A differential Lyapunov framework for contraction analysis*, IEEE Transactions on Automatic Control, 59 (2014), pp. 614–628.

[17] C. Franke and R. Schaback, *Convergence order estimates of meshless collocation methods using radial basis functions*, Adv. Comput. Math., 8 (1998), pp. 381–399.
[18] E. Fuselier, *Sobolev-type approximation rates for divergence-free and curl-free RBF interpolants*, Math. Comput., 77 (2008), pp. 1407–1423.
[19] P. Giesl, *Construction of global Lyapunov functions using radial basis functions*, vol. 1904 of Lecture Notes in Mathematics, Springer-Verlag, Heidelberg, 2007.
[20] ———, *Converse theorems on contraction metrics for an equilibrium*, J. Math. Anal. Appl., 424 (2015), pp. 1380–1403.
[21] P. Giesl and S. Hafstein, *Construction of a CPA contraction metric for periodic orbits using semidefinite optimization*, Nonlinear Anal., 86 (2013), pp. 114–134.
[22] P. Giesl and S. Hafstein, *Computation and verification of Lyapunov functions*, SIAM J. Appl. Dyn. Syst., 14 (2015), pp. 1663–1698.
[23] ———, *Review on computational methods for Lyapunov functions*, Discrete and Continuous Dynamical Systems - Series B (DCDS-B), 20 (2015), pp. 2291 – 2331.
[24] P. Giesl and H. Wendland, *Meshless collocation: Error estimates with application to dynamical systems*, SIAM J. Numer. Anal., 45 (2007), pp. 1723–1741.
[25] ———, *Construction of a contraction metric by meshless collocation*. Preprint Bayreuth/Sussex, to appear in Discrete and Continuous Dynamical Systems Series B, 2016.
[26] W. Hahn, *Stability of Motion*, Springer, Berlin, 1967.
[27] P. Hartman, *Ordinary Differential Equations*, Wiley, New York, 1964.
[28] E. J. Kansa, *Multiquadrics - A scattered data approximation scheme with applications to computational fluid-dynamics I. Surface approximations and partial derivative estimates*, Comput. Math. Appl., 19 (1990), pp. 127–145.
[29] H. Khalil, *Nonlinear systems*, Macmillan Publishing Company, New York, 1992.
[30] N. N. Krasovskii, *Problems of the Theory of Stability of Motion*, Mir, Moscow, 1959.
[31] G. A. Leonov, I. M. Burkin, and A. I.Shepelyavyi, *Frequency Methods in Oscillation Theory*, vol. 357 of Ser. Math. and its Appl., Kluwer, 1996.
[32] W. Lohmiller and J.-J. E. Slotine, *On contraction analysis for non-linear systems*, Automatica J. IFAC, 34 (1998), pp. 683–696.
[33] S. Lowitzsch, *Approximation and Interpolation Employing Divergence-Free Radial Basis Functions with Applications*, PhD thesis, Texas A&M University, College Station, USA, 2002.
[34] S. Mardare, *On systems of first order linear partial differential equations with L^p coefficients*, Adv. Differential Equations, 12 (2007), pp. 301–360.
[35] C. A. Micchelli and M. Pontil, *Learning the kernel function via regularization*, J. of Mach. Learning Research, 6 (2005), pp. 1099–1125.
[36] F. J. Narcowich and J. D. Ward, *Generalized Hermite interpolation via matrix-valued conditionally positive definite functions*, Math. Comput., 63 (1994), pp. 661–687.
[37] F. J. Narcowich, J. D. Ward, and H. Wendland, *Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting*, Math. Comput., 74 (2005), pp. 643–763.
[38] R. Schaback, *The missing Wendland functions*, Adv. Comput. Math., 34 (2011), pp. 67–81.
[39] R. Schaback and H. Wendland, *Kernel techniques: From machine learning to meshless methods*, in Acta Numerica, A. Iserles, ed., vol. 15, Cambridge University Press, 2006, pp. 543–639.
[40] B. Schölkopf and A. J. Smola, *Learning with Kernels – Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, Cambridge, Massachusetts, 2002.
[41] I. Steinwart and A. Christmann, *Support Vector Machines*, Springer, New York, 2008.
[42] B. Stenström, *Dynamical systems with a certain local contraction property*, Math. Scand., 11 (1962), pp. 151–155.
[43] V. Vapnik, *Statistical Learning Theory*, John Wiley and Sons, 1998.
[44] J. Weickert and T. Brox, *Diffusion and regularization of vector- and matrix-valued images*, in Inverse problems, image analysis, and medical imaging (New Orleans, LA, 2001), vol. 313 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2002, pp. 251–268.
[45] H. Wendland, *Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree*, Adv. Comput. Math., 4 (1995), pp. 389–396.
[46] ———, *Meshless Galerkin methods using radial basis functions*, Math. Comput., 68 (1999), pp. 1521–1531.
[47] ———, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, UK, 2005.
[48] ———, *Divergence-free kernel methods for approximating the Stokes problem*, SIAM J. Numer. Anal., 47 (2009), pp. 3158–3179.