

WEAKLY NORMAL BASIS VECTOR FIELDS IN RKHS WITH AN  
APPLICATION TO SHAPE NEWTON METHODS\*ALBERTO PAGANINI<sup>†</sup> AND KEVIN STURM<sup>‡</sup>

**Abstract.** We construct a space of vector fields that are normal to differentiable curves in the plane. Its basis functions are defined via saddle point variational problems in reproducing kernel Hilbert spaces (RKHSs). First, we study the properties of these basis vector fields and show how to approximate them. Then, we employ this basis to discretize shape Newton methods and investigate the impact of this discretization on convergence rates.

**Key words.** shape optimization, shape Newton method, numerical analysis, radial basis functions, reproducing kernel Hilbert spaces

**AMS subject classifications.** 49Q10, 49M15, 93B40, 65Dxx

**DOI.** 10.1137/17M1131623

**1. Introduction.** Shape optimization studies how to design a domain  $\Omega$  such that a shape functional  $J$  is minimized. Shape optimization problems arise naturally in numerous industrial applications. For example, shape optimization can be used to improve the design of electrodes [4], bridges [2], breakwaters [20], optical lenses [23], airplane components [24], and 3D printers [7]. In light of its wide range of applications, it is not a surprise that several commercial products offer shape optimization features.

Shape optimization problems are inherently infinite dimensional. Therefore, any numerical shape optimization algorithm must rely on a discretization of the space of shapes. Additionally, shape optimization problems are, in most instances, nonconvex. Therefore, numerical shape optimization software commonly relies on iterative optimization algorithms. In particular, in light of the large number of optimization parameters, it is advantageous to use optimization strategies that involve derivatives of the functional  $J$ , such as steepest descent or the Newton algorithm.

The derivative of  $J$  with respect to perturbations of  $\Omega$  is called its shape derivative, and its technical definition depends on how these domain perturbations are modeled. Historically, domain perturbations have been defined in terms of velocity flows or perturbations of the identity [9, 18, 30]. Recently, Allaire, Cancès, and Vié [1] introduced domain perturbations in terms of Hamilton–Jacobi equations. These definitions lead to the same formulas for first-order shape derivatives but differ for second-order ones. The equivalence of the three approaches for the first shape derivative stems from the fact that each approach constructs domain perturbations from vector fields, and the Taylor series expansions of these constructions agree to first order. However, the second-order terms of these series do not match. Therefore, the corresponding shape Newton methods potentially deliver different shape updates.

Regardless of the definition used, both first- and second-order shape derivatives have a nontrivial kernel. For instance, shape derivatives are zero in the direction of vector fields that vanish on  $\partial\Omega$ . The Hadamard–Zolésio theorem and its second-order counterpart [22] provide a complete characterization of the kernel of shape derivatives.

---

\*Received by the editors May 24, 2017; accepted for publication (in revised form) September 27, 2018; published electronically January 3, 2019.

<http://www.siam.org/journals/sinum/57-1/M113162.html>

<sup>†</sup>Mathematical Institute, University of Oxford, Oxford OX2 6HG, UK (paganini@maths.ox.ac.uk).

<sup>‡</sup>Johann Radon Institute, Linz 4040, Austria (kevin.sturm@oeaw.ac.at).

This information needs to be taken into account to prove convergence rates of shape Newton methods [27, 31] (because one must “quotient” the kernel from the search space) and has a direct consequence for numerical computations: if the discretization of domain perturbations is not chosen carefully, the matrix that results by restricting the second-order shape derivative to the discrete trial space may fail to be positive definite (it is usually positive semidefinite). In particular, the basis vector fields used to discretize domain perturbations must necessarily have normal component on the boundary of  $\Omega$ .

Our goal is to construct basis vector fields that are normal to a given boundary  $\partial\Omega$  and that have good approximation properties. The literature contains articles that address this issue and rely on vector fields with this geometric feature. In [13, 14, 16], the authors consider star-shaped domains whose parametrization formula is known analytically and construct the basis vector fields using this parametrization. At the same time, this approach is ideal for and limited to shape optimization problems for which all computations can be restricted to the boundary  $\partial\Omega$ , because the basis vector fields constructed in [13, 14, 16] are defined only on  $\partial\Omega$ . To restrict all computations to the boundary, the domain  $\Omega$  must satisfy certain regularity requirements, and if the shape functional  $J$  is constrained to a boundary value problem, its solution must also be sufficiently smooth. If these conditions are not fulfilled, restricting all computations to  $\partial\Omega$  is not possible, and one needs to employ basis vector fields whose support has nonzero measure. In practice, it is not difficult to construct such vector fields by replacing  $\Omega$  with a triangulation and using finite elements [28, 29]. However, the approximation properties of the constructed space are unclear, and it is difficult to know whether the Newton updates are well defined in the zero mesh-width limit.

In our work, we use reproducing kernel Hilbert spaces (RKHSs) to define a class of basis vector fields that are *weakly normal* to the boundary of a domain  $\Omega$  and whose support has nonzero measure. We use the term weakly normal because this geometric property is enforced with the help of saddle point variational problems. This approach allows us to ensure that the resulting vector fields are linearly independent and normal to the shape boundary  $\partial\Omega$ . Therefore, they can be used to formulate discrete shape Newton methods. Moreover, these vector fields span a dense subset of the space of normal vector fields, and it is possible to analyze the error between exact and approximate shape Newton updates.

For simplicity, we restrict the exposition to the two-dimensional case. However, the approach can be adapted to higher dimensions, although the implementation aspects may become more challenging.

## 2. Normal vector fields in RKHS.

**2.1. RKHSs, Cartesian products, and trace spaces.** We begin with some basic definitions of reproducing kernels.

**DEFINITION 2.1.** Let  $\mathcal{X} \subset \mathbf{R}^2$  be an arbitrary set. A function  $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbf{R}$  is called a *reproducing kernel* for the Hilbert space  $(\mathcal{H}(\mathcal{X}), (\cdot, \cdot)_\mathcal{H})$  of functions  $f : \mathcal{X} \rightarrow \mathbf{R}$  if

- (a)  $k(\mathbf{x}, \cdot) \in \mathcal{H}(\mathcal{X})$  for every  $\mathbf{x} \in \mathcal{X}$ ,
- (b)  $(k(\mathbf{x}, \cdot), f)_\mathcal{H} = f(\mathbf{x})$  for every  $\mathbf{x} \in \mathcal{X}$  and every  $f \in \mathcal{H}(\mathcal{X})$ .

**DEFINITION 2.2.** A kernel  $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbf{R}$  is said to be *positive definite* on  $\mathcal{X}$  if, for every finite subset of pairwise distinct points  $\{\mathbf{x}_i\}_{i=1}^n \subset \mathcal{X}$ , the matrix  $(k(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n$  is positive definite.

DEFINITION 2.3. A kernel  $\mathbf{k} : \mathcal{X} \times \mathcal{X} \rightarrow \mathbf{R}$  is said to be symmetric if

$$\mathbf{k}(\mathbf{x}, \mathbf{y}) = \mathbf{k}(\mathbf{y}, \mathbf{x}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

The main reason to consider symmetric positive-definite reproducing kernels is that they can be used to define certain Hilbert spaces.

THEOREM 2.4. To a positive-definite and symmetric kernel  $\mathbf{k} : \mathcal{X} \times \mathcal{X} \rightarrow \mathbf{R}$  defined on an arbitrary set  $\mathcal{X} \subset \mathbf{R}^2$  there corresponds a unique RKHS  $(\mathcal{H}(\mathcal{X}), (\cdot, \cdot)_{\mathcal{H}})$ , which is called the native space of  $\mathbf{k}$ . Note that  $(\cdot, \cdot)_{\mathcal{H}}$  depends on  $\mathbf{k}$ .

There is a variety of symmetric positive-definite reproducing kernels [33] (each tailored to specific applications). In this work, we are particularly interested in Wendland kernels with compact support because, for applications in shape optimization, compact support is useful (it implies sparsity of the discretized shape Hessian) and sometimes necessary (when part of the domain is not “free to move” [21]). In the next example, we list three Wendland kernels and recall their native spaces.

*Example 2.5.* Classical examples of positive-definite kernels can be found in [33]. In this work, we are particularly interested in

$$(1) \quad \mathbf{k}_4^\sigma(x, y) := \left(1 - \frac{|x - y|}{\sigma}\right)_+^4 \left(4 \frac{|x - y|}{\sigma} + 1\right), \quad \sigma > 0,$$

where

$$(\cdot)_+ : \mathbf{R} \rightarrow \mathbf{R}_0^+, \quad x \mapsto (x)_+ := \max(0, x).$$

When  $\sigma = 1$ , (1) is the lowest order Wendland kernel with compact support that is positive-definite on  $\mathbf{R}^d$  (for  $d \leq 3$ ) and is of class  $C^2$ . When  $\mathcal{X} = \mathbf{R}^2$ , its native space is  $H^{2.5}(\mathbf{R}^2)$  [33, Thm. 10.35, p. 160]. When  $\mathcal{X} = \Omega$  is a bounded smooth domain, its native space is  $H^{2.5}(\Omega)$  (because every function  $f \in \mathcal{H}(\Omega)$  can be extended to a function  $\tilde{f} \in \mathcal{H}(\mathbf{R}^2) = H^{2.5}(\mathbf{R}^2)$  [33, Thm. 10.46, p. 169]). Similar properties can be proved for the  $C^4$ -kernel  $\mathbf{k}_6^\sigma$  and the  $C^6$ -kernel  $\mathbf{k}_8^\sigma$ , which are given by [33]

$$(2) \quad \mathbf{k}_6^\sigma(x, y) := \frac{1}{3} \left(1 - \frac{|x - y|}{\sigma}\right)_+^6 \left(35 \frac{|x - y|^2}{\sigma^2} + 18 \frac{|x - y|}{\sigma} + 3\right),$$

$$(3) \quad \mathbf{k}_8^\sigma(x, y) := \left(1 - \frac{|x - y|}{\sigma}\right)_+^8 \left(32 \frac{|x - y|^3}{\sigma^3} + 25 \frac{|x - y|^2}{\sigma^2} + 8 \frac{|x - y|}{\sigma} + 1\right),$$

respectively. Their corresponding native spaces in dimension two are  $H^{3.5}(\mathbf{R}^2)$  and  $H^{4.5}(\mathbf{R}^2)$ , respectively [33, Thm. 10.35, p. 160].

Scalar reproducing kernels can be used to also define Hilbert spaces of vector valued functions [11, Lemma 3.6].

PROPOSITION 2.6. The Cartesian product  $[\mathcal{H}(\mathcal{X})]^2 := \mathcal{H}(\mathcal{X}) \times \mathcal{H}(\mathcal{X})$  of an RKHS  $\mathcal{H}(\mathcal{X})$  (with reproducing kernel  $\mathbf{k}$ ) is itself an RKHS. Its (matrix valued) reproducing kernel is  $\mathbf{k}(\cdot, \cdot)\mathbf{I}$ , where  $\mathbf{I} \in \mathbf{R}^2$  denotes the identity matrix. Henceforth, we denote by  $(\cdot, \cdot)_{[\mathcal{H}]^2}$  the (canonical) inner product of  $[\mathcal{H}(\mathcal{X})]^2$ .

Next, we describe how to construct RKHSs on boundaries of bounded subdomains of  $\mathbf{R}^2$ . Let  $\mathbf{k}$  be a positive-definite kernel on  $\mathbf{R}^2$ , and let  $\mathcal{H}(\mathbf{R}^2)$  be its native space.

Let  $\Omega \subset \mathbf{R}^2$  be a nonempty, bounded, open, and smooth set (e.g.,  $C^1$ ). We denote by  $\boldsymbol{\nu}$  and  $\boldsymbol{\tau}$  the unit normal and unit tangential vector fields along  $\partial\Omega$ .

The restriction of  $k$  onto  $\partial\Omega$  defines a positive-definite kernel  $k|_{\partial\Omega}$ , which uniquely characterizes the RKHS [33, p. 169]

$$(4) \quad \mathcal{H}(\partial\Omega) = \{f|_{\partial\Omega} : f \in \mathcal{H}(\mathbf{R}^2)\}.$$

Similarly, it is possible to define the space of vector fields restricted onto  $\partial\Omega$  by

$$(5) \quad [\mathcal{H}(\partial\Omega)]^2 = \{\mathbf{f}|_{\partial\Omega} : \mathbf{f} \in [\mathcal{H}(\mathbf{R}^2)]^2\}.$$

The next example identifies the space  $\mathcal{H}(\partial\Omega)$  for the kernel (1).

*Example 2.7.* The native space of the restriction onto  $\partial\Omega$  of the Wendland kernel  $k_4^\sigma$  (with  $\sigma$  fixed) is  $H^2(\partial\Omega)$ . This follows from  $\mathcal{H}(\mathbf{R}^2) = H^{2.5}(\mathbf{R}^2)$  and the standard Sobolev trace theorem.

We conclude this section by defining the RKHS of normal and tangential vector fields:

$$(6) \quad [\mathcal{H}(\partial\Omega)]_\boldsymbol{\nu}^2 := \{\mathbf{f} \in [\mathcal{H}(\mathbf{R}^2)]^2 : \mathbf{f} \cdot \boldsymbol{\tau} = 0 \quad \text{on } \partial\Omega\},$$

$$(7) \quad [\mathcal{H}(\partial\Omega)]_\boldsymbol{\tau}^2 := \{\mathbf{f} \in [\mathcal{H}(\mathbf{R}^2)]^2 : \mathbf{f} \cdot \boldsymbol{\nu} = 0 \quad \text{on } \partial\Omega\}.$$

**2.2. Weakly normal basis functions.** In this section, we introduce a novel class of normal vector fields. We begin with the following assumption on the reproducing kernel  $k$ .

**ASSUMPTION 2.8.** Let  $k \geq 1$  be an integer, and let  $\partial\Omega$  be of class  $C^{k+1}$ . We assume that  $k : \mathbf{R}^2 \times \mathbf{R}^2 \rightarrow \mathbf{R}$  is a symmetric positive-definite kernel on  $\mathbf{R}^2$  with the property

$$(8) \quad k(\mathbf{x}, \mathbf{x}) = 1 \quad \text{for all } \mathbf{x} \in \partial\Omega.$$

We assume that for every  $\mathbf{x} \in \mathbf{R}^2$  we have  $k(\mathbf{x}, \cdot) \in C^k(\mathbf{R}^2)$ . Furthermore, we assume that there is a constant  $c_\Omega > 0$ , such that for every  $\phi \in C^k(\partial\Omega)$  and every  $f \in \mathcal{H}(\partial\Omega)$ ,

$$(9) \quad \phi f \in \mathcal{H}(\partial\Omega) \quad \text{and} \quad \|\phi f\|_{\mathcal{H}} \leq c_\Omega \|\phi\|_{C^k(\partial\Omega)} \|f\|_{\mathcal{H}}.$$

Finally, we assume that the constant function  $1 : \partial\Omega \rightarrow \mathbf{R}$  belongs to  $\mathcal{H}(\partial\Omega)$  and, for simplicity, that the support of  $k$  is connected.

*Example 2.9.* The kernels  $k_4^\sigma$ ,  $k_6^\sigma$ , and  $k_8^\sigma$  satisfy Assumption 2.8 for  $k = 2, 3, 4$ , respectively.

*Remark 2.10.* Property (9) implies that  $C^k(\partial\Omega) \subset \mathcal{H}(\partial\Omega)$ , and that  $[\mathcal{H}(\partial\Omega)]^2 = [\mathcal{H}(\partial\Omega)]_\boldsymbol{\nu}^2 \oplus [\mathcal{H}(\partial\Omega)]_\boldsymbol{\tau}^2$  (because every vector field  $\mathbf{v} \in [\mathcal{H}(\partial\Omega)]^2$  can be decomposed into  $\mathbf{v} = (\mathbf{v} \cdot \boldsymbol{\nu})\boldsymbol{\nu} + (\mathbf{v} \cdot \boldsymbol{\tau})\boldsymbol{\tau}$  on  $\partial\Omega$ ).

**DEFINITION 2.11.** For a point  $\mathbf{x} \in \partial\Omega$ , let  $(\mathbf{r}_\mathbf{x}, p_\mathbf{x}) \in [\mathcal{H}(\partial\Omega)]^2 \times \mathcal{H}(\partial\Omega)$  be the solution of

$$(10a) \quad (\mathbf{r}_\mathbf{x}, \boldsymbol{\varphi})_{[\mathcal{H}]^2} + (\boldsymbol{\varphi} \cdot \boldsymbol{\tau}, p_\mathbf{x})_{\mathcal{H}} = \boldsymbol{\nu}(\mathbf{x}) \cdot \boldsymbol{\varphi}(\mathbf{x}) \quad \text{for all } \boldsymbol{\varphi} \in [\mathcal{H}(\partial\Omega)]^2,$$

$$(10b) \quad (\mathbf{r}_\mathbf{x} \cdot \boldsymbol{\tau}, \psi)_{\mathcal{H}} = 0 \quad \text{for all } \psi \in \mathcal{H}(\partial\Omega).$$

The function  $\mathbf{r}_\mathbf{x}$  is the weakly normal basis function associated with  $\mathbf{x}$ .

To show that Definition 2.11 makes sense, we need to prove that (10) admits a unique and stable solution. This is done in Lemma 2.13, which in turn relies on Lemma 2.12. Before showing Lemma 2.12, we introduce the following notation: henceforth, we denote by  $\mathbf{e}_i, i = 1, 2$ , the canonical basis of  $\mathbf{R}^2$  and use the norm

$$(11) \quad \|\mathbf{f}\|_{C^k} := \sqrt{\|\mathbf{f} \cdot \mathbf{e}_1\|_{C^k}^2 + \|\mathbf{f} \cdot \mathbf{e}_2\|_{C^k}^2} \quad \text{for all } \mathbf{f} \in C^k(\mathbf{R}^2, \mathbf{R}^2).$$

LEMMA 2.12. *The operator  $B : [\mathcal{H}(\partial\Omega)]^2 \rightarrow \mathcal{H}(\partial\Omega)^*$  defined by  $\varphi \mapsto (\varphi \cdot \boldsymbol{\tau}, \cdot)_\mathcal{H}$  is surjective. Moreover,*

$$(12) \quad \|B^*\psi\|_{([\mathcal{H}]^2)^*} \geq \frac{\|\psi\|_\mathcal{H}}{c_\Omega \|\boldsymbol{\tau}\|_{C^k}} \quad \text{for all } \psi \in \mathcal{H}(\partial\Omega).$$

*Proof.* To show that  $B$  is surjective, let  $f \in \mathcal{H}(\partial\Omega)^*$ . By the Riesz representation theorem, there is a function  $\eta_f \in \mathcal{H}(\partial\Omega)$  that satisfies  $f(\psi) = (\eta_f, \psi)_\mathcal{H}$  for all  $\psi \in \mathcal{H}(\partial\Omega)$ . Since the function  $\varphi := \eta_f \boldsymbol{\tau} \in [\mathcal{H}(\partial\Omega)]^2$  satisfies  $B\varphi = f$ , the operator  $B$  is surjective. To verify (12), first note that inequality (9) implies  $\|\boldsymbol{\tau}\psi\|_{[\mathcal{H}]^2} \leq c_\Omega \|\boldsymbol{\tau}\|_{C^k} \|\psi\|_\mathcal{H}$  for every  $\psi \in \mathcal{H}(\partial\Omega)$ . Thus,

$$\|B^*\psi\|_{([\mathcal{H}]^2)^*} = \sup_{\substack{\varphi \in [\mathcal{H}(\partial\Omega)]^2 \\ \|\varphi\|_{[\mathcal{H}]^2} = 1}} |(\boldsymbol{\tau} \cdot \varphi, \psi)_\mathcal{H}| \geq \left( \frac{\boldsymbol{\tau} \cdot (\boldsymbol{\tau}\psi)}{\|\boldsymbol{\tau}\psi\|_{[\mathcal{H}]^2}}, \psi \right)_\mathcal{H} \geq \frac{\|\psi\|_\mathcal{H}}{c_\Omega \|\boldsymbol{\tau}\|_{C^k}}. \quad \square$$

LEMMA 2.13. *Let  $\mathbf{x} \in \partial\Omega$ . The saddle point problem (10) admits a unique solution, which satisfies*

$$(13) \quad \|\mathbf{r}_x\|_{[\mathcal{H}]^2} \leq \frac{1}{\|\boldsymbol{\nu}\|_{[\mathcal{H}]^2}} \quad \text{and} \quad \|p_x\|_\mathcal{H} \leq \frac{2c_\Omega \|\boldsymbol{\tau}\|_{C^k}}{\|\boldsymbol{\nu}\|_{[\mathcal{H}]^2}}.$$

*Proof.* Since  $B$  is surjective and  $(\cdot, \cdot)_{[\mathcal{H}]^2}$  is an inner product, existence and uniqueness follow from the classical result [5, Thm. 4.2.1, p. 224]. The continuity estimates (13) follow from [5, Thm. 4.2.3, p. 228], (12), and the estimate

$$(14) \quad \sup_{\substack{\varphi \in [\mathcal{H}]^2, \\ \|\varphi\|_{[\mathcal{H}]^2} = 1}} |\boldsymbol{\nu}(\mathbf{x}) \cdot \varphi(\mathbf{x})| \leq \frac{|\boldsymbol{\nu}(\mathbf{x}) \cdot \boldsymbol{\nu}(\mathbf{x})|}{\|\boldsymbol{\nu}\|_{[\mathcal{H}]^2}} \leq 1/\|\boldsymbol{\nu}\|_{[\mathcal{H}]^2}.$$

Note that  $1 \in \mathcal{H}(\partial\Omega)$  by Assumption 2.8, and thus  $\boldsymbol{\nu} \in [\mathcal{H}(\partial\Omega)]^2$ .  $\square$

The following remark provides an alternative definition of the function  $\mathbf{r}_x$ .

Remark 2.14. Let  $\mathbf{x} \in \partial\Omega$ . The solution  $\mathbf{r}_x \in [\mathcal{H}(\partial\Omega)]^2$  of (10) is the unique minimizer of

$$(15) \quad \min_{\substack{\varphi \in [\mathcal{H}(\partial\Omega)]^2, \\ \varphi \cdot \boldsymbol{\tau} = 0 \text{ on } \partial\Omega}} \frac{1}{2} \|\varphi\|_{[\mathcal{H}]^2}^2 - \boldsymbol{\nu}(\mathbf{x}) \cdot \varphi(\mathbf{x}).$$

*Proof.* By differentiating the Lagrangian functional

$$(16) \quad \mathcal{L} : [\mathcal{H}(\partial\Omega)]^2 \times \mathcal{H}(\partial\Omega) \rightarrow \mathbf{R}, \quad (\varphi, \psi) \mapsto \frac{1}{2} \|\varphi\|_{[\mathcal{H}]^2}^2 - \boldsymbol{\nu}(\mathbf{x}) \cdot \varphi(\mathbf{x}) + (\varphi \cdot \boldsymbol{\tau}, \psi)_\mathcal{H},$$

it is easy to see that equations (10) are the first-order optimality conditions of (15). This is the so-called Lagrange multiplier rule. Since  $\varphi \mapsto 1/2 \|\varphi\|_{[\mathcal{H}]^2}^2 - \boldsymbol{\nu}(\mathbf{x}) \cdot \varphi(\mathbf{x})$  is a convex and coercive functional,  $\mathbf{r}_x$  is a minimizer of (15), where  $(\mathbf{r}_x, p_x)$  is a solution of the saddle point equation (10).  $\square$

In the next theorem, we analyze in detail the properties of  $\mathbf{r}_x$  and show that, in general,  $\mathbf{r}_x$  is neither  $\nu(\cdot)k(x, \cdot)$  nor  $\nu(\cdot)$ . To facilitate the interpretation, in Figure 1 we include some plots of the function  $\mathbf{r}_x$ .

**THEOREM 2.15.** *Let  $x \in \partial\Omega$ . The solution  $(\mathbf{r}_x, p_x)$  of (10) has the following properties:*

- (a)  $\mathbf{r}_x \in [\mathcal{H}(\partial\Omega)]_\nu^2$ .
- (b)  $\|\mathbf{r}_x\|_{[\mathcal{H}]^2} \leq 1$ .
- (c)  $\|\mathbf{r}_x\|_{[\mathcal{H}]^2} > 0$ , and the surface measure of  $\text{supp}(\mathbf{r}_x)$  is strictly greater than zero.
- (d) The equality  $\|\mathbf{r}_x\|_{[\mathcal{H}]^2} = 1$  holds if and only if  $\partial\Omega \cap \text{supp}(\mathbf{r}_x)$  is a straight segment. In particular,  $(\mathbf{r}_x, p_x) = (k(x, \cdot)\nu(x), 0)$ .
- (e)  $(\mathbf{r}_x, \nu)_{[\mathcal{H}]^2} = 1$  and  $\|\nu\|_{[\mathcal{H}]^2}\|\mathbf{r}_x\|_{[\mathcal{H}]^2} = 1$ . In particular, if  $\|\nu\|_{[\mathcal{H}]^2} = 1$ , then  $\partial\Omega$  is a straight segment.
- (f) If  $\partial\Omega \cap \text{supp}(\mathbf{r}_x)$  is not a straight segment, then

$$(17) \quad \mathbf{r}_x \neq \nu(\cdot)k(x, \cdot) \quad \text{and} \quad \mathbf{r}_x \neq \nu.$$

- (g) The basis functions  $\mathbf{r}_x, \mathbf{r}_y$  associated with the points  $x, y \in \partial\Omega$  satisfy

$$(18) \quad \mathbf{r}_x(y) \cdot \nu(y) = \mathbf{r}_y(x) \cdot \nu(x).$$

Moreover, for each  $x \in \partial\Omega$  there is a neighborhood  $U_x$  of  $x$  such that  $\mathbf{r}_x(y) \cdot \nu(y) > 0$  for all  $y \in U_x$ .

*Proof.* Item (a) follows from (10b) because  $\mathbf{r}_x(y) \cdot \tau(y) = (\mathbf{r}_x \cdot \tau, k(y, \cdot))_{\mathcal{H}} = 0$  for every  $y \in \partial\Omega$ .

To show item (b), note that choosing  $\varphi = \mathbf{r}_x$  in (10a) gives

$$(19) \quad \|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2 = \mathbf{r}_x(x) \cdot \nu(x) \leq |\mathbf{r}_x(x)| = \sqrt{|\mathbf{r}_x(x) \cdot \mathbf{e}_1|^2 + |\mathbf{r}_x(x) \cdot \mathbf{e}_2|^2}.$$

The reproducing kernel property of  $k$  and (8) imply

$$(20) \quad \|k(x, \cdot)\|_{\mathcal{H}}^2 = (k(x, \cdot), k(x, \cdot))_{\mathcal{H}} = k(x, x) = 1.$$

Therefore, by the reproducing kernel property of  $k$  and the Cauchy–Schwarz inequality,

$$(21) \quad |\mathbf{r}_x(x) \cdot \mathbf{e}_i| = |(k(x, \cdot), \mathbf{r}_x \cdot \mathbf{e}_i)_{\mathcal{H}}| \leq \|k(x, \cdot)\|_{\mathcal{H}} \|\mathbf{r}_x \cdot \mathbf{e}_i\|_{\mathcal{H}} = \|\mathbf{r}_x \cdot \mathbf{e}_i\|_{\mathcal{H}},$$

and thus

$$|\mathbf{r}_x(x)|^2 = |\mathbf{r}_x(x) \cdot \mathbf{e}_1|^2 + |\mathbf{r}_x(x) \cdot \mathbf{e}_2|^2 \leq \|\mathbf{r}_x \cdot \mathbf{e}_1\|_{\mathcal{H}}^2 + \|\mathbf{r}_x \cdot \mathbf{e}_2\|_{\mathcal{H}}^2 = \|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2.$$

In light of (19), this yields

$$(22) \quad \|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2 \leq |\mathbf{r}_x(x)| \leq \|\mathbf{r}_x\|_{[\mathcal{H}]^2},$$

which, in turn, implies  $\|\mathbf{r}_x\|_{[\mathcal{H}]^2} \leq 1$ .

To show item (c), suppose that  $\mathbf{r}_x = 0$ . Then,  $p_x = 0$  because (10a) implies that for any  $y \in \partial\Omega$ ,

$$\begin{aligned} p_x(y) &= (k(y, \cdot), p_x)_{\mathcal{H}} = (k(y, \cdot)\tau(\cdot) \cdot \tau(\cdot), p_x)_{\mathcal{H}} \\ &= \nu(x) \cdot k(y, x)\tau(x) - (\mathbf{r}_x, k(y, \cdot)\tau(\cdot))_{[\mathcal{H}]^2} = 0. \end{aligned}$$

However,  $(\mathbf{r}_x, p_x) = 0$  does not satisfy (10a) when the test function  $\varphi$  is  $\varphi = k(x, \cdot)\nu(\cdot) \in [\mathcal{H}(\partial\Omega)]^2$ . Finally,  $\mathbf{r}_x \neq \mathbf{0}$  implies that there is at least a point  $\mathbf{y} \in \partial\Omega$  such that  $\mathbf{r}_x(\mathbf{y}) \neq 0$ . Since  $\mathbf{r}_x$  is continuous on  $\partial\Omega$ ,  $\text{supp}(\mathbf{r}_x)$  contains a nonempty neighborhood  $U_y$  of  $\mathbf{y}$ .

To show item (d), let us first assume that  $\|\mathbf{r}_x\|_{[\mathcal{H}]^2} = 1$ . Then, (22) implies  $|\mathbf{r}_x(\mathbf{x})| = \|\mathbf{r}_x\|_{[\mathcal{H}]^2}$ , and since

$$|\mathbf{r}_x(\mathbf{x}) \cdot \mathbf{e}_1|^2 + |\mathbf{r}_x(\mathbf{x}) \cdot \mathbf{e}_2|^2 = |\mathbf{r}_x(\mathbf{x})|^2 = \|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2 = \|\mathbf{r}_x \cdot \mathbf{e}_1\|_{\mathcal{H}}^2 + \|\mathbf{r}_x \cdot \mathbf{e}_2\|_{\mathcal{H}}^2,$$

(21) becomes an equality (because  $|\mathbf{r}_x(\mathbf{x}) \cdot \mathbf{e}_i| \leq \|\mathbf{r}_x \cdot \mathbf{e}_i\|_{\mathcal{H}}$ ). In particular, this implies that

$$|(k(x, \cdot), \mathbf{r}_x \cdot \mathbf{e}_i)_{\mathcal{H}}| = \|k(x, \cdot)\|_{\mathcal{H}} \|\mathbf{r}_x \cdot \mathbf{e}_i\|_{\mathcal{H}}.$$

For nonzero vectors, the Cauchy–Schwarz inequality becomes an equality if and only if the two vectors are linearly dependent. Since  $k(x, \cdot)$  and  $\mathbf{r}_x$  are both nonzero, there is a nonzero vector  $\alpha \in \mathbf{R}^2$  such that  $\mathbf{r}_x(\cdot) = k(x, \cdot)\alpha$  (in fact,  $|\alpha| = 1$  because  $|\mathbf{r}_x(\mathbf{x})| = 1 = k(x, \mathbf{x})$ ). This, combined with  $\mathbf{r}_x \in [\mathcal{H}(\partial\Omega)]_{\nu}^2$  (see item (a)), implies that  $\nu$  restricted to  $\partial\Omega \cap \text{supp}(k(x, \cdot))$  is either  $\alpha$  or  $-\alpha$  and thus constant.

On the other hand, note that

$$(k(x, \cdot)\nu(x), \varphi)_{[\mathcal{H}]^2} = \nu(x) \cdot \mathbf{e}_1(k(x, \cdot), \varphi \cdot \mathbf{e}_1)_{\mathcal{H}} + \nu(x) \cdot \mathbf{e}_2(k(x, \cdot), \varphi \cdot \mathbf{e}_2)_{\mathcal{H}} = \nu(x) \cdot \varphi(x),$$

which implies that  $(\mathbf{r}_x, p_x) = (k(x, \cdot)\nu(x), 0)$  satisfies (10a). If  $\partial\Omega \cap \text{supp}(\mathbf{r}_x)$  is a straight segment, then  $\nu(x) \cdot \tau(y) = 0$  for all  $y \in \partial\Omega \cap \text{supp}(k(x, \cdot))$ . Since  $k(x, y) = 0$  for all  $y \notin \text{supp}(k(x, \cdot))$ ,

$$(k(x, \cdot)\nu(x) \cdot \tau(\cdot), k(y, \cdot))_{\mathcal{H}} = k(x, y)\nu(x) \cdot \tau(y) = 0 \quad \text{for every } y \in \partial\Omega,$$

and (10b) is also satisfied. Finally,  $\|k(x, \cdot)\nu(x)\|_{[\mathcal{H}]^2}^2 = |\nu(x)|^2 \|k(x, \cdot)\|_{\mathcal{H}}^2 = 1$  by (20).

To show item (e), recall that  $1 \in [\mathcal{H}(\partial\Omega)]^2$  by Assumption 2.8. Therefore,  $\nu \in [\mathcal{H}(\partial\Omega)]^2$ . Replacing  $\varphi$  with  $\nu$  in (10a) shows that  $(\mathbf{r}_x, \nu)_{[\mathcal{H}]^2} = 1$ . So, by the Cauchy–Schwarz inequality and (13),

$$(23) \quad 1 = (\mathbf{r}_x, \nu)_{[\mathcal{H}]^2} \leq \|\nu\|_{[\mathcal{H}]^2} \|\mathbf{r}_x\|_{[\mathcal{H}]^2} \leq 1.$$

Moreover, if  $\|\nu\|_{[\mathcal{H}]^2} = 1$ , then  $\|\mathbf{r}_x\|_{[\mathcal{H}]^2} = 1$  for every  $x \in \partial\Omega$ . By item (d), it follows that  $\partial\Omega \cap \text{supp}(\mathbf{r}_x)$  is a straight segment. Since this is the case for every  $x \in \partial\Omega$ , it follows that  $\partial\Omega$  itself is a straight segment.

To show item (f), let  $y \in \partial\Omega$ . Plugging  $\nu(\cdot)k(y, \cdot)$  into (10a) shows that

$$(24) \quad (\mathbf{r}_x, \nu(\cdot)k(y, \cdot))_{[\mathcal{H}]^2} = k(x, y).$$

In particular,  $(\mathbf{r}_x, \nu(\cdot)k(x, \cdot))_{[\mathcal{H}]^2} = k(x, x) = 1$ . Therefore,

$$(25) \quad 0 \leq \|\mathbf{r}_x - \nu(\cdot)k(x, \cdot)\|_{[\mathcal{H}]^2}^2 = \|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2 + \|\nu(\cdot)k(x, \cdot)\|_{[\mathcal{H}]^2}^2 - 2$$

$$(26) \quad = (\|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2 - 1) + (\|\nu(\cdot)k(x, \cdot)\|_{[\mathcal{H}]^2}^2 - 1).$$

Since item (d) implies that  $\|\mathbf{r}_x\|_{[\mathcal{H}]^2} < 1$ , we conclude that  $\|\nu(\cdot)k(x, \cdot)\|_{[\mathcal{H}]^2} > 1$  and  $\mathbf{r}_x \neq \nu(\cdot)k(x, \cdot)$ . Finally, item (e) states that  $(\mathbf{r}_x, \nu)_{[\mathcal{H}]^2} = 1$ . Therefore,

$$(27) \quad 0 \leq \|\mathbf{r}_x - \nu\|_{[\mathcal{H}]^2}^2 = (\|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2 - 1) + (\|\nu\|_{[\mathcal{H}]^2}^2 - 1),$$

and  $\mathbf{r}_x \neq \boldsymbol{\nu}$  because  $\|\boldsymbol{\nu}\|_{[\mathcal{H}]^2} > 1$ .

To show item (g), we note that  $(\mathbf{r}_x, \mathbf{r}_y)_{[\mathcal{H}]^2} = (\mathbf{r}_y, \mathbf{r}_x)_{[\mathcal{H}]^2}$ . Therefore, by (10a),

$$(28) \quad \boldsymbol{\nu}(x) \cdot \mathbf{r}_y(x) = (\mathbf{r}_x, \mathbf{r}_y)_{[\mathcal{H}]^2} = (\mathbf{r}_y, \mathbf{r}_x)_{[\mathcal{H}]^2} = \boldsymbol{\nu}(y) \cdot \mathbf{r}_x(y).$$

For  $x = y$ , this implies  $\mathbf{r}_x(x) \cdot \boldsymbol{\nu}(x) = \|\mathbf{r}_x\|_{[\mathcal{H}]^2}^2 > 0$ . Since  $y \mapsto \mathbf{r}_x(y) \cdot \boldsymbol{\nu}(y)$  is continuous, there is a neighborhood  $U_x$  of  $x$  such that  $\mathbf{r}_x(y) \cdot \boldsymbol{\nu}(y) > 0$  for all  $y \in U_x$ .  $\square$

*Remark 2.16.* The vector field  $\mathbf{r}_x$  depends on the metric induced by the kernel  $k$ . For instance, varying the parameters of the Wendland kernel (1) results in different  $\mathbf{r}_x$ s; see Figure 1. In particular, we observe that the decay of  $\mathbf{r}_x$  away from  $x$  is influenced by the support of the kernel  $k_p^\sigma$ .

*Remark 2.17.* To extend Definition 2.11 to the three-dimensional (3D) case, we note that any vector field  $\boldsymbol{\varphi} \in [\mathcal{H}(\partial\Omega)]^3$  can be rewritten as  $\boldsymbol{\varphi} = (\boldsymbol{\varphi} \cdot \boldsymbol{\nu})\boldsymbol{\nu} + \mathbf{P}_\tau(\boldsymbol{\varphi})$ , where  $\mathbf{P}_\tau(\boldsymbol{\varphi}) := \boldsymbol{\varphi} - (\boldsymbol{\varphi} \cdot \boldsymbol{\nu})\boldsymbol{\nu}$ . The operator  $\mathbf{P}_\tau : [\mathcal{H}(\partial\Omega)]^3 \rightarrow [\mathcal{H}(\partial\Omega)]^3$  is a linear continuous projection with closed range  $[\mathcal{H}(\partial\Omega)]_\tau^3$ . Therefore, to construct weakly normal vector fields in three dimensions, one can replace (10) with the saddle point problem

$$(29a) \quad (\mathbf{r}_x, \boldsymbol{\varphi})_{[\mathcal{H}]^3} + (\mathbf{P}_\tau(\boldsymbol{\varphi}), \mathbf{p}_x)_{[\mathcal{H}]^3} = \boldsymbol{\nu}(x) \cdot \boldsymbol{\varphi}(x) \quad \text{for all } \boldsymbol{\varphi} \in [\mathcal{H}(\partial\Omega)]^3,$$

$$(29b) \quad (\mathbf{P}_\tau(\mathbf{r}_x), \boldsymbol{\psi})_{[\mathcal{H}]^3} = 0 \quad \text{for all } \boldsymbol{\psi} \in [\mathcal{H}(\partial\Omega)]^3.$$

Similarly to Lemma 2.12, we consider the operator  $B : [\mathcal{H}(\partial\Omega)]^3 \rightarrow ([\mathcal{H}(\partial\Omega)]^3)^*$  defined by  $\boldsymbol{\varphi} \mapsto (\mathbf{P}_\tau(\boldsymbol{\varphi}), \cdot)_{[\mathcal{H}]^3}$ . In contrast to the 2D case, here this operator  $B$  is not surjective. However, it has closed range. Therefore, following [5, sect. 4.2.4, p. 230], it is straightforward to show that (29a) admits a solution. The function  $\mathbf{r}_x$  (which is our main target) is uniquely determined and satisfies a stability condition similar to (13), whereas the multiplier  $\mathbf{p}_x$  is uniquely defined only as an element of the quotient space  $[\mathcal{H}(\partial\Omega)]^3 / \ker B^*$ . Finally, it is possible to derive a counterpart of Theorem 2.15 by repeating its proof and replacing products of the form  $\mathbf{f} \cdot \boldsymbol{\tau}$  with  $\mathbf{P}_\tau(\mathbf{f})$  in the proof and in the definition (6).

**2.3. Density of weakly normal functions.** In this section we study the approximation properties of the functions  $\mathbf{r}_x$ s with respect to  $[\mathcal{H}(\partial\Omega)]_\boldsymbol{\nu}^2$ .

Let  $N \geq 1$  be an integer, and let  $\mathcal{X}_N := \{x_1, \dots, x_N\} \subset \partial\Omega$  be a collection of pairwise distinct points.

**LEMMA 2.18.** *The functions  $\mathbf{r}_{x_1}, \dots, \mathbf{r}_{x_N}$  associated with the points in  $\mathcal{X}_N$  are linearly independent.*

*Proof.* Let  $\gamma_1, \dots, \gamma_N \in \mathbf{R}$  be such that  $\mathbf{r} := \sum_{i=1}^N \gamma_i \mathbf{r}_{x_i} = 0$ . Setting  $p := \sum_{i=1}^N \gamma_i p_{x_i}$  and summing (10a) over  $i = 1, \dots, N$  yields

$$(30) \quad (\boldsymbol{\varphi} \cdot \boldsymbol{\tau}, p)_\mathcal{H} = \sum_{i=1}^N \gamma_i \boldsymbol{\nu}(x_i) \cdot \boldsymbol{\varphi}(x_i) = \left( \sum_{i=1}^N \gamma_i k(x_i, \cdot), \boldsymbol{\nu} \cdot \boldsymbol{\varphi} \right)_\mathcal{H} \quad \text{for all } \boldsymbol{\varphi} \in [\mathcal{H}(\partial\Omega)]^2.$$

Choosing  $\boldsymbol{\varphi} = (\sum_{i=1}^N \gamma_i k(x_i, \cdot))\boldsymbol{\nu} \in [\mathcal{H}(\partial\Omega)]_\boldsymbol{\nu}^2$  (cf. Assumption 2.8) yields

$$(31) \quad \left\| \sum_{i=1}^N \gamma_i k(x_i, \cdot) \right\|_\mathcal{H}^2 = 0 \quad \text{and thus} \quad \sum_{i=1}^N \gamma_i k(x_i, y) = 0 \quad \text{for all } y \in \mathbf{R}^2.$$

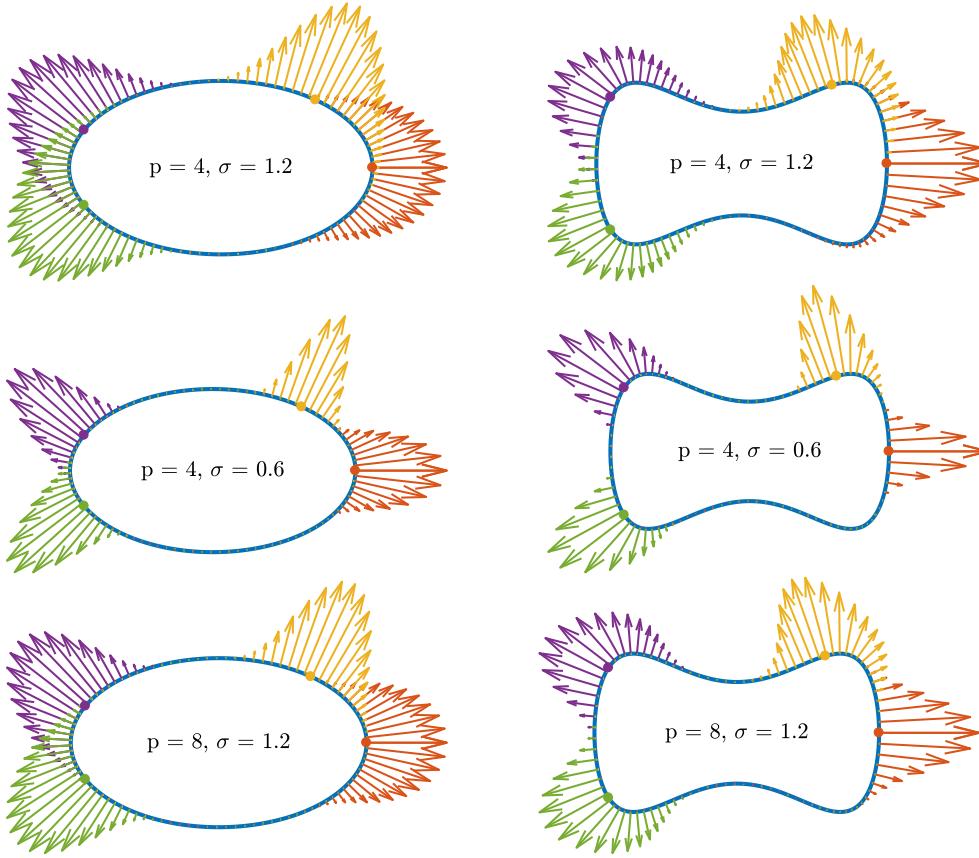


FIG. 1. Some weakly normal basis functions based on compactly supported Wendland kernels (1) and (3) when  $\partial\Omega$  is the ellipse  $\{\gamma(\phi) = (1.4 \cos(\phi), 0.8 \sin(\phi)) : \phi \in [0, 2\pi]\}$  (left) and the nonconvex domain  $\{\gamma(\phi) = (1.4 \cos(\phi), 0.8 \sin(\phi) + 0.3 \sin(3\phi)) : \phi \in [0, 2\pi]\}$  (right). The dots denote the points to which the weakly normal basis functions are associated. We observe that different choices of kernel parameters result in different weakly normal basis functions.

In particular, this implies that

$$(32) \quad \sum_{i=1}^N \gamma_i k(\mathbf{x}_i, \mathbf{x}_j) = 0 \quad \text{for all } \mathbf{x}_j \in \mathcal{X}_N,$$

and thus  $\gamma_1 = \dots = \gamma_N = 0$ , because  $k$  is positive-definite on  $\mathbf{R}^2$ . Therefore, the functions  $\mathbf{r}_{\mathbf{x}_1}, \dots, \mathbf{r}_{\mathbf{x}_N}$  are linearly independent.  $\square$

Lemma 2.18 allows the definition of the interpolation operator onto  $\mathcal{R}_N := \text{span}\{\mathbf{r}_{\mathbf{x}} : \mathbf{x} \in \mathcal{X}_N\}$ . The following lemma and remark clarify the approximation properties of  $\mathcal{R}_N$  on  $[\mathcal{H}(\partial\Omega)]^2$ .

LEMMA 2.19. *The interpolation operator  $[\mathcal{H}(\partial\Omega)]^2 \ni \mathbf{r} \mapsto \mathbf{r}_N \in \mathcal{R}_N$ , defined by*

$$(33) \quad \mathbf{r}(\mathbf{x}) = \mathbf{r}_N(\mathbf{x}) \quad \text{for every } \mathbf{x} \in \mathcal{X}_N,$$

*is an  $[\mathcal{H}]^2$ -orthogonal projection.*

*Proof.* First, recall that  $\mathbf{r}_x \in [\mathcal{H}(\partial\Omega)]_\nu^2$  for every  $x \in \mathcal{X}_N$ . Therefore,  $\mathbf{r}_N \in [\mathcal{H}(\partial\Omega)]_\nu^2$ . For  $x \in \mathcal{X}_N$  fixed, using  $\mathbf{r}$  and  $\mathbf{r}_N$  as test functions in (10a) gives

$$(34) \quad (\mathbf{r}_x, \mathbf{r})_{[\mathcal{H}]^2} = \boldsymbol{\nu}(x) \cdot \mathbf{r}(x) \stackrel{(33)}{=} \boldsymbol{\nu}(x) \cdot \mathbf{r}_N(x) = (\mathbf{r}_x, \mathbf{r}_N)_{[\mathcal{H}]^2}.$$

Therefore,

$$(35) \quad (\mathbf{r}_x, \mathbf{r} - \mathbf{r}_N)_{[\mathcal{H}]^2} = 0 \quad \text{for all } x \in \mathcal{X}_N. \quad \square$$

*Remark 2.20.* Lemma 2.19 implies that  $\mathbf{r}_N$  is the best approximation of  $\mathbf{r}$  in  $\mathcal{R}_N$  with respect to the  $[\mathcal{H}]^2$ -norm, and that  $\|\mathbf{r}_N\|_{[\mathcal{H}]^2} \leq \|\mathbf{r}\|_{[\mathcal{H}]^2}$ .

We conclude the section by showing that any function in  $[\mathcal{H}(\partial\Omega)]_\nu^2$  can be approximated arbitrarily well using sufficiently many  $\mathbf{r}_x$ 's.

**THEOREM 2.21.** *Let  $\{\mathcal{X}_N\}_{N \in \mathbb{N}}$  be a nested sequence of finite subsets (that contain pairwise distinct points) of  $\partial\Omega$  such that the union  $\cup_{N \in \mathbb{N}} \mathcal{X}_N$  is dense in  $\partial\Omega$ . The set  $\cup_{N \geq 1} \mathcal{R}_N$  is dense in  $[\mathcal{H}(\partial\Omega)]_\nu^2$ .*

*Proof.* For an arbitrary  $\mathbf{r} \in [\mathcal{H}(\partial\Omega)]_\nu^2$ , denote by  $\mathbf{r}_N$  its interpolant on  $\mathcal{R}_N$  (see Lemma 2.19). The sequence  $\{\mathbf{r}_N\}_{N \in \mathbb{N}}$  is bounded in  $[\mathcal{H}(\partial\Omega)]^2$  because  $\|\mathbf{r}_N\|_{[\mathcal{H}]^2} \leq \|\mathbf{r}\|_{[\mathcal{H}]^2}$  for every  $N$ . Therefore, there is a subsequence  $\{\mathbf{r}_{N_i}\}_{i \in \mathbb{N}}$  that converges weakly to a function  $\mathbf{f} \in [\mathcal{H}(\partial\Omega)]_\nu^2$ . By Mazur's lemma [12, p. 61], there is a sequence  $\{\hat{\mathbf{r}}_n\}_{n \in \mathbb{N}}$  of convex combinations of  $\{\mathbf{r}_{N_i}\}_{i \in \mathbb{N}}$  that converges strongly in  $[\mathcal{H}(\partial\Omega)]^2$  to  $\mathbf{f}$ . Note that there is a function  $\mathcal{N} : \mathbb{N} \rightarrow \mathbb{N}$  such that  $\mathcal{N}(n) \geq n$  and  $\hat{\mathbf{r}}_n \in \mathcal{R}_{\mathcal{N}(n)}$  for every  $n \in \mathbb{N}$ .

We recall that in RKHSs, strong convergence implies pointwise convergence. Therefore, for every  $z \in \partial\Omega$ ,

$$(36) \quad \hat{\mathbf{r}}_n(z) \rightarrow \mathbf{f}(z) \quad \text{as } n \rightarrow \infty.$$

Additionally,

$$(37) \quad \hat{\mathbf{r}}_n(z) = \mathbf{r}(z) \quad \text{for every } z \in \mathcal{X}_{\mathcal{N}(n)}$$

because  $\hat{\mathbf{r}}_n$  is a convex combination of interpolants. Therefore,  $\mathbf{f} = \mathbf{r}$  on  $\cup_{n \in \mathbb{N}} \mathcal{X}_{\mathcal{N}(n)}$ , and in turn, since  $\cup_{n \in \mathbb{N}} \mathcal{X}_{\mathcal{N}(n)}$  is dense in  $\partial\Omega$  and  $\mathbf{f}$  and  $\mathbf{r}$  are continuous, we conclude that  $\mathbf{f} = \mathbf{r}$  on  $\partial\Omega$ . It follows that  $\hat{\mathbf{r}}_n$  converges strongly to  $\mathbf{r}$ .  $\square$

**2.4. Approximation of weakly normal basis functions.** The  $\mathbf{r}_x$  functions introduced by Definition 2.11 live in the infinite dimensional space  $[\mathcal{H}(\partial\Omega)]^2$ . In this section, we explain how to approximate them using finitely many collocation points, and we analyze the numerical error of this approximation.

In this section, the sets  $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$  and  $\mathcal{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_M\}$ ,  $M \leq N$ , denote two subsets of  $\partial\Omega$  (containing pairwise distinct points) with  $\mathcal{Y} \subset \mathcal{X}$ . These sets are used to define the finite dimensional spaces

$$(38) \quad V_h(\mathcal{X}) = \text{span}\{\mathbf{k}(\mathbf{x}, \cdot)|_{\partial\Omega} : \mathbf{x} \in \mathcal{X}\} \quad \text{and} \quad V_h(\mathcal{Y}) = \text{span}\{\mathbf{k}(\mathbf{y}, \cdot)|_{\partial\Omega} : \mathbf{y} \in \mathcal{Y}\}.$$

Finally, we denote by  $\mathcal{I}_{V_h} : \mathcal{H}(\partial\Omega) \rightarrow V_h(\mathcal{X})$  and  $\mathcal{I}_{[V_h]^2} : [\mathcal{H}(\partial\Omega)]^2 \rightarrow [V_h(\mathcal{X})]^2$  the standard pointwise interpolation operators.

**DEFINITION 2.22.** *Let  $\mathbf{x} \in \partial\Omega$ . The finite dimensional approximation  $(\mathbf{r}_x^h, p_x^h) \in [V_h(\mathcal{X})]^2 \times [V_h(\mathcal{Y})]$  of  $(\mathbf{r}_x, p_x)$  is characterized by*

$$(39a) \quad (\mathbf{r}_x^h, \varphi)_{[\mathcal{H}]^2} + (\varphi \cdot \boldsymbol{\tau}, p_x^h)_\mathcal{H} = \boldsymbol{\nu}(x) \cdot \varphi(x) \quad \text{for all } \varphi \in [V_h(\mathcal{X})]^2,$$

$$(39b) \quad (\mathbf{r}_x^h \cdot \boldsymbol{\tau}, \psi)_\mathcal{H} = 0 \quad \text{for all } \psi \in V_h(\mathcal{Y}).$$

Theorem 2.27 shows that Definition 2.22 is well defined. Its proof relies on Lemma 2.26, which is a discrete counterpart of Lemma 2.12. To prove this lemma, we use the following properties of the interpolation operators  $\mathcal{I}_{V_h}$  and  $\mathcal{I}_{[V_h]^2}$ .

**LEMMA 2.23.** *The interpolation operator  $\mathcal{I}_{V_h}$  is an  $\mathcal{H}$ -orthogonal projection.*

*Proof.* Let  $\mathbf{x}$  be a point from  $\mathcal{X}$ , and choose a  $\psi \in \mathcal{H}(\partial\Omega)$ . Since  $\mathbf{x}$  is an interpolation point, it holds that

$$(40) \quad (\mathcal{I}_{V_h}(\psi), \mathbf{k}(\mathbf{x}, \cdot))_{\mathcal{H}} = \mathcal{I}_{V_h}(\psi)(\mathbf{x}) = \psi(\mathbf{x}) = (\psi, \mathbf{k}(\mathbf{x}, \cdot))_{\mathcal{H}},$$

and thus

$$(41) \quad (\mathcal{I}_{V_h}(\psi) - \psi, \phi)_{\mathcal{H}} = 0 \quad \text{for all } \phi \in V_h(\mathcal{X}). \quad \square$$

**COROLLARY 2.24.** *If Assumption 2.8 holds true, then the interpolation operator  $\mathcal{I}_{V_h}$  satisfies*

$$(42) \quad \|\mathcal{I}_{V_h}(f\psi)\|_{\mathcal{H}} \leq c_{\Omega} \|f\|_{C^k(\partial\Omega)} \|\psi\|_{\mathcal{H}}$$

for every  $f \in C^k(\partial\Omega)$  and every  $\psi \in \mathcal{H}(\partial\Omega)$ .

*Proof.* Choose an  $f \in C^k(\partial\Omega)$  and a  $\psi \in \mathcal{H}(\partial\Omega)$ . Since  $\mathcal{I}_{V_h}$  is an  $\mathcal{H}$ -orthogonal projection, since  $V_h(\mathcal{X}) \subset \mathcal{H}(\partial\Omega)$ , and by (9), it holds that

$$(43) \quad \|\mathcal{I}_{V_h}(f\psi)\|_{\mathcal{H}} \leq \|f\psi\|_{\mathcal{H}} \leq c_{\Omega} \|f\|_{C^k(\partial\Omega)} \|\psi\|_{\mathcal{H}}. \quad \square$$

**COROLLARY 2.25.** *If Assumption 2.8 holds true, then the interpolation operator  $\mathcal{I}_{[V_h]^2}$  satisfies*

$$(44) \quad \|\mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi)\|_{[\mathcal{H}]^2} \leq c_{\Omega} \|\boldsymbol{\tau}\|_{C^k} \|\psi\|_{\mathcal{H}} \quad \text{for all } \psi \in \mathcal{H}(\partial\Omega).$$

*Proof.* Note that  $\mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi) = \mathcal{I}_{V_h}(\boldsymbol{\tau} \cdot \mathbf{e}_1 \psi) \mathbf{e}_1 + \mathcal{I}_{V_h}(\boldsymbol{\tau} \cdot \mathbf{e}_2 \psi) \mathbf{e}_2$ . Therefore, by Corollary 2.24,

$$\begin{aligned} \|\mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi)\|_{[\mathcal{H}]^2}^2 &= \|\mathcal{I}_{V_h}(\boldsymbol{\tau} \cdot \mathbf{e}_1 \psi)\|_{\mathcal{H}}^2 + \|\mathcal{I}_{V_h}(\boldsymbol{\tau} \cdot \mathbf{e}_2 \psi)\|_{\mathcal{H}}^2 \\ &\leq c_{\Omega}^2 \|\boldsymbol{\tau} \cdot \mathbf{e}_1\|_{C^k}^2 \|\psi\|_{\mathcal{H}}^2 + c_{\Omega}^2 \|\boldsymbol{\tau} \cdot \mathbf{e}_2\|_{C^k}^2 \|\psi\|_{\mathcal{H}}^2 \\ &= c_{\Omega}^2 \|\boldsymbol{\tau}\|_{C^k}^2 \|\psi\|_{\mathcal{H}}^2. \end{aligned} \quad \square$$

**LEMMA 2.26** (discrete inf-sup condition). *The following holds:*

$$(45) \quad \sup_{\substack{\boldsymbol{\varphi} \in [V_h(\mathcal{X})]^2, \\ \|\boldsymbol{\varphi}\|_{[\mathcal{H}]^2} \leq 1}} (\boldsymbol{\varphi} \cdot \boldsymbol{\tau}, \psi)_{\mathcal{H}} \geq \frac{\|\psi\|_{\mathcal{H}}}{c_{\Omega} \|\boldsymbol{\tau}\|_{C^k}} \quad \text{for all } \psi \in V_h(\mathcal{Y}).$$

*Proof.* Let  $\psi \in V_h(\mathcal{Y})$ . By Corollary 2.25,

$$(46) \quad \sup_{\substack{\boldsymbol{\varphi} \in [V_h(\mathcal{X})]^2, \\ \|\boldsymbol{\varphi}\|_{[\mathcal{H}]^2} \leq 1}} (\boldsymbol{\varphi} \cdot \boldsymbol{\tau}, \psi)_{\mathcal{H}} \geq \left( \boldsymbol{\tau} \cdot \frac{\mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi)}{\|\mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi)\|_{[\mathcal{H}]^2}}, \psi \right)_{\mathcal{H}} \geq \frac{(\boldsymbol{\tau} \cdot \mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi), \psi)_{\mathcal{H}}}{c_{\Omega} \|\boldsymbol{\tau}\|_{C^k} \|\psi\|_{\mathcal{H}}}.$$

Let  $\mathbf{y} \in \mathcal{Y}$ . Then, since  $\mathcal{Y} \subset \mathcal{X}$ ,

$$\begin{aligned} (\boldsymbol{\tau} \cdot \mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi), \mathbf{k}(\mathbf{y}, \cdot))_{\mathcal{H}} &= \boldsymbol{\tau}(\mathbf{y}) \cdot \mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi)(\mathbf{y}) \\ &= \boldsymbol{\tau}(\mathbf{y}) \cdot (\boldsymbol{\tau}\psi)(\mathbf{y}) = \psi(\mathbf{y}) = (\psi, \mathbf{k}(\mathbf{y}, \cdot))_{\mathcal{H}}. \end{aligned}$$

Therefore, (45) follows by plugging  $(\boldsymbol{\tau} \cdot \mathcal{I}_{[V_h]^2}(\boldsymbol{\tau}\psi), \psi)_{\mathcal{H}} = (\psi, \psi)_{\mathcal{H}} = \|\psi\|_{\mathcal{H}}^2$  into (46).  $\square$

**THEOREM 2.27.** *The finite dimensional problem (39) has a unique solution, which satisfies*

$$(47) \quad \|\mathbf{r}_{\mathbf{x}}^h\|_{[\mathcal{H}]^2} \leq 1 \quad \text{and} \quad \|p_{\mathbf{x}}^h\|_{\mathcal{H}} \leq 2c_{\Omega}\|\boldsymbol{\tau}\|_{C^k}.$$

Moreover,

$$(48) \quad \|\mathbf{r}_{\mathbf{x}}^h - \mathbf{r}_{\mathbf{x}}\|_{[\mathcal{H}]^2} \leq \|\boldsymbol{\nu}\|_{C^k}^2 \inf_{\varphi \in [V_h(\mathcal{X})]^2} \|\mathbf{r}_x - \varphi\|_{[\mathcal{H}]^2} + \|\boldsymbol{\tau}\|_{C^k} \inf_{\psi \in V_h(\mathcal{Y})} \|p_{\mathbf{x}} - \psi\|_{\mathcal{H}}$$

and

$$(49) \quad \|p_{\mathbf{x}}^h - p_{\mathbf{x}}\|_{\mathcal{H}} \leq c_{\Omega}\|\boldsymbol{\tau}\|_{C^k} \|\mathbf{r}_{\mathbf{x}}^h - \mathbf{r}_{\mathbf{x}}\|_{[\mathcal{H}]^2} + (1 + c_{\Omega}^2\|\boldsymbol{\tau}\|_{C^k}^2) \inf_{\psi \in V_h(\mathcal{Y})} \|p_{\mathbf{x}} - \psi\|_{\mathcal{H}}.$$

*Proof.* The functions  $\mathbf{r}_{\mathbf{x}}^h$  and  $p_{\mathbf{x}}^h$  can be written as

$$(50) \quad \mathbf{r}_{\mathbf{x}}^h = \sum_{i=1}^N \boldsymbol{\alpha}_i \mathbf{k}(\mathbf{x}_i, \cdot), \quad p_{\mathbf{x}}^h = \sum_{i=1}^M \beta_i \mathbf{k}(\mathbf{y}_i, \cdot),$$

where  $\boldsymbol{\alpha}_i = (\alpha_i^1, \alpha_i^2)^\top \in \mathbf{R}^2$  and  $\beta_i \in \mathbf{R}$  are certain unknown coefficients.

The vectors  $\mathbf{a} := (\alpha_1^1, \dots, \alpha_N^1, \alpha_1^2, \dots, \alpha_N^2)^\top$  and  $\mathbf{b} := (\beta_1, \dots, \beta_M)^\top$  satisfy

$$(51) \quad \begin{pmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} = \begin{pmatrix} \mathbf{L} \\ \mathbf{0} \end{pmatrix},$$

where

$$(52) \quad \mathbf{A} = \begin{pmatrix} \tilde{\mathbf{A}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{A}} \end{pmatrix}, \quad \mathbf{B} = (\mathbf{B}^1 \mathbf{B}^2), \quad \mathbf{L} = \begin{pmatrix} \mathbf{L}^1 \\ \mathbf{L}_2 \end{pmatrix},$$

and (with  $\boldsymbol{\tau} = (\tau^1, \tau^2)^\top$  and  $\boldsymbol{\nu} = (\nu^1, \nu^2)^\top$ )

$$(53) \quad (\tilde{\mathbf{A}})_{n,m} = \mathbf{k}(\mathbf{x}_n, \mathbf{x}_m), \quad n, m = 1, \dots, N,$$

$$(54) \quad (\mathbf{B}^\ell)_{m,n} = \mathbf{k}(\mathbf{x}_n, \mathbf{y}_m) \tau^\ell(\mathbf{x}_n), \quad m = 1, \dots, M, \quad n = 1, \dots, N,$$

$$(55) \quad (\mathbf{L}^\ell)_n = \nu^\ell(\mathbf{x}) \mathbf{k}(\mathbf{x}_n, \mathbf{x}), \quad n = 1, \dots, N.$$

The linear system (51) has a unique solution because  $\tilde{\mathbf{A}}$  (and thus  $\mathbf{A}$ ) is positive definite and  $\mathbf{B}$  has maximal rank  $M$ . The maximal rank of  $\mathbf{B}$  follows directly from Lemma 2.26. For a more direct proof, recall that  $\mathcal{Y} \subset \mathcal{X}$ . Therefore, the matrix  $\mathbf{B}$  contains the  $M$  column vectors  $\{(\mathbf{k}(\mathbf{y}_i, \mathbf{y}_j)c_i)_{j=1, \dots, M} : i = 1, \dots, M\}$ , where  $c_i = \tau^1(\mathbf{y}_i)$  if  $\tau^1(\mathbf{y}_i) \neq 0$  and  $c_i = \tau^2(\mathbf{y}_i)$  otherwise (so that  $c_i \neq 0$  for every  $i$ ). These vectors are linearly independent because the kernel  $\mathbf{k}$  is positive definite.

The estimate (47) follows by inserting  $\varphi = \mathbf{r}_{\mathbf{x}}^h$  into (39a) and using the Cauchy–Schwarz inequality and  $\|\mathbf{k}(\mathbf{x}, \cdot)\|_{\mathcal{H}} = 1$  (see (20)). To prove the estimate on the multiplier  $p_{\mathbf{x}}^h$ , note that, by (45),

$$(56) \quad \|p_{\mathbf{x}}^h\| \leq c_{\Omega}\|\boldsymbol{\tau}\|_{C^k} \sup_{\substack{\varphi \in [V_h(\mathcal{X})]^2, \\ \|\varphi\|_{[\mathcal{H}]^2} \leq 1}} (\varphi \cdot \boldsymbol{\tau}, p_{\mathbf{x}}^h)_{\mathcal{H}},$$

and that by (39a), for every  $\varphi \in [V_h(\mathcal{X})]^2$  with  $\|\varphi\|_{[\mathcal{H}]^2} \leq 1$ ,

$$\begin{aligned} (\varphi \cdot \boldsymbol{\tau}, p_{\mathbf{x}}^h)_{\mathcal{H}} &= \boldsymbol{\nu}(\mathbf{x}) \cdot \varphi(\mathbf{x}) - (\mathbf{r}_{\mathbf{x}}^h, \varphi)_{[\mathcal{H}]^2} \\ &\leq |\boldsymbol{\nu}(\mathbf{x})| + \|\mathbf{r}_{\mathbf{x}}^h\|_{[\mathcal{H}]^2} \|\varphi\|_{[\mathcal{H}]^2} \leq \|\boldsymbol{\nu}\|_{[\mathcal{H}]^2} + 1 \leq 2. \end{aligned}$$

Finally, quasi-optimality results similar to (48) and (49) can be derived by applying directly [5, Prop. 5.2.1, p. 274]. Here, we give explicit details on the derivation of (48) and (49).

To show (48), note that the differences  $\mathbf{r}_x^h - \mathbf{r}_x$  and  $p_x^h - p_x$  satisfy

$$(57a) \quad (\mathbf{r}_x^h - \mathbf{r}_x, \varphi)_{[\mathcal{H}]^2} + (\varphi \cdot \boldsymbol{\tau}, p_x^h - p_x)_{\mathcal{H}} = 0 \quad \text{for all } \varphi \in [V_h(\mathcal{X})]^2,$$

$$(57b) \quad ((\mathbf{r}_x^h - \mathbf{r}_x) \cdot \boldsymbol{\tau}, \psi)_{\mathcal{H}} = 0 \quad \text{for all } \psi \in V_h(\mathcal{Y}).$$

This implies that

$$\begin{aligned} \|\mathbf{r}_x^h - \mathbf{r}_x\|_{[\mathcal{H}]^2}^2 &= (\mathbf{r}_x^h - \mathbf{r}_x, \mathbf{r}_x^h - \mathbf{r}_x)_{[\mathcal{H}]^2} \\ &= (\mathbf{r}_x^h - \mathbf{r}_x, \mathbf{r}_x^h)_{[\mathcal{H}]^2} - (\mathbf{r}_x^h - \mathbf{r}_x, \mathbf{r}_x)_{[\mathcal{H}]^2} \\ &= -(\mathbf{r}_x^h \cdot \boldsymbol{\tau}, p_x^h - p_x)_{\mathcal{H}} - (\mathbf{r}_x^h - \mathbf{r}_x, \mathbf{r}_x)_{[\mathcal{H}]^2} \\ &= -(\mathbf{r}_x^h \cdot \boldsymbol{\tau}, p_x^h)_{\mathcal{H}} + (\mathbf{r}_x^h \cdot \boldsymbol{\tau}, p_x)_{\mathcal{H}} - (\mathbf{r}_x^h - \mathbf{r}_x, \mathbf{r}_x)_{[\mathcal{H}]^2} \\ &= (\mathbf{r}_x^h \cdot \boldsymbol{\tau}, p_x)_{\mathcal{H}} - (\mathbf{r}_x^h - \mathbf{r}_x, \mathbf{r}_x)_{[\mathcal{H}]^2} \\ &= (\mathbf{r}_x^h \cdot \boldsymbol{\tau}, p_x)_{\mathcal{H}} - (\mathbf{r}_x^h - \mathbf{r}_x, (\mathbf{r}_x \cdot \boldsymbol{\nu}) \cdot \boldsymbol{\nu})_{[\mathcal{H}]^2}. \end{aligned}$$

The latter result combined with (57) implies that, for any  $\varphi \in [V_h(\mathcal{X})]^2$  and any  $\psi \in V_h(\mathcal{Y})$ ,

$$(58) \quad \|\mathbf{r}_x^h - \mathbf{r}_x\|_{[\mathcal{H}]^2}^2 = (\mathbf{r}_x^h \cdot \boldsymbol{\tau} - \mathbf{r}_x \cdot \boldsymbol{\tau}, p_x - \psi)_{\mathcal{H}} - (\mathbf{r}_x^h - \mathbf{r}_x, (\mathbf{r}_x \cdot \boldsymbol{\nu}) \cdot \boldsymbol{\nu} - (\varphi \cdot \boldsymbol{\nu}) \boldsymbol{\nu})_{[\mathcal{H}]^2}.$$

Then, the Cauchy–Schwarz inequality and (9) imply

$$\begin{aligned} \|\mathbf{r}_x^h - \mathbf{r}_x\|_{[\mathcal{H}]^2}^2 &\leq \|\mathbf{r}_x^h \cdot \boldsymbol{\tau} - \mathbf{r}_x \cdot \boldsymbol{\tau}\|_{\mathcal{H}} \|p_x - \psi\|_{\mathcal{H}} + \|\mathbf{r}_x^h - \mathbf{r}_x\|_{[\mathcal{H}]^2} \|((\mathbf{r}_x - \varphi) \cdot \boldsymbol{\nu}) \boldsymbol{\nu}\|_{[\mathcal{H}]^2} \\ (59) \quad &\leq \|\mathbf{r}_x^h - \mathbf{r}_x\|_{[\mathcal{H}]^2} (\|\boldsymbol{\tau}\|_{C^k} \|p_x - \psi\|_{\mathcal{H}} + \|\boldsymbol{\nu}\|_{C^k}^2 \|\mathbf{r}_x - \varphi\|_{[\mathcal{H}]^2}), \end{aligned}$$

from which (48) follows easily.

To show (49), we note that for any  $\psi \in V_h(\mathcal{Y})$ , (45) implies

$$(60) \quad \|p_x^h - \psi\|_{\mathcal{H}} \leq c_{\Omega} \|\boldsymbol{\tau}\|_{C^k} \sup_{\substack{\varphi \in [V_h(\mathcal{X})]^2, \\ \|\varphi\|_{[\mathcal{H}]^2} \leq 1}} (\varphi \cdot \boldsymbol{\tau}, p_x^h - \psi)_{\mathcal{H}}.$$

For any  $\varphi \in [V_h(\mathcal{X})]^2$ , (10) and (39) imply

$$\begin{aligned} (\varphi \cdot \boldsymbol{\tau}, p_x^h - \psi)_{\mathcal{H}} &= (\varphi \cdot \boldsymbol{\tau}, p_x^h)_{\mathcal{H}} - (\varphi \cdot \boldsymbol{\tau}, \psi)_{\mathcal{H}} \\ &= \boldsymbol{\nu}(\mathbf{x}) \cdot \varphi(\mathbf{x}) - (\mathbf{r}_x^h, \varphi)_{[\mathcal{H}]^2} - (\varphi \cdot \boldsymbol{\tau}, \psi)_{\mathcal{H}} \\ &= (\mathbf{r}_x, \varphi)_{[\mathcal{H}]^2} + (\varphi \cdot \boldsymbol{\tau}, p_x)_{\mathcal{H}} - (\mathbf{r}_x^h, \varphi)_{[\mathcal{H}]^2} - (\varphi \cdot \boldsymbol{\tau}, \psi)_{\mathcal{H}} \\ &= (\mathbf{r}_x - \mathbf{r}_x^h, \varphi)_{[\mathcal{H}]^2} + (\varphi \cdot \boldsymbol{\tau}, p_x - \psi)_{\mathcal{H}} \\ &\leq \|\psi\|_{[\mathcal{H}]^2} (\|\mathbf{r}_x - \mathbf{r}_x^h\|_{[\mathcal{H}]^2} + c_{\Omega} \|\boldsymbol{\tau}\|_{C^k} \|p_x - \psi\|_{\mathcal{H}}), \end{aligned}$$

where the last inequality follows from the Cauchy–Schwarz inequality and (9). Therefore,

$$(61) \quad \|p_x^h - \psi\|_{\mathcal{H}} \leq c_{\Omega} \|\boldsymbol{\tau}\|_{C^k} (\|\mathbf{r}_x - \mathbf{r}_x^h\|_{[\mathcal{H}]^2} + c_{\Omega} \|\boldsymbol{\tau}\|_{C^k} \|p_x - \psi\|_{\mathcal{H}}),$$

and, by the triangle inequality,

$$\begin{aligned} \|p_x - p_x^h\|_{\mathcal{H}} &\leq \|p_x - \psi\|_{\mathcal{H}} + \|p_x^h - \psi\|_{\mathcal{H}} \\ &\leq c_{\Omega} \|\boldsymbol{\tau}\|_{C^k} \|\mathbf{r}_x - \mathbf{r}_x^h\|_{[\mathcal{H}]^2} + (1 + c_{\Omega}^2 \|\boldsymbol{\tau}\|_{C^k}^2) \|p_x - \psi\|_{\mathcal{H}}. \end{aligned} \quad \square$$

The following corollary follows directly from the quasi-optimality estimates (48) and (49).

**COROLLARY 2.28.** *Let  $\{\mathcal{Y}_N\}_{N \in \mathbb{N}}$  be a nested sequence of finite subsets (that contain pairwise distinct points) of  $\partial\Omega$  such that the union  $\cup_{N \in \mathbb{N}} \mathcal{Y}_N$  is dense in  $\partial\Omega$ . Then, the sequence of functions  $(\mathbf{r}_x^{h_N}, p_x^{h_N})$  converges strongly to  $(\mathbf{r}_x, p_x)$ .*

Before studying convergence rates, we mention that in [31] Sturm proved superlinear convergence of a shape Newton method in the Micheletti space based on functions of the form  $k(\mathbf{x}, \cdot)\boldsymbol{\nu}(\mathbf{x})$ . The next lemma shows that such functions arise from a particular discretization of (10).

**PROPOSITION 2.29.** *If  $M = 1$  and  $\mathcal{Y} = \{\mathbf{x}\} \subset \mathcal{X}$ , then the solution of (39) is*

$$\mathbf{r}_x^h(\cdot) = k(\mathbf{x}, \cdot)\boldsymbol{\nu}(\mathbf{x}), \quad p_x^h = 0.$$

*Proof.* By Theorem 2.27, we need only verify that the pair  $(\mathbf{r}_x^h, p_x^h)$  satisfies (39). By the reproducing kernel property,

$$(\mathbf{r}_x^h, \varphi)_{[\mathcal{H}]^2} = \boldsymbol{\nu}(\mathbf{x}) \cdot \varphi(\mathbf{x}) \quad \text{for all } \varphi \in [\mathcal{H}(\partial\Omega)]^2,$$

which, together with  $p_x^h = 0$ , implies (39a). Finally,

$$(62) \quad (\mathbf{r}_x^h \cdot \boldsymbol{\tau}, k(\mathbf{x}, \cdot))_{\mathcal{H}} = k(\mathbf{x}, \mathbf{x})\boldsymbol{\tau}(\mathbf{x}) \cdot \boldsymbol{\nu}(\mathbf{x}) = 0,$$

together with  $V_h(\mathcal{Y}) = \text{span}\{k(\mathbf{x}, \cdot)\}$ , implies (39b).  $\square$

Henceforth, we restrict ourselves to radial kernels, for which a solid interpolation theory on submanifolds is available.

**ASSUMPTION 2.30.** *The kernel  $k$  satisfies  $k(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x} - \mathbf{y})$  for a positive-definite function  $\Phi \in L_2(\mathbf{R}^d) \cap C(\mathbf{R}^d)$ . Assume that the Fourier transform  $\hat{\Phi}$  of  $\Phi$  satisfies*

$$(63) \quad c_1(1 + \|\mathbf{x}\|^2)^{-\xi} \leq \hat{\Phi}(\mathbf{x}) \leq c_2(1 + \|\mathbf{x}\|^2)^{-\xi} \quad \text{for all } \mathbf{x} \in \mathbf{R}^2,$$

with  $\xi > 1$  and two constants  $c_1, c_2 > 0$ ,  $c_1 \leq c_2$ .

**Remark 2.31.** Assumption 2.30 implies that the native space of  $k$  on  $\mathbf{R}^2$  is  $\mathcal{H}(\mathbf{R}^2) = H^\xi(\mathbf{R}^2)$  with equivalent norms. Thus, by the Sobolev trace theorem, the native space of  $k$  on  $\partial\Omega$  is  $\mathcal{H}(\partial\Omega) = H^{\xi-1/2}(\partial\Omega)$  [17, Prop. 5].

For such kernels, we can derive convergence rates of  $\mathbf{r}_x^h$  to  $\mathbf{r}_x$  and  $p_x^h$  to  $p_x$  in certain Sobolev spaces. For brevity, we focus on the more interesting term  $\mathbf{r}_x$  (convergence rates of  $p_x^h$  can be derived in exactly the same way). Before stating the theorem, we recall the concept of fill distance of discrete sets.

**DEFINITION 2.32.** *The fill distance associated with a set  $\mathcal{X} \subset \partial\Omega$  is defined by*

$$(64) \quad h_{\mathcal{X}, \partial\Omega} := \sup_{\mathbf{y} \in \partial\Omega} \inf_{\mathbf{x} \in \mathcal{X}} d_{\partial\Omega}(\mathbf{x}, \mathbf{y}),$$

where  $d_{\partial\Omega}$  denotes the geodesic distance associated with  $\partial\Omega$ . Note that  $d_{\partial\Omega}$  is equivalent to the Euclidean metric restricted on  $\partial\Omega$  [17, Thm. 6]; that is, there are two constants  $c_1, c_2 > 0$  (that depend only on  $\partial\Omega$ ) such that

$$(65) \quad c_1\|\mathbf{x} - \mathbf{y}\| \leq d_{\partial\Omega}(\mathbf{x}, \mathbf{y}) \leq c_2\|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in \partial\Omega.$$

Given a continuous kernel  $\kappa$  we define the integral operator  $T : L_2(\partial\Omega) \rightarrow L_2(\partial\Omega)$  via

$$(66) \quad f \mapsto \left( Tf : \mathbf{x} \mapsto \int_{\partial\Omega} \kappa(\mathbf{x}, \mathbf{s}) f(\mathbf{s}) ds \right).$$

We define similarly the integral operator  $\mathbf{T}$  that acts on  $L_2(\partial\Omega, \mathbf{R}^2)$ .

**THEOREM 2.33.** *Let  $\mathcal{Y} = \mathcal{X}$ ,  $\xi > 1$ , and  $s = \xi - 1/2$ . Let  $\kappa$  satisfy Assumption 2.30 so that  $\mathcal{H}(\partial\Omega) := H^s(\partial\Omega)$ . Assume that  $T^{-1}(p_{\mathbf{x}}) \in L_2(\partial\Omega)$  and  $\mathbf{T}^{-1}(\mathbf{r}_{\mathbf{x}}) \in L_2(\partial\Omega, \mathbf{R}^2)$ . Then, there is a constant  $c_{\partial\Omega} > 0$ , such that*

$$(67) \quad \|\mathbf{r}_{\mathbf{x}}^{h_{\mathcal{X}, \partial\Omega}} - \mathbf{r}_{\mathbf{x}}\|_{H^s(\partial\Omega)} \leq c_{\partial\Omega} h_{\mathcal{X}, \partial\Omega}^s (\|\boldsymbol{\nu}\|_{C^k}^2 \|\mathbf{T}^{-1}\mathbf{r}_{\mathbf{x}}\|_{L_2(\partial\Omega, \mathbf{R}^2)} + \|\boldsymbol{\tau}\|_{C^k} \|T^{-1}p_{\mathbf{x}}\|_{L_2(\partial\Omega)})$$

for all  $h_{\mathcal{X}, \partial\Omega} \leq h_{\partial\Omega}$ .

*Proof.* In Corollary 15 from [17] it is shown that there is a constant  $c > 0$ , such that

$$(68) \quad \|\mathcal{I}_{V_h}(\psi) - \psi\|_{\mathcal{H}} \leq ch_{\mathcal{X}, \partial\Omega}^s \|T^{-1}\psi\|_{L_2(\partial\Omega)} \quad \text{for all } \psi \in \mathcal{H}(\partial\Omega).$$

A similar result holds for the operator  $\mathbf{T}$ . Hence, replacing  $\boldsymbol{\varphi} = \mathcal{I}_{[V_h]^2}(\mathbf{r}_x)$  and  $\psi = \mathcal{I}_{V_h}(p_{\mathbf{x}})$  in (48) and using the estimate (68) gives the desired results.  $\square$

**COROLLARY 2.34.** *Let the hypotheses of Theorem 2.33 be satisfied. Then, there is a constant  $c$  such that*

$$(69) \quad \|\mathbf{r}_{\mathbf{x}}^{h_{\mathcal{X}, \partial\Omega}} - \mathbf{r}_{\mathbf{x}}\|_{L_2(\partial\Omega, \mathbf{R}^2)} \leq ch_{\mathcal{X}, \partial\Omega}^{2s}$$

for all  $h_{\mathcal{X}, \partial\Omega} \leq h_{\partial\Omega}$ .

*Proof.* Applying [17, Lem. 10] yields

$$(70) \quad \|\boldsymbol{\varphi} - \mathcal{I}_{[V_h]^2}(\boldsymbol{\varphi})\|_{L_2(\partial\Omega, \mathbf{R}^2)} \leq ch_{\mathcal{X}, \partial\Omega}^s \|\boldsymbol{\varphi} - \mathcal{I}_{[V_h]^2}(\boldsymbol{\varphi})\|_{\mathcal{H}}$$

for all  $\boldsymbol{\varphi} \in [\mathcal{H}(\partial\Omega)]^2$  (and similarly for  $\mathcal{I}_{V_h}$ ). Hence combining this inequality with (67) gives the desired estimate.  $\square$

**NUMERICAL EXPERIMENT 2.35.** *We perform a numerical experiment to verify the convergence rates derived in Theorem 2.33. We choose  $\partial\Omega$  to be the ellipse depicted in Figure 1 (left) and consider the sequence of subsets*

$$(71) \quad \mathcal{X}_N = \{\gamma(2\pi\ell/N) : \ell = 1, \dots, N\} \quad \text{with} \quad N = 2^4, 2^5, \dots, 2^{10}.$$

*Then, for each  $N$ , we compute  $\mathbf{r}_{\gamma(0)}^{h_N}$ , where  $h_N$  denotes the fill distance of  $\mathcal{X}_N$ . Finally, we compute (with sufficiently many quadrature points) the  $L_2(\partial\Omega)$ -absolute error using  $\mathbf{r}_{\gamma(0)}^{h_{2^{10}}}$  as a reference solution. In Figure 2, we plot these  $L_2$ -errors versus  $h_N$  (measured with the Euclidean distance) for the kernels  $\kappa_4^{1,2}$ ,  $\kappa_6^{1,2}$ , and  $\kappa_8^{1,2}$ . The measured convergence rates read (approximately) 4.51, 6.79, and 8.77, respectively, and are slightly better than but still in line with Corollary 2.34. Comparable results are obtained for the nonconvex domain depicted in Figure 1 (left), as well as for different values of  $\sigma$ .*

**Remark 2.36.** It is known from scattered data approximation theory [25] that the condition number of  $(\kappa(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n$  can be very large for certain collections of points

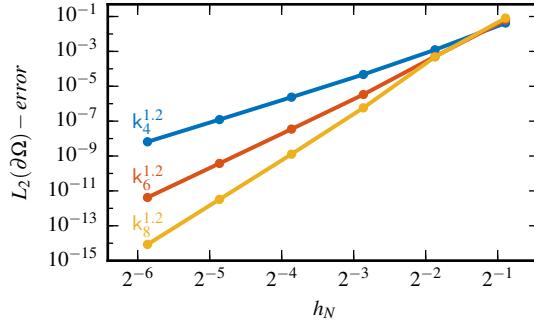


FIG. 2. Numerical Experiment 2.35 confirms the convergence rates predicted by Corollary 2.34: the error  $\|\mathbf{r}_x^{h_N} - \mathbf{r}_x\|_{L_2(\partial\Omega)}$  decays with the algebraic rates 4.51 (for  $k_4^{1,2}$ ), 6.79 (for  $k_6^{1,2}$ ), and 8.77 (for  $k_8^{1,2}$ ).

$\{\mathbf{x}_i\}_{i=1}^n$ . We have computed the condition number of the saddle point matrix from (51) for different  $\sigma$ 's and  $h_N$  and have observed that, similarly to the condition number of  $(k(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n$ , it behaves roughly as  $(\sigma/h_N)^p$ , where  $p$  is the polynomial order of the Wendland kernel. Note that by combining Gershgorin's theorem [33, Chap. 12, p. 207] and the result listed in [33, Chap. 12.2, Table 12.1], it is possible to show that the condition number of  $(k(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n$  is bounded from above by  $C(h_N)^{-p}$ , where  $C$  is a positive constant.

**3. Application to shape Newton methods.** The formulation of shape Newton methods is a delicate task, and not all second-order shape derivatives are suited for this job. The study of second-order shape derivatives is beyond the scope of this work, and we refer the interested reader to [26, 31]. In this section, we briefly recall the shape Newton method introduced in [31], discuss its discretization with the weakly normal  $\mathbf{r}_x$  functions, and perform some numerical simulations for an unconstrained shape optimization test case. This test case is kept simple on purpose in order to restrict the discretization error to the use of weakly-normal basis functions. In this section, we use  $\mathcal{X} = \mathcal{Y}$  in Definition 2.22.

**3.1. Shape derivatives and their structure theorem.** We begin by recalling the definition of shape derivatives and their structure theorem from [31]. For a given set  $D \subset \mathbf{R}^2$ , we denote by  $\wp(D)$  the powerset of  $D$ . Furthermore we use the notation  $\mathring{C}^1(\bar{D}, \mathbf{R}^d)$  to indicate the Banach space of functions  $\mathbf{X} \in C^1(\bar{D}, \mathbf{R}^d)$  that satisfy  $\mathbf{X} = 0$  on  $\partial D$ .

DEFINITION 3.1. Let  $J : \Xi \subset \wp(D) \rightarrow \mathbf{R}$  be a shape function, and let  $\mathbf{X}, \mathbf{Y} \in \mathring{C}^1(\bar{D}, \mathbf{R}^2)$  be two vector fields.

(i) The directional derivative of  $J$  at  $\Omega \in \Xi$  in the direction  $\mathbf{X}$  is defined by

$$(72) \quad DJ(\Omega)(\mathbf{X}) := \lim_{t \rightarrow 0^+} \frac{J((\mathbf{id} + t\mathbf{X})(\Omega)) - J(\Omega)}{t}.$$

(ii) The second directional derivative of  $J$  at  $\Omega$  in the direction  $(\mathbf{X}, \mathbf{Y})$  is defined by

$$(73) \quad \mathfrak{D}^2 J(\Omega)(\mathbf{X})(\mathbf{Y}) = \lim_{t \rightarrow 0^+} \frac{DJ((\mathbf{id} + t\mathbf{Y})(\Omega))(\mathbf{X} \circ (\mathbf{id} + t\mathbf{Y})^{-1}) - DJ(\Omega)(\mathbf{X})}{t}$$

(assuming that  $DJ((\mathbf{id} + t\mathbf{Y})(\Omega))(\mathbf{X} \circ (\mathbf{id} + t\mathbf{Y})^{-1})$  exists for all small  $t$ ).

*Remark 3.2.* Assume that the map  $J_\Omega : \mathring{C}^1(\bar{\Omega}, \mathbf{R}^d) \rightarrow \mathbf{R}$ ,  $\mathbf{X} \mapsto J((\mathbf{id} + \mathbf{X})(\Omega))$  is twice Fréchet differentiable at 0. Then, the bilinear operator  $(\mathbf{X}, \mathbf{Y}) \mapsto \mathfrak{D}^2 J(\Omega)(\mathbf{X})(\mathbf{Y})$  is the second derivative of  $J_\Omega$  at 0, and thus symmetric [3, Thm. 5.2, p. 189].

We denote by  $\nabla^\tau f$  the tangential gradient of function  $f$  and by  $\mathbf{X}_\tau := \mathbf{X}|_{\partial\Omega} - (\mathbf{X}|_{\partial\Omega} \cdot \boldsymbol{\nu})\boldsymbol{\nu}$  the tangential component of a vector field  $\mathbf{X}$  on  $\partial\Omega$ . The shape derivatives  $DJ$  and  $\mathfrak{D}^2 J$  are characterized by the following structure theorem (see [22, 31]).

**THEOREM 3.3.** *Assume that  $J$  is twice differentiable at  $\Omega$  and that  $\Omega$  is of class  $C^3$ . Then, there are continuous functionals  $\mathfrak{g} : C^1(\partial\Omega) \rightarrow \mathbf{R}$  and  $\mathfrak{h} : C^1(\partial\Omega) \times C^1(\partial\Omega) \rightarrow \mathbf{R}$ , such that*

$$(74) \quad DJ(\Omega)(\mathbf{X}) = \mathfrak{g}(\mathbf{X}|_{\partial\Omega} \cdot \boldsymbol{\nu}) \quad \text{for all } \mathbf{X} \in \mathring{C}^1(\bar{\Omega}, \mathbf{R}^2)$$

and

$$(75) \quad \begin{aligned} \mathfrak{D}^2 J(\Omega)(\mathbf{X})(\mathbf{Y}) &= \mathfrak{h}(\mathbf{X}|_{\partial\Omega} \cdot \boldsymbol{\nu}, \mathbf{Y}|_{\partial\Omega} \cdot \boldsymbol{\nu}) - \mathfrak{g}(\boldsymbol{\nu} \cdot \partial^\tau \mathbf{X}_\tau \mathbf{Y}_\tau) \\ &\quad - \mathfrak{g}(\nabla^\tau(\mathbf{Y} \cdot \boldsymbol{\nu}) \cdot \mathbf{X}_\tau) - \mathfrak{g}(\nabla^\tau(\mathbf{X} \cdot \boldsymbol{\nu}) \cdot \mathbf{Y}_\tau) \end{aligned}$$

for all  $\mathbf{X}, \mathbf{Y} \in \mathring{C}^2(\bar{\Omega}, \mathbf{R}^2)$ .

*Remark 3.4.* The function  $\mathfrak{h}$  corresponds to a Banach space version of the Riemannian Hessian on shape spaces used in [27, Def. 2.2, p. 488].

*Remark 3.5.* Notice that (75) is equivalent to (see [22])

$$(76) \quad \mathfrak{D}^2 J(\Omega)(\mathbf{X})(\mathbf{Y}) = \mathfrak{h}(\mathbf{X}|_{\partial\Omega} \cdot \boldsymbol{\nu}, \mathbf{Y}|_{\partial\Omega} \cdot \boldsymbol{\nu}) - \mathfrak{g}(\boldsymbol{\nu} \cdot \partial^\tau \mathbf{Y} \mathbf{X}_\tau) - \mathfrak{g}(\mathbf{Y}_\tau \cdot \partial^\tau \boldsymbol{\nu} \mathbf{X}_\tau) - \mathfrak{g}(\boldsymbol{\nu} \cdot \partial^\tau \mathbf{Y} \mathbf{Y}_\tau).$$

**3.2. Shape Newton descent directions and their approximation.** In this section, we define the shape Newton equation and its approximation with weakly normal basis functions. For a particular test case, we derive quasi-optimality of the approximate Newton update and verify the resulting convergence rates with a numerical experiment. Henceforth,  $\Omega \subset \mathbf{R}^2$  is a set with  $C^{k+1}$ -boundary (with  $k \geq 1$ ), and  $\mathcal{H}(\mathbf{R}^2)$  is an RKHS that satisfies Assumption 2.8.

**DEFINITION 3.6.** *The  $\mathcal{H}(\mathbf{R}^2)$ -Newton descent direction at  $\Omega$  is the solution  $\mathbf{X}_\boldsymbol{\nu} \in [\mathcal{H}(\partial\Omega)]_\boldsymbol{\nu}$  of*

$$(77) \quad \mathfrak{D}^2 J(\Omega)(\mathbf{X}_\boldsymbol{\nu})(\mathbf{Y}) = -DJ(\Omega)(\mathbf{Y}) \quad \text{for all } \mathbf{Y} \in [\mathcal{H}(\partial\Omega)]_\boldsymbol{\nu}^2.$$

*Remark 3.7.* If  $J$  satisfies the assumptions of Theorem 3.3, and if the functions in  $\mathcal{H}(\partial\Omega)$  are sufficiently smooth (for instance,  $\mathcal{H}(\partial\Omega) \subset C^2(\partial\Omega)$ ), then formulas (74) and (75) satisfy

$$\mathfrak{D}^2 J(\Omega)(\mathbf{r}_x)(\mathbf{r}_y) = \mathfrak{h}(\mathbf{r}_x \cdot \boldsymbol{\nu}, \mathbf{r}_y \cdot \boldsymbol{\nu}) \quad \text{and} \quad DJ(\Omega)(\mathbf{r}_y) = \mathfrak{g}(\mathbf{r}_y \cdot \boldsymbol{\nu}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \partial\Omega,$$

because item (a) from Theorem 2.15 implies  $(\mathbf{r}_x)_\tau = (\mathbf{r}_y)_\tau = 0$ . This, in light of Theorem 2.21, implies that (77) is equivalent to

$$(78) \quad \mathfrak{h}(\mathbf{X}_\boldsymbol{\nu} \cdot \boldsymbol{\nu}, \mathbf{r}_y \cdot \boldsymbol{\nu}) = -\mathfrak{g}(\mathbf{r}_y \cdot \boldsymbol{\nu}) \quad \text{for all } \mathbf{y} \in \partial\Omega.$$

Since  $\mathfrak{h}$  and  $\mathfrak{g}$  are (bi)linear and continuous in  $C^1(\mathbf{R}^2)$ , (78) can also be used to define  $\mathcal{H}(\mathbf{R}^2)$ -Newton descent directions for RKHSs that satisfy only  $\mathcal{H}(\mathbf{R}^2) \subset C^1(\mathbf{R}^2)$ . For instance, as mentioned in Example 2.5, the native space of the kernel  $k_2^\sigma$  is  $\mathcal{H}(\mathbf{R}^2) = H^{2.5}(\mathbf{R}^2)$ , and  $H^{2.5}(\mathbf{R}^2) \subset C^{1,1/2}(\mathbf{R}^2)$  by the Sobolev embedding theorem (where  $C^{1,1/2}(\mathbf{R}^2)$  denotes the usual Hölder space).

In general, it can be difficult to compute an exact solution of (77) because  $\mathfrak{D}^2 J$  and  $DJ$  are typically given by very complicated formulas. For instance, when  $J$  is constrained by a PDE, the formula of  $\mathfrak{D}^2 J$  contains the material derivative of the solution of the PDE constraint [8]. Nevertheless, approximate  $\mathcal{H}(\mathbf{R}^2)$ -Newton descent directions can be computed by discretizing (77) with the weakly-normal basis functions  $\mathbf{r}_x$ s.

**DEFINITION 3.8.** Let  $\mathcal{X}_N = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \partial\Omega$  be a collection of pairwise distinct points, and let  $h = h_{\mathcal{X}_N, \partial\Omega}$  denote its fill distance. The approximate  $\mathcal{H}(\mathbf{R}^2)$ -Newton descent direction at  $\Omega$  with respect to  $\mathcal{H}(\partial\Omega)$  is the solution  $\mathbf{X}_\nu^h \in \mathcal{R}_N := \text{span}\{\mathbf{r}_x : \mathbf{x} \in \mathcal{X}_N\}$  of

$$(79) \quad \mathfrak{D}^2 J(\Omega)(\mathbf{X}_\nu^h)(\mathbf{Y}^h) = -DJ(\Omega)(\mathbf{Y}^h) \quad \text{for all } \mathbf{Y}^h \in \mathcal{R}_N.$$

**Example 3.9.** We consider the shape function  $J(\Omega) = \int_\Omega f \, dx$  with  $f \in C^2(\mathbf{R}^2)$ . This shape function  $J$  can be used as a regularization in image segmentation; see [19]. Its shape derivatives read [31]

$$(80) \quad DJ(\Omega)(\mathbf{X}) = \int_{\partial\Omega} f \mathbf{X} \cdot \boldsymbol{\nu} \, ds,$$

$$(81) \quad \begin{aligned} \mathfrak{D}^2 J(\Omega)(\mathbf{X})(\mathbf{Y}) &= \int_{\partial\Omega} (\nabla f \cdot \boldsymbol{\nu} + \kappa f)(\mathbf{X} \cdot \boldsymbol{\nu})(\mathbf{Y} \cdot \boldsymbol{\nu}) \, ds - \int_{\partial\Omega} f \boldsymbol{\nu} \cdot \partial^\tau \mathbf{X}_\tau \mathbf{Y}_\tau \, ds \\ &\quad - \int_{\partial\Omega} f \nabla^\tau (\mathbf{Y} \cdot \boldsymbol{\nu}) \cdot \mathbf{X} + f \nabla^\tau (\mathbf{X} \cdot \boldsymbol{\nu}) \cdot \mathbf{Y} \, ds, \end{aligned}$$

where  $\kappa$  denotes the curvature of  $\partial\Omega$ . Hence the  $\mathcal{H}(\mathbf{R}^2)$ -Newton descent direction at  $\Omega$  is the solution  $\mathbf{X}_\nu \in [\mathcal{H}(\partial\Omega)]_\nu$  of

$$(82) \quad \int_{\partial\Omega} (\nabla f \cdot \boldsymbol{\nu} + \kappa f)(\mathbf{X}_\nu \cdot \boldsymbol{\nu})(\mathbf{Y} \cdot \boldsymbol{\nu}) \, ds = - \int_{\partial\Omega} f(\mathbf{Y} \cdot \boldsymbol{\nu}) \, ds \quad \text{for all } \mathbf{Y} \in [\mathcal{H}(\partial\Omega)]_\nu.$$

Note that, when it exists, the solution of (82) is  $\mathbf{X}_\nu = (-f/(\nabla f \cdot \boldsymbol{\nu} + \kappa f))\boldsymbol{\nu}$ . Clearly, it is possible to approximate  $\mathbf{X}_\nu$  by interpolating it on  $\mathcal{R}_N$ . However, note that this interpolant would not satisfy (79). Finally, we point out that (80) implies that zero-level sets of  $f$  are local optima for this shape optimization problems. This implies that, close to the optimum, the term  $\kappa f$  in (82) plays a minor role compared to  $\nabla f \cdot \boldsymbol{\nu}$ . For this reason, some authors omit the term  $\kappa f$  in the computation of Newton updates [27, 31].

Proposition 3.10 shows quasi-optimality of the approximate shape Newton updates when the shape derivatives are as in Example 3.9.

**PROPOSITION 3.10.** Let  $J(\Omega) = \int_\Omega f \, dx$  with  $f \in C^2(\mathbf{R}^2)$ , and further assume that  $\nabla f \cdot \boldsymbol{\nu} + \kappa f > 0$  on  $\partial\Omega$ . Then, the solution  $\mathbf{X}_\nu^h \in \mathcal{R}_N$  of (79) satisfies

$$(83) \quad \|\mathbf{X}_\nu - \mathbf{X}_\nu^h\|_{L_2(\partial\Omega)} \leq c_{f,\Omega} \inf_{\mathbf{Y} \in \mathcal{R}_N} \|\mathbf{X}_\nu - \mathbf{Y}\|_{L_2(\partial\Omega)}.$$

*Proof.* The bilinear form

$$(84) \quad L_2(\partial\Omega) \times L_2(\partial\Omega) \rightarrow \mathbb{R}, \quad (x, y) \mapsto \int_{\partial\Omega} (\nabla f \cdot \boldsymbol{\nu} + \kappa f) xy \, ds$$

is  $L_2(\partial\Omega)$ -continuous, symmetric, and  $L_2(\partial\Omega)$ -elliptic. Let  $\mathcal{Z}_N := \text{span}\{\mathbf{X} \cdot \boldsymbol{\nu} : \mathbf{X} \in \mathcal{R}_N\}$ . Clearly,  $\mathcal{Z}_N \subset L_2(\partial\Omega)$ . Therefore, the solution  $z_h \in \mathcal{Z}_N$  of

$$(85) \quad \int_{\partial\Omega} (\nabla f \cdot \boldsymbol{\nu} + \kappa f) z_h y \, ds = - \int_{\partial\Omega} f y \, ds \quad \text{for all } y \in \mathcal{Z}_N$$

satisfies

$$(86) \quad \|z - z_h\|_{L_2(\partial\Omega)} \leq c_{f,\Omega} \inf_{y \in \mathcal{Z}_N} \|z - y\|_{L_2(\partial\Omega)},$$

where  $z \in L_2(\partial\Omega)$  is the solution of

$$(87) \quad \int_{\partial\Omega} (\nabla f \cdot \nu + \kappa f) zy \, ds = - \int_{\partial\Omega} fy \, ds \quad \text{for all } y \in L_2(\partial\Omega).$$

Finally,  $\mathbf{X}_\nu^h = z_h \nu$  because  $\mathcal{R}_N$  satisfies  $\mathcal{R}_N = \{z\nu : z \in \mathcal{Z}_N\}$ , and  $\mathbf{X}_\nu = z\nu$  because the solution of (87) is  $z = -f/(\nabla f \cdot \nu + \kappa f)$ .  $\square$

**NUMERICAL EXPERIMENT 3.11.** *We perform a numerical experiment to investigate the approximation error (83) for  $f(\mathbf{x}) = |\mathbf{x}|^2 - 1$ . We choose  $\partial\Omega$  to be the ellipse depicted in Figure 1 (left) and consider the sequence of subsets*

$$(88) \quad \mathcal{X}_N = \{\gamma(2\pi\ell/N) : \ell = 1, \dots, N\} \quad \text{with} \quad N = 2^4, 2^5, \dots, 2^8.$$

For each  $N$ , we construct the finite dimensional space  $\tilde{\mathcal{R}}_N := \text{span}\{\mathbf{r}_x^{h_N} : \mathbf{x} \in \mathcal{X}_N\}$  (where  $h_N$  denotes the fill distance of  $\mathcal{X}_N$ ) and compute the solution  $\tilde{\mathbf{X}}_\nu^{h_N} \in \tilde{\mathcal{R}}_N$  of

$$(89) \quad \int_{\partial\Omega} (\nabla f \cdot \nu + \kappa f)(\tilde{\mathbf{X}}_\nu^{h_N} \cdot \nu)(\mathbf{Y}^{h_N} \cdot \nu) \, ds = - \int_{\partial\Omega} f(\mathbf{Y}^{h_N} \cdot \nu) \, ds \quad \text{for all } \mathbf{Y}^{h_N} \in \tilde{\mathcal{R}}_N.$$

Although the normal vector field  $\nu$  and curvature  $\kappa$  can be computed exactly, for ease of implementation we replace the curve  $\gamma$  with a highly accurate interpolant (using Chebfun [10]) and evaluate  $\nu$  and  $\kappa$  of this interpolant. For a result on this kind of approximation, see [32]. Finally, we compute (with sufficiently many quadrature points) the  $L_2(\partial\Omega)$ -absolute error  $\|\mathbf{X}_\nu - \tilde{\mathbf{X}}_\nu^{h_N}\|_{L_2(\partial\Omega)}$ . In Figure 3, we plot this  $L_2$ -errors versus  $h_N$  (measured with the Euclidean distance) for the kernels  $k_4^{0.7}$ ,  $k_6^{0.7}$ , and  $k_8^{0.7}$ . The measured convergence rates read (approximately) 4.20, 5.98, and 8.20 (when  $h_N > 2^{-4}$ ), respectively. Comparable results are obtained for the nonconvex domain depicted in Figure 1 (left), as well as for different values of  $\sigma$ . To interpret these results, let us first point out that  $\tilde{\mathcal{R}}_N$  is only an approximation of  $\mathcal{R}_N$ . It may not be so simple to prove a proposition for  $\tilde{\mathcal{R}}_N$  that is analogous to Proposition 3.10, because  $\tilde{\mathcal{R}}_N$  does not contain the image of the operator  $\mathbf{X} \mapsto (\mathbf{X} \cdot \nu)\nu$ . Note also that it is true that  $\mathbf{r}_x^{h_N}$  converges to  $\mathbf{r}_x$  as  $h_N \rightarrow 0$ , but Theorem 2.33 does not guarantee that this convergence is uniform in  $x$ . However, if we assume that a proposition analogous to Proposition 3.10 exists for  $\tilde{\mathcal{R}}_N$ , we can explain the measured convergence rates with [17, Cor. 15], because interpolating a vector field  $\mathbf{f}$  into  $\tilde{\mathcal{R}}_N$  or into  $\mathcal{H}(\mathcal{X}_N) \times \mathcal{H}(\mathcal{X}_N)$  returns exactly the same interpolant if the vector field  $\mathbf{f}$  is normal to  $\partial\Omega$ .

Finally, the convergence history stagnates for  $k_8^{0.7}$  because the condition number of the discretized shape Hessian is very large for  $N = 2^8$  (the MATLAB function cond returns the value 4.826490e+18). In a simple numerical experiment, we observed that the condition number of the discretized shape Hessian behaves as  $h_N^{\alpha_h} \sigma^{\alpha_\sigma}$  with  $\alpha_h = -8.7$  and  $\alpha_\sigma = 7.2$  when  $p = 4$ , and with  $\alpha_h = -12.48$  and  $\alpha_\sigma = 10.18$  when  $p = 6$ . We believe that these condition numbers deteriorate because we rely on radial basis functions. Indeed, following closely [6, sect. 5.2], it is possible to bound from above the condition number of a mass matrix that arises from a discretization based on compactly supported Wendland kernels, and this upper bound reads  $Ch_N^{-2p}$ , where  $C$  is a positive constant.

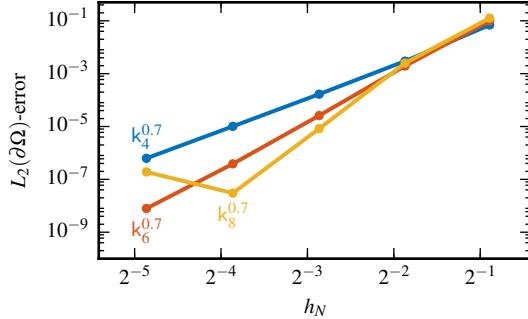


FIG. 3. Numerical Experiment 3.11 shows that  $\|\mathbf{X}_\nu - \tilde{\mathbf{X}}_\nu^{h_N}\|_{L_2(\partial\Omega)}$  decays with the algebraic rates 4.20 (for  $k_4^{0.7}$ ), 5.98 (for  $k_6^{0.7}$ ), and 8.20 (for  $k_8^{0.7}$  when  $h_N > 2^{-4}$ ).

**3.3. Discrete shape Newton algorithm based on  $\mathbf{r}_x$ .** In this section, we use the basis functions  $\mathbf{r}_x$  to formulate a discrete shape Newton algorithm and test it on Example 3.9. In particular, we investigate the impact of the Newton update approximation on the convergence rates of the shape Newton method. To measure these convergence rates, we use the norm on the tangent spaces, which is simply the norm of the RKHS.

Let  $J$  be a twice differentiable shape function, let  $\Omega \subset \mathbf{R}^2$  be an initial guess, and let  $\mathcal{X}_N = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$  be a finite number of points on  $\partial\Omega$ . We begin by specifying how to transport the basis functions  $\mathbf{r}_x$  when  $\Omega$  is perturbed by a geometric transformation.

**DEFINITION 3.12.** Let  $\mathbf{r}_{\partial\Omega, \mathbf{x}}$  denote the weakly normal function associated with  $\partial\Omega$  and  $\mathbf{x} \in \partial\Omega$ , and let  $F_t : \mathbf{R}^2 \rightarrow \mathbf{R}^2$  be a family of  $C^1$ -diffeomorphisms such that  $F_0 = \text{id}$ . We define the transport  $\mathfrak{T}_s$  by

$$(90) \quad \mathfrak{T}_s(\mathbf{r}_{\partial\Omega, \mathbf{x}}) := \mathbf{r}_{F_s(\partial\Omega), F_s(\mathbf{x})}$$

and extend it by linearity to  $\mathfrak{T}_s : \text{span}\{\mathbf{r}_{\partial\Omega, \mathbf{x}}, \mathbf{x} \in \partial\Omega\} \rightarrow \text{span}\{\mathfrak{T}_s(\mathbf{r}_{\partial\Omega, \mathbf{x}}), \mathbf{x} \in \partial\Omega\}$ . By construction,  $\mathfrak{T}_s : \text{span}\{\mathbf{r}_{\partial\Omega, \mathbf{x}}, \mathbf{x} \in \partial\Omega\} \subset [\mathcal{H}(\partial\Omega)]_\nu^2 \rightarrow [\mathcal{H}(F_t(\partial\Omega))]_\nu^2$ .

Using this transport, we formulate the following discrete shape Newton method. Set  $\Omega_0 = \Omega$ , and  $\tilde{\mathcal{R}}_N^0 := \text{span}\{\mathbf{r}_{\partial\Omega, \mathbf{x}}^h : \mathbf{x} \in \mathcal{X}_N\}$ . The optimization algorithm proceeds as follows: first, it computes the solution  $\mathbf{g}_0 \in \tilde{\mathcal{R}}_N^0$  of

$$(91) \quad \mathfrak{D}^2 J(\Omega_0)(\mathbf{g}_0)(\varphi) = -D J(\Omega_0)(\varphi) \quad \text{for all } \varphi \in \tilde{\mathcal{R}}_N^0;$$

then, it defines the transformation  $F_0(\mathbf{x}) := \mathbf{x} + \mathbf{g}_0(\mathbf{x})$  and the new domain  $\Omega_1 := F_0(\Omega_0)$ ; finally, it defines the space  $\tilde{\mathcal{R}}_N^1$  by transporting  $\tilde{\mathcal{R}}_N^0$  with  $F_0$ , that is,

$$(92) \quad \tilde{\mathcal{R}}_N^1 := \text{span}\{\mathbf{r}_{F_0(\partial\Omega), F_0(\mathbf{x})}^h : \mathbf{x} \in \mathcal{X}_N\}.$$

This procedure is repeated until convergence.

**Remark 3.13.** The shape  $\Omega_1$  that results by updating  $\Omega_0$  with  $\mathbf{g}_0$  has the same regularity of  $\mathbf{g}_0$ . Therefore, from a theoretical point of view, it may be necessary to regularize the update  $\mathbf{g}_0$  to guarantee that Assumption 2.8 remains fulfilled during the optimization process. However, such a regularization is not necessary in practice when one employs the Wendland kernels from Example 2.5. In this case, regularity

is lost only at the points  $\{\mathbf{x} \pm \sigma : \mathbf{x} \in \mathcal{X}_N\}$ , whereas to construct the approximate weakly normal function  $\mathbf{r}_{\mathbf{x}}^h$ s it is necessary to evaluate  $\boldsymbol{\nu}$  and  $\boldsymbol{\tau}$  only at  $\mathcal{X}_N$ . Therefore, if  $\{\mathbf{x} \pm \sigma : \mathbf{x} \in \mathcal{X}_N\} \cap \mathcal{X}_N = \emptyset$ , the discrete algorithm does not perceive the loss of regularity of the domain.

NUMERICAL EXPERIMENT 3.14. *We test this discrete shape Newton method by performing five optimization steps to solve Example 3.9 with  $f(\mathbf{x}) = |\mathbf{x}|^2 - 1$  and*

$$(93) \quad \Omega = \{\gamma(\phi) = (0.2 + 1.15 \cos(\phi), 0.15 + 0.9 \sin(\phi)) : \phi \in [0, 2\pi)\}.$$

*The (approximate) weakly normal basis functions  $\mathbf{r}_{\mathbf{x}}^h$  are constructed with the kernel  $k_4^{0.7}$  on the sets*

$$(94) \quad \mathcal{X}_N = \{\gamma(2\pi\ell/N) : \ell = 1, \dots, N\} \quad \text{with} \quad N = 2^4, 2^5, \dots, 2^8.$$

*To simplify the implementation, the shape iterates  $\{\Omega_{\ell+1} = \Omega_\ell + \mathbf{g}_\ell(\Omega_\ell)\}_{\ell=0}^4$  are replaced by highly accurate interpolants computed with Chebfun [10]. Normal and tangential vector fields  $\boldsymbol{\nu}$  and  $\boldsymbol{\tau}$ , as well as the curvature term  $\kappa$ , are approximated by evaluating their corresponding analytic formulas with these interpolants.*

To replicate the results from [27, 31], we neglect the curvature term  $\kappa f$  in (82) and compute the corresponding Newton descent direction. Note that a counterpart of Proposition 3.10 can be proved provided that  $\nabla f \cdot \boldsymbol{\nu} > 0$  (which is the case in these experiments). In Figure 4, we plot the evolution of the shape iterates  $\{\Omega_\ell\}_{\ell=0}^2$  for  $N = 2^4$  (top left) and for  $N = 2^5$  (top right), and for each  $N$  we plot the values of the sequence  $\{\|\mathbf{g}_\ell\|_{[\mathcal{H}]^2}\}_{\ell=0}^4$  (bottom left) and the measure of the quadratic rate of convergence  $\{\|\mathbf{g}_{\ell+1}\|_{[\mathcal{H}]^2}/\|\mathbf{g}_\ell\|_{[\mathcal{H}]^2}^2\}_{\ell=0}^3$  (bottom right).

We observe that the sequence of shapes converges quickly to the minimizer (a circle of radius 1 centered in the origin) and that the discretization error on the retrieved shape is not visible for  $N \geq 2^5$ . We also observe that the discretized shape Newton method converges quadratically if sufficiently many weakly normal vector fields are employed.

Remark 3.15. We have repeated the previous experiment, this time by taking the curvature term  $\kappa f$  into account. Surprisingly, we observed that the quadratic rate of convergence deteriorates to superlinear although the sequence of shape iterates converge to the correct minimizer. To further investigate this unexpected behavior, we consider a simpler scenario that can be solved analytically. Let the initial domain  $\Omega_0$  be a disc of radius  $r_0$  centered in the origin. In this case, the shape iterates are concentric discs of radii,

$$(95) \quad r_{\ell+1} = r_\ell - \frac{r_\ell^2 - 1}{2r_\ell + \kappa_\ell(r_\ell^2 - 1)}, \quad \ell = 0, 1, \dots,$$

where  $\kappa_\ell = r_\ell^{-1}$  is the curvature of the  $\ell$ th disc. Equation (95) can be rewritten as

$$(96) \quad r_{\ell+1} = \frac{2r_\ell^3}{3r_\ell^2 - 1}.$$

From this equation, we see that if  $r_\ell < 1/\sqrt{3}$ , then  $r_{\ell+1} < 0$ , which is not feasible. On the other hand, if  $r_\ell > 1/\sqrt{3}$ , then  $r_{\ell+1} > 1$ . Moreover,  $r_{\ell+1} < r_\ell$  if  $r_\ell > 1$ , which implies that  $(r_\ell)$  decreases monotonically and converges to one. By direct computation, we can show that

$$(97) \quad |1 - r_{\ell+1}| = \left| \frac{(2r_\ell + 1)(r_\ell - 1)}{3r_\ell^2 - 1} \right| |1 - r_\ell|,$$

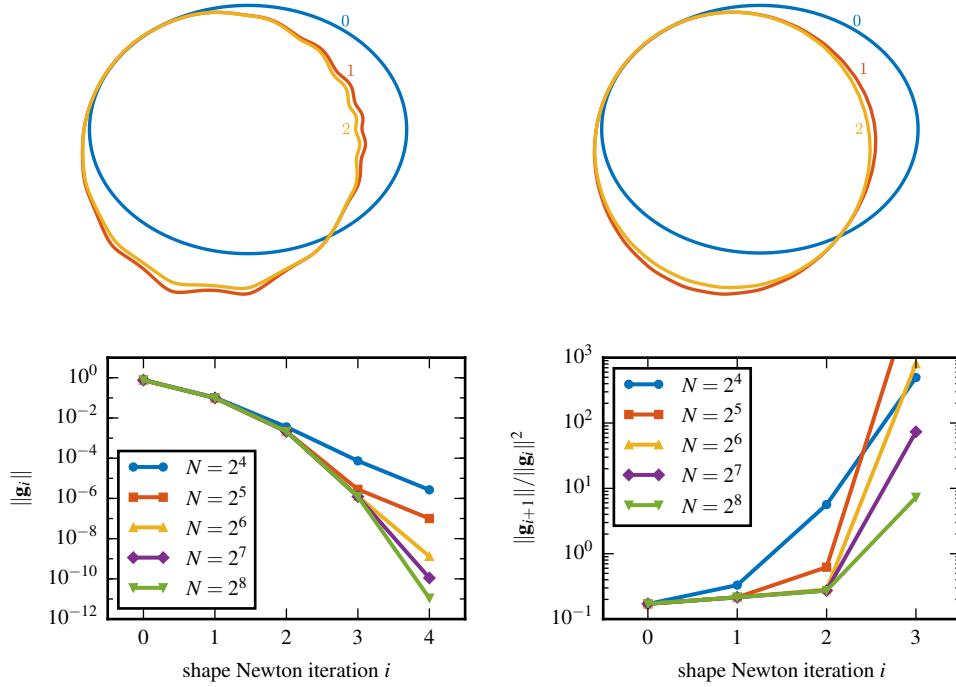


FIG. 4. Numerical Experiment 3.14 shows that the discretized shape Newton method converges quadratically if sufficiently many weakly normal vector fields are employed.

which implies that the error is decreasing only if  $r_\ell > \frac{1+\sqrt{41}}{10} > \frac{1+\sqrt{36}}{10} = 0.7$ . Since  $r_\ell > 1/\sqrt{3}$  implies  $r_{\ell+1} > 1$ , the sequence  $\{r_\ell\}$  converges quadratically for every initial radius  $r_0 > 1/\sqrt{3}$  after the first iteration. However, we obtain a better algorithm if we set  $\kappa_\ell = 0$ , because equation (95) becomes

$$r_{\ell+1} = \frac{1}{2} \left( r_\ell + \frac{1}{r_\ell} \right),$$

which converges quadratically for any  $r_0 > 0$ . To be more precise, in this case

$$(98) \quad |r_{\ell+1} - 1| = \left| \frac{r_\ell - 1}{2r_\ell} \right| |r_\ell - 1|,$$

which converges faster than (97) because  $\frac{2r_\ell+1}{3r_\ell^2-1} > \frac{1}{2r_\ell}$  for  $r_\ell > \frac{1}{\sqrt{3}}$ . This example shows that the denominator in (95) plays a key role in the convergence analysis, and it may be for this reason that the shape Newton algorithm without a curvature term performed better in our previous numerical experiment.

NUMERICAL EXPERIMENT 3.16. *To further highlight the effect of the discretization parameter  $N$  on approximate shape Newton methods, we consider the shape optimization problem from Example 3.9 with  $f(\mathbf{x}) = f(x, y) = (x^2 + 16y^2 - 1)(x^2 + (y - 0.45)^2 - 0.04)$ . The zero-level set of this function is the union of a disc and an ellipse, and it has a cusp at  $(0, 0.25)$  (see Figure 5). Therefore, the optimal shape is not*

even Lipschitz, and retrieving it by solving the optimization problem is particularly challenging.

We perform 15 discrete shape Newton steps with the same numerical setup of Numerical Experiment 3.14. In Figure 5, we plot the evolution of shape iterates  $\{\Omega_\ell\}_{\ell=0}^{15}$  for  $N = 2^6$  (left) and for  $N = 2^8$  (right). These plots clearly show that increasing  $N$  allows retrieving a much better approximation of the optimal shape. Note that, after 15 steps, the algorithm has fully converged to an approximate optimum.

Note that in the computation of these shape iterates, we have neglected the curvature term  $\kappa f$  in (82). Similarly to Numerical Experiment 3.14, we have observed that the performance of the shape Newton method deteriorates if the curvature term is taken into account. In particular, it is necessary both to regularize the Newton equation to enforce coercivity [19] and to switch to the damped Newton method to ensure that every shape iterate is feasible.

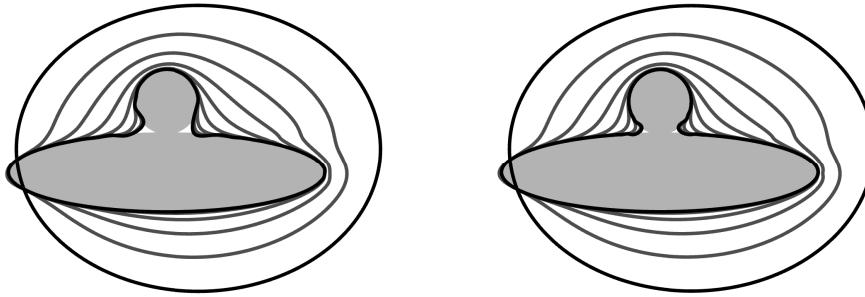


FIG. 5. Numerical Experiment 3.16 shows that increasing the discretization parameter  $N$  allows us to retrieve a much better approximation of the optimal shape. In these plots, black lines denote the initial and the final shape, lighter lines denote intermediate shape iterates, and the gray area denotes the zero-level set of  $f$ .

**4. Conclusion.** In sections 1 and 2, we introduced the class of weakly normal vector fields, which are defined as solutions of saddle point variational problems in RKHSs. Besides investigating their properties, we have discussed their approximation and proved the related convergence rates. These vector fields can be used to discretize shape Newton methods because they have purely normal components.

In section 3, we formulated a discrete shape Newton method based on these weakly normal vector fields and, for a specific test case, proved that the discrete shape Newton direction converges to the continuous one in the metric associated with the shape Hessian. Finally, we showed that the discrete shape Newton method exhibits quadratic convergence if the discretization error is sufficiently small.

We have observed that our approach leads to matrices whose condition numbers deteriorate as the number of degrees of freedom increases. This is a common issue of discretizations based on radial basis functions. In [6], the authors describe a multilevel residual correction algorithm that leads to approximation matrices with bounded condition numbers. Although there is debate on its convergence properties [34], this algorithm performed well in numerical experiments, and it would be interesting to investigate whether a similar technique could be used in our context.

The weakly normal vector fields have another property that has not been exploited in this work: although defined on the boundary  $\partial\Omega$ , they can be straightforwardly

extended to  $\mathbf{R}^2$ . This implies that the discrete shape Newton method remains well-posed even when formulated with volume-based expressions of the shape derivative and the shape Hessian. This is a great advantage for PDE-constrained shape optimization problems, in which case volume-based formulas are easier to derive and impose fewer regularity requirements on the solution of the PDE constraint. However, using volume-based formulas requires the integration of the weakly normal vector fields on  $\Omega$ , which introduces an additional complexity in the implementation as well as additional sources of discretization error that need to be analyzed. For this reason, we postpone to a subsequent work the study of the shape Newton method based on weakly normal vector fields in the framework of PDE-constrained shape functionals.

Finally, we mention that certain shape optimization problems present an additional difficulty: their second-order shape derivative is continuous on a function space and coercive on a weaker one (this is commonly known as the “norm gap”). When this is the case, the shape optimization problem under consideration is not well-posed [15]. Although discretization is a form of regularization, in the case of an ill-posed problem, one should pay additional attention to the discretization level. The weakly normal vector fields introduced in this work may carry extra regularization properties because their smoothness can be tuned by choosing appropriate reproducing kernels. We defer the investigation of this interesting aspect to future work.

**Appendix A.** In this appendix, we show that  $\mathcal{H}(\partial\Omega) \subset C^0(\partial\Omega)$  when  $k \in C^0(\mathbf{R}^2 \times \mathbf{R}^2)$ . We believe that this result is not new, but we could not find its proof in the literature.

**LEMMA A.1.** *Let  $k$  be a symmetric positive-definite kernel on  $\mathbf{R}^2$ . Denote by  $\mathcal{H}(\partial\Omega)$  the RKHS associated with the restriction of  $k$  to  $\partial\Omega$ . If  $k \in C^0(\mathbf{R}^2 \times \mathbf{R}^2)$ , then  $\mathcal{H}(\partial\Omega) \subset C^0(\partial\Omega)$ .*

*Proof.* For a generic  $f \in \mathcal{H}(\partial\Omega)$ , the reproducing kernel properties of  $\mathcal{H}(\partial\Omega)$  imply

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{y})| &= |(k(\mathbf{x}, \cdot), f)_\mathcal{H} - (k(\mathbf{y}, \cdot), f)_\mathcal{H}| \\ (99) \quad &= |(k(\mathbf{x}, \cdot) - k(\mathbf{y}, \cdot), f)_\mathcal{H}| \\ &\leq \|k(\mathbf{x}, \cdot) - k(\mathbf{y}, \cdot)\|_\mathcal{H} \|f\|_\mathcal{H} \end{aligned}$$

for all  $\mathbf{x}, \mathbf{y} \in \partial\Omega$ . Moreover,

$$(100) \quad \|k(\mathbf{x}, \cdot) - k(\mathbf{y}, \cdot)\|_\mathcal{H}^2 = k(\mathbf{x}, \mathbf{x}) + k(\mathbf{y}, \mathbf{y}) - 2k(\mathbf{x}, \mathbf{y}).$$

Therefore, the right-hand side of (100) goes to zero as  $\mathbf{x}$  goes to  $\mathbf{y}$  because  $k \in C^0(\mathbf{R}^2 \times \mathbf{R}^2)$ . In view of estimates (99) and (65), this shows that  $f$  is continuous on  $\partial\Omega$ .  $\square$

## REFERENCES

- [1] G. ALLAIRE, E. CANCEÙS, AND J.-L. VIÉ, *Second-order shape derivatives along normal trajectories, governed by Hamilton-Jacobi equations*, Struct. Multidiscip. Optim., 54 (2016), pp. 1245–1266.
- [2] G. ALLAIRE, C. DAPOGNY, AND P. FREY, *Shape optimization with a level set based mesh evolution method*, Comput. Methods Appl. Mech. Engrg., 282 (2014), pp. 22–53.
- [3] H. AMANN AND J. ESCHER, *Analysis II*, Birkhäuser, Basel, 2006.
- [4] K. BANDARA, F. CIRAK, G. OF, O. STEINBACH, AND J. ZAPLETAL, *Boundary element based multiresolution shape optimisation in electrostatics*, J. Comput. Phys., 297 (2015), pp. 584–598.

- [5] D. BOFFI, F. BREZZI, AND M. FORTIN, *Mixed Finite Element Methods and Applications*, Springer, Heidelberg, 2013.
- [6] A. CHERNIH AND Q. LE GIA, *Multiscale methods with compactly supported radial basis functions for Galerkin approximation of elliptic PDEs*, IMA J. Numer. Anal., 34 (2014), pp. 569–591.
- [7] C. DAPOGNY, R. ESTEVEZ, A. FAURE, AND G. MICHAELIDIS, *Shape and Topology Optimization Considering Anisotropic Features Induced by Additive Manufacturing Processes*, working paper, <https://hal.archives-ouvertes.fr/hal-01660850>, 2018.
- [8] M. C. DELFOUR AND J.-P. ZOLÉSIO, *Velocity method and Lagrangian formulation for the computation of the shape Hessian*, SIAM J. Control Optim., 29 (1991), pp. 1414–1442, <https://doi.org/10.1137/0329072>.
- [9] M. C. DELFOUR AND J.-P. ZOLÉSIO, *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*, 2nd ed., Adv. Des. Control 22, SIAM, Philadelphia, 2011, <https://doi.org/10.1137/1.9780898719826>.
- [10] T. A. DRISCOLL, N. HALE, AND L. N. TREFETHEN, EDS., *Chebfun Guide*, Pafnuty Publications, Oxford, 2014, <https://www.chebfun.org/docs/guide/>.
- [11] M. EIGEL AND K. STURM, *Reproducing kernel Hilbert spaces and variable metric algorithms in PDE-constrained shape optimization*, Optim. Methods Software, 33 (2018), pp. 268–296, <https://doi.org/10.1080/10556788.2017.1314471>.
- [12] I. EKELAND AND R. TÉMAM, *Convex Analysis and Variational Problems*, Classics Appl. Math. 28, SIAM, Philadelphia, 1999.
- [13] K. EPPLER, *Second derivatives and sufficient optimality conditions for shape functionals*, Control Cybernet., 29 (2000), pp. 485–511.
- [14] K. EPPLER AND H. HARBRECHT, *A regularized Newton method in electrical impedance tomography using shape Hessian information*, Control Cybernet., 34 (2005), pp. 203–225.
- [15] K. EPPLER AND H. HARBRECHT, *Shape optimization for free boundary problems—analysis and numerics*, in Constrained Optimization and Optimal Control for Partial Differential Equations, Internat. Ser. Numer. Math. 160, Birkhäuser, Basel, 2012, pp. 277–288.
- [16] K. EPPLER, H. HARBRECHT, AND R. SCHNEIDER, *On convergence in elliptic shape optimization*, SIAM J. Control Optim., 46 (2007), pp. 61–83, <https://doi.org/10.1137/05062679X>.
- [17] E. FUSELIER AND G. B. WRIGHT, *Scattered data interpolation on embedded submanifolds with restricted positive definite kernels: Sobolev error estimates*, SIAM J. Numer. Anal., 50 (2012), pp. 1753–1776, <https://doi.org/10.1137/110821846>.
- [18] A. HENROT AND M. PIERRE, *Variation et optimisation de formes. Une analyse géométrique*, Springer, Berlin, 2005.
- [19] M. HINTERMÜLLER AND W. RING, *A second order shape optimization approach for image segmentation*, SIAM J. Appl. Math., 64 (2003), pp. 442–467, <https://doi.org/10.1137/S0036139902403901>.
- [20] M. KEUTHEN AND D. KRAFT, *Shape optimization of a breakwater*, Inverse Probl. Sci. Eng., 24 (2016), pp. 936–956.
- [21] A. LAURAIN AND K. STURM, *Distributed shape derivative via averaged adjoint method and applications*, ESAIM M2AN, 50 (2016), pp. 1241–1267.
- [22] A. NOVRUZI AND M. PIERRE, *Structure of shape derivatives*, J. Evol. Equ., 2 (2002), pp. 365–382.
- [23] A. PAGANINI, S. SARGHEINI, R. HIPTMAIR, AND C. HAFNER, *Shape optimization of microlenses*, Opt. Express, 23 (2015), pp. 13099–13107.
- [24] C. C. D. RONCO, R. PONZA, AND E. BENINI, *Aerodynamic shape optimization of aircraft components using an advanced multi-objective evolutionary approach*, Comput. Methods Appl. Mech. Engrg., 285 (2015), pp. 255–290.
- [25] R. SCHABACK, *Error estimates and condition numbers for radial basis function interpolation*, Adv. Comput. Math., 3 (1995), pp. 251–264.
- [26] A. SCHIELA AND J. ORTIZ, *Second Order Directional Shape Derivatives*, preprint 3251, EPub Bayreuth, Universität Bayreuth, <https://epub.uni-bayreuth.de/id/eprint/3251>, 2017.
- [27] V. H. SCHULZ, *A Riemannian view on shape optimization*, Found. Comput. Math., 14 (2014), pp. 483–501.
- [28] V. H. SCHULZ, M. SIEBENBORN, AND K. WELKER, *PDE constrained shape optimization as optimization on shape manifolds*, in Geometric Science of Information, Lecture Notes in Comput. Sci. 9389, Springer, Cham, 2015, pp. 499–508.
- [29] V. H. SCHULZ, M. SIEBENBORN, AND K. WELKER, *Structured inverse modeling in parabolic diffusion problems*, SIAM J. Control Optim., 53 (2015), pp. 3319–3338, <https://doi.org/10.1137/140985883>.

- [30] J. SOKOŁOWSKI AND J.-P. ZOLÉSIO, *Introduction to Shape Optimization. Shape Sensitivity Analysis*, Springer, Berlin, 1992.
- [31] K. STURM, *Convergence Analysis of Newton's Method in Shape Optimisation via Approximate Normal Functions*, preprint, <https://arxiv.org/abs/1608.02699>, 2016.
- [32] Y. N. SUBBOTIN, *Uniform approximation of curvature for smooth classes of plane curves*, Proc. Steklov Inst. Math., 297 (2017), pp. 182–185.
- [33] H. WENDLAND, *Scattered Data Approximation*, Cambridge University Press, Cambridge, UK, 2005.
- [34] H. WENDLAND, *Solving partial differential equations with multiscale radial basis functions*, in *Contemporary Computational Mathematics—A Celebration of the 80th Birthday of Ian Sloan*, J. Dick, F. Kuo, and H. Woźniakowski, eds., Springer, Cham, 2018, pp. 1191–1213, [https://doi.org/10.1007/978-3-319-72456-0\\_55](https://doi.org/10.1007/978-3-319-72456-0_55).