SIAM J. NUMER. ANAL.
Vol. 45, No. 1, pp. 333–351

CONVERGENCE OF UNSYMMETRIC KERNEL-BASED MESHLESS COLLOCATION METHODS*

ROBERT SCHABACK†

Abstract. This paper proves convergence of variations of the unsymmetric kernel-based collocation method introduced by Kansa in 1986. Since then, this method has been very successfully used in many applications, though it may theoretically fail in special situations, and though it had no error bound or convergence proof up to now. Thus it is necessary to add assumptions or to make modifications. Our modifications prevent numerical failure by dropping strict collocation and allow a rigorous mathematical analysis proving error bounds and convergence rates. These rates improve with the smoothness of the solution, the domain, and the kernel providing the trial spaces, but they are currently not yet optimal and deserve refinement. They are based on rates of approximation to the residuals by nonstationary meshless kernel-based trial spaces, and they are independent of the type of differential operator. The results are applicable to large classes of linear problems in strong form, provided that there is a smooth solution and the test and trial discretizations are chosen with some care. Our analysis does not require assumptions like ellipticity, and it can be extended to ill-posed problems.

Key words. Kansa method, error bounds, stability, ill-posed problems, greedy adaptive solver

AMS subject classifications. 65N12, 65N35, 65N22, 65F22

DOI. 10.1137/050633366

1. Introduction.

The final goal of this paper is to prove error bounds and convergence of certain numerical techniques that approximately solve a PDE problem via an unsymmetric or even nonsquare system of linear collocation equations involving meshless kernel-based trial functions. The most popular method of this kind was first proposed by Kansa [8] in 1986, and there are many follow-up papers in engineering journals (see, e.g., [5] for a selection) that can easily be retrieved via the Internet. This is why this paper does not supply additional numerical examples. So far, the method is quite successful in applications with smooth solutions, but it can fail [7] in specially constructed situations. Consequently, it has neither error bounds nor convergence proofs for its original form, and a rigorous mathematical analysis will either require some additional assumptions or make changes to the method itself. We shall do both, but we shall stay general enough not to spoil the applicability to elliptic, parabolic, and hyperbolic problems. Therefore we need a somewhat nonstandard framework, which we sketch here first, to make sure that the reader does not get lost in the technical details we have to provide later.

Consider a linear operator equation

$$L(u) = f,\ L\ :\ U \to F \tag{1.1}$$

between normed linear spaces U and F which is to be solved for any given $f \in F$. The map L takes a *solution* $u \in U$ to its *data* $L(u)$ in F. Thus F will usually be a Cartesian product of *trace* spaces of functions prescribed on the domain or on parts of the boundary. We shall consider a large class of unsymmetric discretization methods to solve such equations approximately, and we need five essential ingredients.

*Received by the editors June 10, 2005; accepted for publication (in revised form) August 18, 2006; published electronically February 9, 2007.
http://www.siam.org/journals/sinum/45-1/63336.html

†Institut für Numerische und Angewandte Mathematik, Lotzestraße 16–18, D–37083 Göttingen, Germany (schaback@math.uni-goettingen.de, http://www.num.math.uni-goettingen.de/schaback).

The first ingredient requires the problem to be *continuously dependent on the data*. In quantitative form, this means that the inverse L^{-1} is a bounded linear map from F to U. In particular, we assume an inequality of the form

$$\|u\|_U \le C_a \|L(u)\|_F \text{ for all } u \in U \tag{1.2}$$

with a positive constant C_a describing the stability of the problem. In practical applications this will imply that the *solution space* U and the *data space* F have to be chosen with some care. In particular, U and F must often be chosen on a theoretical basis, e.g., as quite large spaces in which certain general existence results hold and which carry only rather weak norms. Usually, U will be a Sobolev space $W_2^\mu(\Omega)$ while F is a Cartesian product of Sobolev spaces which provide the right-hand sides for the differential equation and the boundary data via trace operators. Continuous dependence serves here as a replacement for more specific analytic assumptions like coercivity of a bilinear form. However, in section 9 we shall abandon the assumption of continuous dependence to be able to treat a certain class of ill-posed problems.

The second ingredient is some *additional regularity*. The actual solution u of a specific problem will often have more regularity than needed for the spaces U and F defining continuous dependence, and therefore we shall focus on a subspace $U_R \subseteq W_2^m(\Omega) \subset W_2^\mu(\Omega) =: U$ of U which we call the *regularity* subspace. The additional regularity of order $m - \mu > 0$ will be the driving force behind convergence rates, as we shall prove later.

The third ingredient is a scale of finite-dimensional *trial* subspaces U_r of U for a *trial discretization parameter* $r > 0$ which uses the additional regularity to provide a *convergent scheme for data approximation*. This is formalized by not necessarily linear maps $I_r : U_R \to U_r$ with error bounds

$$\|L(u - I_r(u))\|_F \le \epsilon_r(u) \text{ for all } u \in U_R. \tag{1.3}$$

It will be this approximation property that yields our convergence rates driven by additional regularity. Note that we do not use a single discretization parameter like the usual h here, because we need two different discretization parameters r and s for trial and test discretization. The trial spaces U_r can be chosen independent of the operator L, and we do not approximate the solution directly but rather the data via the linear operator L.

The fourth ingredient is a scale of *stable test discretizations* F_s of the data space F with respect to the scale of trial spaces U_r. This is formalized by a *test discretization parameter* $s > 0$ and linear maps $\Pi_s : F \to F_s$ into a scale of finite-dimensional *test data* spaces F_s such that the inequalities

$$\begin{aligned} \|L(u_r)\|_F &\le C(r,s)\|\Pi_s L(u_r)\|_{F_s} \text{ for all } u_r \in U_r, \\ c(s)\|\Pi_s L(u)\|_{F_s} &\le \|L(u)\|_F \text{ for all } u \in U \end{aligned} \tag{1.4}$$

hold. We call a specific choice of trial and test discretization schemes *uniformly stable* if both constants can be chosen independent of r and s for a certain range of these parameters. When restricted to the finite-dimensional trial spaces U_r, the inequalities express equivalence of discrete and nondiscrete norms on the data $L(u_r)$. The second inequality will be easily satisfied by discretization, but the first one will be hard, because it bounds a nondiscrete norm by a discrete norm, and this can work only for finite-dimensional spaces. It also implies uniqueness of solutions of the discretized finite systems $\Pi_s L(u_r) = \Pi_s L(u)$, which is a serious problem [7].

To simplify some of the later arguments, we outline here how we proceed to prove the first inequality of (1.4).

THEOREM 1.1. *Assume a* Poincaré–Friedrichs *inequality*

$$\|f\|_F \le c_1(s)\|f\|_{F_R} + c_2(s)\|\Pi_s f\|_{F_s} \text{ for all } f \in F_R \subset F \tag{1.5}$$

on a regularity subspace F_R *of the data space* F. *Second, assume a* Markov–Bernstein *inequality*

$$\|Lu_r\|_{F_R} \le c_3(r)\|Lu_r\|_F \text{ for all } u_r \in U_r \subset U_R \tag{1.6}$$

on a scale of trial spaces $U_r \subset U_R$ *with* $L(U_R) \subseteq F_R$. *Third, let the trial and test discretization parameters* r *and* s *satisfy the* stability criterion

$$c_1(s)c_3(r) < \frac{1}{2}. \tag{1.7}$$

Then the first inequality of (1.4) *holds with* $C(r,s) \le 2c_2(s)$.

Proof. Just consider

$$\begin{aligned} \|L(u_r)\|_F &\le c_1(s)\|L(u_r)\|_{F_R} + c_2(s)\|\Pi_s L(u_r)\|_{F_s} \\ &\le c_1(s)c_3(r)\|Lu_r\|_F + c_2(s)\|\Pi_s L(u_r)\|_{F_s} \\ &\le \tfrac{1}{2}\|Lu_r\|_F + c_2(s)\|\Pi_s L(u_r)\|_{F_s} \end{aligned}$$

for all $u_r \in U_R$. ∎

Note that in (1.5) we shall have $c_1(s) \to 0$ for $s \to 0$, because the inequality means that a function is small in a weak norm if it is bounded in a strong norm and is small on a large discrete set. In (1.6) we have to expect $c_3(r) \to \infty$ for $r \to 0$ because c_3 bounds a strong norm by a weak one on a finite-dimensional space. Thus the stability criterion (1.7) will usually be satisfied if the test discretization is "fine enough" with respect to the trial discretization.

The final ingredient is the class of *numerical methods*. We do not specify details in this overview, but we can always find nonunique trial functions $u^*_{r,s} \in U_r$ with

$$\|\Pi_s L(u - u^*_{r,s})\|_{F_s} \le \delta_{r,s} \tag{1.8}$$

solving $\Pi_s L(u_r) = \Pi_s L(u)$ approximately with a small tolerance. In fact, the approximation $I_r(u)$ is a solution if we have (1.4) and

$$c(s)\epsilon_r(u) \le \delta_{r,s}. \tag{1.9}$$

Note that we do not attempt to solve the discrete system $\Pi_s L(u_r) = \Pi_s L(u)$ exactly, because it will be overdetermined and unsolvable in general. However, under the assumption (1.9) we know that the relaxed problem (1.8) is solvable. In section 8 we show how to tackle such problems. Altogether, we replace strict collocation by a generalized form of "almost interpolation." Now we can formulate the core result of this paper.

THEOREM 1.2. *If the analytic problem is solvable by* $u \in U_R$ *and if we solve* (1.8) *by some* $u^*_{r,s} \in U_r$, *then the following error bound holds:*

$$\|u - u^*_{r,s}\|_U \le C_a \left(\epsilon_r(u) \left(1 + \frac{C(r,s)}{c(s)} \right) + c(s)\delta_{r,s} \right)$$

provided that all of the above ingredients are available. If the discretization is uniformly stable with constants

$$\begin{array}{rcl} & C(r,s) & \le C, \\ 0 < c \le & c(s) & \le \tilde{c}, \end{array}$$

and if we choose $\tilde{c}\epsilon_r(u) \le \delta_{r,s} \le 2\tilde{c}\epsilon_r(u)$ *to satisfy* (1.9)*, we get the bound*

$$\|u - u^*_{r,s}\|_U \le \epsilon_r(u) C_a \left(1 + \frac{C}{c} + 2\tilde{c}^2\right),$$

which behaves asymptotically like the trial approximation error $\epsilon_r(u)$*.*

Proof. The assertion follows from a simple chain of inequalities:

$$\begin{array}{rcl} \frac{1}{C_a}\|u - u^*_{r,s}\|_U & \le & \|L(u - u^*_{r,s})\|_F \\ & \le & \|L(u - I_r(u))\|_F + \|L(I_r(u) - u^*_{r,s})\|_F \\ & \le & \epsilon_r(u) + c(s)\|\Pi_s L(I_r(u) - u^*_{r,s})\|_{F_s} \\ & \le & \epsilon_r(u) + c(s)\|\Pi_s L(u - u^*_{r,s})\|_{F_s} \\ & & + c(s)\|\Pi_s L(I_r(u) - u)\|_{F_s} \\ & \le & \epsilon_r(u) + c(s)\delta_{r,s} + \frac{C(r,s)}{c(s)}\|L(I_r(u) - u)\|_F \\ & \le & \epsilon_r(u)\left(1 + \frac{C(r,s)}{c(s)}\right) + c(s)\delta_{r,s}. \end{array}$$ ∎

But now we shall have to show how this abstract machinery can be set to work. We shall finally derive specific convergence rates for unsymmetric collocation techniques solving strongly posed problems with continuous dependence in Sobolev spaces, including the Poisson problem with Dirichlet data as an illustration. But note that the above formalism is much more general, and there may be various other future ways to apply the framework, e.g., to unsymmetric methods solving problems in weak form.

The following sections will treat the above ingredients one by one, and then we shall patch the results together. Our key tools will be *nonstationary meshless kernel-based trial spaces* which allow approximation schemes with high-order convergence rates while maintaining stability if paired with sufficiently rich test discretizations. It turns out that the use of smooth kernels makes the final convergence order dependent mainly on the regularity of the solution and the problem. The numerical methods for solving (1.8) will consist of certain variations of the original unsymmetric collocation method, and we already have solvability via (1.9). This saves us from the degeneration problems of the standard unsymmetric collocation technique [7].

2. Well-posed problems and regularity. For example, consider a standard Poisson boundary value problem

$$\begin{array}{rcll} -\Delta u & = & f_\Omega & \text{in } \Omega, \\ u & = & f_D & \text{on } \partial\Omega \end{array} \tag{2.1}$$

on a bounded domain $\Omega \subset \mathbb{R}^d$ with Dirichlet data f_D on the piecewise smooth boundary $\partial\Omega$. In such problems, we consider the equations as being given in strong form; i.e., we assume the solution u to be regular enough to pose the equations pointwise as

$$\begin{array}{rcccll} (-\Delta u)(x) & = & (\delta_x \circ (-\Delta))(u) & = & f_\Omega(x) & \text{for all } x \in \Omega, \\ u(x) & = & (\delta_x \circ Id)(u) & = & f_D(x) & \text{for all } x \in \partial\Omega. \end{array} \tag{2.2}$$

This leads to (1.1) if we define $L(u) := (-\Delta u_{|\Omega}, u_{|\partial\Omega})$ on $U := W_2^\mu(\Omega)$ with values in $F := W_2^{\mu-2}(\Omega) \times W_2^{\mu-1/2}(\partial\Omega)$, with given data $f = (f_\Omega, f_D) \in F$.

But we allow much more general linear equations and boundary value operators. Formally, we follow the notation of [3] and others to combine differential and boundary operators into just one equation and write the latter as (1.1) where u is a function from some normed space U of functions. The mapping $L : U \to F$ maps solutions $u \in U$ to their data $L(u) \in F$, and the given problem consists in the inversion of L.

When aiming at methods with error bounds and convergence, we have to take a closer look at the given analytic problem (1.1). In particular, we shall assume that the problem (1.1) is well-posed in the sense that the solution u depends continuously on the data f of the right-hand side of (1.1). But we have to make this more precise. This can be done in various ways, e.g., by *total sets* of data functionals, but this is not quantitative. For later use we impose a norm $\|.\|_F$ on $F := L(U)$ in a suitable way and assume (1.2) with a positive "analytic" constant C_a which describes the norm of the linear map L^{-1} that takes the data $f \in F$ and maps them back to the solution u in the function space U. Clearly, for such a priori inequalities we must be careful with the choice of norms, because they depend on regularity theory, and they always imply that the homogeneous equation has only the trivial solution. The numerical methods following below will work on discretized versions F_s of F, and thus the proper choice of F will also have practical consequences.

So far we have not mentioned any specific numerical algorithm. But if any numerical method has produced an approximate solution $\tilde{u} \in U$ to the problem (1.1), one can calculate the data $\tilde{f} = L(\tilde{u}) \in F$ and the norm $\|\tilde{f} - f\|_F$ to get the a posteriori error bound

$$\|u - \tilde{u}\|_U \le C_a \cdot \|L(u - \tilde{u})\|_F = C_a \cdot \|f - \tilde{f}\|_F \tag{2.3}$$

for free, since the *residuals* $L(u - \tilde{u}) = f - \tilde{f}$ are explicitly known. It means that errors in the solution are bounded by the norm of the residuals, multiplied with the analytic constant. *Thus any numerical technique that produces approximate solutions of well-posed problems with small residuals will automatically guarantee small errors in the solution.* This trivial observation is well known in numerical analysis and serves as a basis for *defect correction* and *residual minimization* techniques, and is important for providing a safe a posteriori foundation for many unsafe and ad hoc numerical calculations published in science and engineering journals. If the underlying problem is continuously dependent on the data and if the naive user at least checks the residuals carefully, the calculations are on the safe side. But, unfortunately, there is no handbook listing all known inequalities of the form (1.2) for typical applications in science and engineering. In particular, it would be very useful to have proven upper bounds for the analytic constants.

Guided by regularity theory for elliptic problems, we focus on operator equations (1.1) where the linear map L splits into maps $L^1, \dots, L^n$ with $L^i : U \to F^i$, $1 \le i \le n$, such that $F = F^1 \times \cdots \times F^n$ is the data space. We assume $U = W_2^\mu(\Omega)$ for some bounded domain $\Omega \subset \mathbb{R}^d$ and $F^i := W_2^{\mu-\mu_i}(\Omega^i)$, where μ_i is defined via a trace theorem by the order of the operator L^i and the dimension $d_i \le d$ of the partial domain $\Omega^i \subset \overline{\Omega} \subset \mathbb{R}^d$. The space F is then equipped with the sup of the norms of the spaces F^i. The regularity subspace U_R occurring later will then be a space $U_R \subseteq W_2^m(\Omega) \subset U := W_2^\mu(\Omega)$ for some $m \ge \mu$.

In the standard Poisson problem with Dirichlet data we take $L(u) = (-\Delta u, u_{\partial\Omega})$ mapping $U \subseteq W_2^\mu(\Omega)$ into $F = W_2^{\mu-2}(\Omega) \times W_2^{\mu-1/2}(\partial\Omega)$. This is a well-established

continuous dependence setting if the domain is smooth enough to make trace theorems and the regularity order μ valid. See, e.g., [3, 12] for early references which also allow distributional data and negative μ.

3. Approximation from trial spaces. The second ingredient of our framework is some additional regularity defined via a subspace U_R of the space U occurring in the continuous dependence bound (1.2). At this point, the regularity space can be quite general, but we also want an approximation property like (1.3) to hold. Thus we now have to consider our third ingredient, i.e., techniques that construct approximate solutions $\tilde{u}_r$ from a scale of *trial spaces* $U_r \subseteq U$ with a trial discretization parameter r. Note that this includes plenty of methods, with or without meshes, like finite elements, Petrov–Galerkin schemes, spectral methods, and all variations of collocation. It is trivial that the choice of the trial space should be such that the true solution u can be approximated easily by functions from the trial space. In case of solutions with singularities, like for Poisson problems on domains with incoming vertices, one should make sure that the trial space contains the expected singular functions.

One way to make this more precise is to assume that there is a mapping $A_r : U_R \to U_r$ with

$$\|u - A_r(u)\|_U =: \delta_r(u) \text{ for all } u \in U_R \tag{3.1}$$

with a certain error $\delta_r(u)$ which will depend on the regularity subspace U_R.

But the previous section teaches us that we do not need to approximate the exact solution u in the space U by functions $u_r \in U_r \subset U$ directly. It suffices to make sure that the residuals $L(u) - L(u_r)$ are small. Thus the crucial quantity is the residual error $\|L(u - u_r)\|_F$ for any $u \in U$ and an approximation $u_r \in U_r$. In contrast to the theory of finite elements, we do not consider optimal approximations of u by u_r here, nor do we attempt to minimize the above error with respect to u_r. Instead, we are satisfied if the trial space U_r contains for each function $u \in U$ an approximation $u_r := I_r(u)$ with small *residual error* $\epsilon_r(u)$ as in (1.3).

Of course, if an L-independent approximation operator A_r with (3.1) is available, one can take $I_r = A_r$ and assume $\epsilon_r(u) \le \|L\|\delta_r(u)$ because of

$$\epsilon_r(u) = \|L(u - I_r(u))\|_W \le \|L\|\|(u - A_r(u))\|_U \le \|L\|\delta_r(u).$$

But there may be better choices of I_r if L and the special structure of the residual space W are taken into account.

Inspection of (1.3) for F being a Cartesian product of Sobolev spaces reveals that the special approximation $I_r(u)$ should approximate u well *including its derivatives*, as far as they occur in the collection of data spaces F^i forming the space F. In fact, if we go back to our special case $U = W_2^\mu(\Omega)$ and $F = F^1 \times \cdots \times F^n$ with $F^i := L^i(U) := W_2^{\mu-\mu_i}(\Omega^i)$ with a regularity subspace $U_R \subseteq W_2^m(\Omega)$ and $m > \mu$, we should expect approximation bounds like

$$\begin{aligned} \|L^i(u) - L^i(I_r(u))\|_{W_2^{\mu-\mu_i}(\Omega^i)} &\le Cr^{m-\mu}\|L^i(u)\|_{W_2^{m-\mu_i}(\Omega^i)} \\ &\le Cr^{m-\mu}\|u\|_{W_2^m(\Omega)} \text{ for all } u \in W_2^m(\Omega), \end{aligned} \tag{3.2}$$

and this should work for a reasonable choice of $0 \le \mu \le m$ and with rates that just depend on the *regularity gap* $m - \mu$ and not on the order of the operators involved.

Note that the standard trial spaces of h-type finite element techniques consisting of piecewise linear functions fail to provide approximations of more than first-order

derivatives. In contrast to this, trial spaces generated by sufficiently smooth kernel functions can contain approximations to derivatives of any order, without any additional work needed. We shall explain this in the following sections.

In contrast to many engineering applications where a rather simple solution function is calculated via a huge finite element method system of millions of unknowns, we tend to argue in favor of small trial spaces designed to capture the essential features of the solution without taking the detour via a fine-grained space discretization. The consequence will be that the linear systems get unsymmetric, because any solution from a small *trial* space must be *tested* on a fine-grained space discretization, asking for many more degrees of freedom on the "test side" than on the "trial side." Unsymmetry of a method can be a feature instead of a bug. In what follows we shall investigate the relation of test and trial spaces more closely.

4. Kernel-based trial spaces. Now it is time to study maps I_r or A_r with good approximation properties for certain trial spaces U_r in the sense of (1.3) and (3.1). This is independent of PDE solving, and we shall see that nonstationary scales of meshless kernel-based trial spaces work perfectly.

DEFINITION 4.1. *A* kernel *is a function of the form* $K : \Omega \times \Omega \to \mathbb{R}$ *with* $\Omega \subseteq \mathbb{R}^d$. *It is* translation-invariant *if* $K(x,y) = \Phi(x-y)$ *with* $\Omega = \mathbb{R}^d$ *and* $\Phi : \mathbb{R}^d \to \mathbb{R}$. *It is* radial *if it is translation-invariant and of the form*

$$K(x,y) = \Phi(x-y) = \phi(\|x-y\|_2) \text{ with } \phi : [0,\infty) \to \mathbb{R} \text{ and } x, y \in \mathbb{R}^d.$$

Radial kernels are also called radial basis functions.

Note that radial basis functions ϕ can in principle be used in any space dimension, but certain properties of the associated translation-invariant kernel Φ on $\mathbb{R}^d$ may depend [6, 19] on the dimension d.

Kernels provide excellent tools in various disciplines, including approximation theory, partial differential equations, and machine learning [17]. The most important kernels are *reproducing kernels* of some Hilbert space which can be called the "native" Hilbert space for the kernel. Any Hilbert space H of functions on a domain Ω with continuous and linearly independent point evaluations has a kernel K with the reproduction property

$$f(x) = (f, K(x,\cdot)) \text{ for all } f \in H,\ x \in \Omega.$$

Conversely, any (strictly) positive definite [6, 19] and continuous kernel K on Ω is the reproducing kernel of a *native* Hilbert space N_K of continuous functions on Ω. We denote the norm on the native space N_K by $\|.\|_K$.

We focus here on *trial spaces* provided by kernels. Like in wavelet theory, the notions of *translation* and *dilation* play an important role. First, a general kernel $K : \Omega \times \Omega \to \mathbb{R}$ can be translated to points $y_1, \ldots, y_M \in \Omega$ called *centers* to provide trial functions $u_j(x) := K(x, y_j)$, $1 \le j \le M$, on Ω. In many cases, the set $Y := \{y_j : 1 \le j \le M\}$ of centers should fill a bounded domain Ω in such a way that the centers get dense when $M \to \infty$. This is expressed by the *fill distance*

$$h := h(Y,\Omega) := \sup_{x\in\Omega} \inf_{y\in Y} \|x-y\|_2$$

depending on Ω and the M centers in Y, which should converge to zero if M tends to infinity. The fill distance is the radius of the largest open ball with center in Ω that contains none of the centers y_j from Y. We use the notation h here, but later

we shall have two different fill distances for trial and test centers, and then we shall use r and s for clarity.

A *nonstationary* scale of kernel-based trial spaces can now be defined as

$$U_r := \text{span}\,\{K(\cdot, y_j) \,:\, 1 \le j \le M\} \text{ with } r := h(Y, \Omega), \tag{4.1}$$

where the dependence on the location and number of the centers is suppressed in the notation.

In the above *nonstationary* situation, only translations, but no dilations are used. The translated kernel is fixed and independent of the fill distance. There is no rescaling, if the fill distance gets small. This is in contrast to the *stationary* technique in standard and general finite elements [4]. There, the basis functions are rescaled when the fill distance changes, and in the translation-invariant kernel-based case this can be described by a scale of trial spaces

$$U_r := \text{span}\,\left\{\Phi\left(\frac{x - y_j}{r}\right) \,:\, 1 \le j \le M\right\},$$

where now the wavelet style interaction of translation and dilation is apparent.

The mathematics of the stationary and nonstationary cases are quite different. This often leads to misunderstandings. The stationary situation, as included in the meshless generalized finite element method [4], uses polynomial reproduction and the Bramble–Hilbert lemma. If centers are on a grid, it applies the Strang–Fix theory. Convergence orders are closely tied to polynomial reproduction properties, and the choice of kernels is quite restricted, because integrable kernels like the Gaussian do not yield convergent stationary approximations for $h \to 0$ [6]. Stability is usually much better than in the nonstationary case, but convergence rates (if they exist at all, e.g., for thin-plate splines or multiquadrics) are much smaller. We focus on nonstationary kernel-based trial spaces here, because condition problems can be overcome [5, 9], and we are heading for methods with high approximation orders.

We define a map $I_r \,:\, u \to u_r := I_r(u) \in U_r$ of (1.3) via *interpolation in the trial centers* by solving the system

$$u_r(y_k) := \sum_{j=1}^{M} \alpha_j K(y_j, y_k) = u(y_k),\ 1 \le k \le M, \tag{4.2}$$

for the coefficients $\alpha_1, \ldots, \alpha_M$ defining the function $u_r := I_r(u)$ in terms of the basis functions of the nonstationary trial space U_r of (4.1). This interpolation problem is solvable by definition if the kernel K is symmetric and positive definite [6, 19], because then the $M \times M$ matrix with entries $K(y_j, y_k)$ is symmetric and positive definite. Table 4.1 gives some examples. We ignore *conditionally* positive definite kernels here and refer to the literature [6, 16, 19] for details.

The interpolation system (4.2) makes sense for all functions u which have well-defined function values at the trial centers y_k. Thus the mapping I_r is at least defined on $C(\Omega)$, but for solutions u of PDE problems in strong form we use it on a regularity subspace U_R of $C(\Omega) \subset U$.

The book [19] contains a fairly complete account of interpolation error bounds in the nonstationary setting, while bounds for stationary and regular cases are in [6]. But in view of (1.3) and (3.2), we need very general error bounds in Sobolev spaces which are not covered in these books. Here (on the trial side) and later (on the test

TABLE 4.1
Radial basis functions $\phi(r)$, positive definite on $\mathbb{R}^d$.

Function	$\phi(r)$	Range	Smoothness β
Gaussian	$\exp(-r^2)$	$d \geq 1$	all β
inverse multiquadric	$(r^2+c^2)^\gamma,\ \gamma < -d/2,\ c > 0$	$d \geq 1$	all β
Sobolev for $W_2^k(\mathbb{R}^d)$	$r^{k-d/2}K_{k-d/2}(r),\ k > d/2$	$d \geq 1$	$\beta = 2k-d$
Wendland C^2 [18]	$(1-r)_+^4(1+4r)$	$d \leq 3$	3
	$(1-r)_+^5(1+5r)$	$d \leq 5$	3
Wendland C^4 [18]	$(1-r)_+^6(3+18r+35r^2)$	$d \leq 3$	5
	$(1-r)_+^7(1+7r+16r^2)$	$d \leq 5$	5

side for proving (1.5)) we use a general result from [13] which was extended in [20], while the range of admissible parameters was enlarged in [14].

THEOREM 4.2. *Suppose $\Omega \subset \mathbb{R}^d$ is a bounded domain with an interior cone condition. Choose $q \in [1,\infty]$ and constants*

$$0 \leq \mu < \mu + d/2 < \lfloor m \rfloor \tag{4.3}$$

with μ being an integer. Then there are positive constants C, h_0 such that

$$|u|_{W_q^\mu(\Omega)} \leq C\left(h^{m-\mu-d(1/2-1/q)_+}|u|_{W_2^m(\Omega)} + h^{-\mu}\|u\|_{\infty,Y_h}\right) \tag{4.4}$$

holds for every discrete set Y_h in Ω with fill distance at most $h \leq h_0$ and every $u \in W_2^m(\Omega)$.

This can be seen as a quantitative Poincaré–Friedrichs inequality for functions which are small on a finite subset, and it is independent of any trial space. If we replace seminorms by norms and extract $h^{-\mu}$ out of the right-hand side, the rest is a μ-independent norm on $W_2^m(\Omega)$ and we can apply interpolation theory to replace (4.4) by

$$\|u\|_{W_q^\mu(\Omega)} \leq C\left(h^{m-\mu-d(1/2-1/q)_+}\|u\|_{W_2^m(\Omega)} + h^{-\mu}\|u\|_{\infty,Y_h}\right) \tag{4.5}$$

under the assumptions (4.3) *without* the restriction of μ being an integer.

Now we take $h = r$ because we discretize the *trial* side, and we interpolate a function u on Y_r by $I_r(u)$ using points from Y_r as translations in (4.1) to get

$$\|u - I_r(u)\|_{W_2^\mu(\Omega)} \leq Cr^{m-\mu}\|u - I_r(u)\|_{W_2^m(\Omega)} \text{ for all } u \in W_2^m(\Omega). \tag{4.6}$$

Then we use the standard fact that the interpolant $I_r(u)$ solves the minimization problem

$$\|v\|_K \to \min,\ v \in K,\ v(y_j) = u(y_j) \text{ for all } y_j \in Y_r,$$

implying that $\|I_r(u)\|_K \leq \|u\|_K$ holds if we assume u to be in the native space N_K for the kernel K.

Therefore we strengthen the requirement on the regularization subspace U_R and on the regularity of our solution u to

$$u \in N_K = U_R \subseteq W_2^m(\Omega) \subseteq U \tag{4.7}$$

with bounded embeddings, where we always assume (4.3). This is easy if the kernel is smooth enough, and for the kernels in Table 4.1 the inequality

$$2m \leq \beta + d \tag{4.8}$$

is a sufficient condition.

We can now replace (4.6) by

$$\|u - I_r(u)\|_{W_2^\mu(\Omega)} \le Cr^{m-\mu}\|u\|_K \text{ for all } u \in N_K$$

with a different constant. This inequality can be coupled with trace theorems for the operators $L^i : W_2^m(\Omega) \to W_2^{m-\mu_i}(\Omega^i)$ to get

$$\|L^i(u - I_r(u))\|_{W_2^{\mu-\mu_i}(\Omega^i)} \le C\|u - I_r(u)\|_{W_2^\mu(\Omega)} \le Cr^{m-\mu}\|u\|_K \text{ for all } u \in N_K$$

which yields (3.2) in a slightly restricted form and our third ingredient (1.3) as

$$\|L(u - I_r(u))\|_F \le Cr^{m-\mu}\|u\|_K =: \epsilon_r(u) \text{ for all } u \in U_R = N_K \tag{4.9}$$

under the assumptions (4.3), (4.7), and (4.8).

5. Stability of kernel-based test discretizations. We now consider the stability conditions (1.4), our fourth ingredient. We do this for meshless kernel-based trial spaces and for our running example generalizing the Poisson equation. The trial discretization via U_r and a set Y_r of centers is chosen as in section 4. We assume (4.7) and have the approximation result (4.9). On the test side, we use a set X_s of test centers which has a fill distance s on all of $\overline{\Omega}$. For all the operators L^i that arise in L, we will have a selection $X_s^i := X_s \cap \Omega^i$ of points with the same fill distance with respect to Ω^i, because we can assume that all Ω^i are subsets of $\overline{\Omega}$. The projectors Π_s^i on F^i just map functions from $F^i = W_2^{\mu-\mu_i}(\Omega^i)$ to their values on X_s^i. We thus have to assume Sobolev embedding conditions

$$2(\mu - \mu_i) > d_i := \dim(\Omega^i) \le d := \dim(\Omega),\ 1 \le i \le n. \tag{5.1}$$

The discretized spaces F_s^i will be $\mathbb{R}^{|X_s^i|}$ with the L_∞ norm, and we have

$$\Pi_s^i L^i(u) = (L^i(u))(X_s^i),\ \|\Pi_s^i L^i(u)\|_{F_s^i} = \|L^i(u)\|_{\infty,X_s^i}.$$

This implies by Sobolev embedding

$$\|\Pi_s^i L^i(u)\|_{F_s^i} = \|L^i(u)\|_{\infty,X_s^i} \le \|L^i(u)\|_{\infty,\Omega^i} \le C_i\|L^i(u)\|_{W_2^{\mu-\mu_i}(\Omega^i)},$$

where the constant is independent of u and s. We now assemble this into a discretization $F_s := F_s^1 \times \cdots \times F_s^n$ with $\Pi_s := \Pi_s^1 \times \cdots \times \Pi_s^n$ of $F = F^1 \times \cdots \times F^n$ and take the sup norm of the components. Then we have

$$\begin{aligned} \|\Pi_s L(u)\|_{F_s} &= \sup_{1\le i\le n} \|\Pi_s^i L^i(u)\|_{F_s^i} \\ &\le C \sup_{1\le i\le n} \|L^i(u)\|_{W_2^{\mu-\mu_i}(\Omega^i)} \\ &= C\|L(u)\|_F \end{aligned}$$

and get the second inequality of (1.4) with a constant that is independent of s and dependent only on Sobolev embedding. This leaves us to prove the first inequality of (1.4) via Theorem 1.1.

Fortunately, the inequality (4.5) holds for general Sobolev spaces, and we can apply it on the test side for different operators. We get

$$\|L^i(u)\|_{W_2^{\mu-\mu_i}(\Omega^i)} \le C\left(s^{m-\mu}\|L^i(u)\|_{W_2^{m-\mu_i}(\Omega^i)} + s^{-(\mu-\mu_i)}\|L^i(u)\|_{\infty,X_s^i}\right)$$

under the assumptions (5.1) and

$$d_i/2 < \mu - \mu_i < \mu - \mu_i + d_i/2 < \lfloor m - \mu_i \rfloor \text{ for all } i,\ 1 \le i \le n, \tag{5.2}$$

which pose no serious problems if m is large enough. Now we define F_R to be the range of $L(U_R)$ in the Cartesian product of all spaces $W_2^{m-\mu_i}(\Omega^i)$, taking the sup of the component norms. With part of the notation

$$\min_i \mu_i =: \underline{\mu} \le \mu_i \le \overline{\mu} := \max_i \mu_i, \tag{5.3}$$

this yields (1.5) in the form

$$\|L(u)\|_F \quad \le \quad C\left(s^{m-\mu}\|L(u)\|_{F_R} + s^{\underline{\mu}-\mu}\|\Pi_s L(u)\|_{F_s}\right).$$

We now want to consider (1.6). Assume K to be a translation-invariant positive definite kernel of finite smoothness which is Fourier-transformable in $\mathbb{R}^d$ with an exact decay

$$c(1+\|\omega\|_2)^{-\beta-d} \le \hat{K}(\omega) \le C(1+\|\omega\|_2)^{-\beta-d} \text{ for all } \omega \in \mathbb{R}^d, \tag{5.4}$$

where the constants β can be read from Table 4.1. If we again assume (4.8), we can cite the Bernstein-type inequality

$$\|u_r\|_{W_2^m(\Omega)} \le C \cdot \|u_r\|_{W_2^{(d+\beta)/2}(\mathbb{R}^d)} \le Cr^{-(d+\beta)/2}\|u_r\|_{L_\infty(\Omega)} \text{ for all } u_r \in U_r$$

from [15] provided that the trial centers in $Y_r \subset \Omega$ are not too wildly scattered in the sense that the minimal separation distance $q(Y_r)$ is uniformly bounded below by the fill distance $h(Y_r, \Omega)$ via

$$q(Y_r) := \min_{y_j \ne y_i \in Y_r} \|y_j - y_i\|_2 \ge C \sup_{y \in \Omega} \min_{y_i \in Y_r} \|y - y_i\|_2 =: h(Y_r, \Omega) \tag{5.5}$$

such that both quantities behave asymptotically like the trial discretization parameter r. This yields

$$\begin{aligned}
\|L(u_r)\|_{F_R} &= \max_i \|L_i(u_r)\|_{W_2^{m-\mu_i}(\Omega^i)} \\
&\le C\|u_r\|_{W_2^m(\Omega)} \\
&\le Cr^{-(d+\beta)/2}\|u_r\|_{L_\infty(\Omega)} \\
&\le Cr^{-(d+\beta)/2}\|u_r\|_{W_2^\mu(\Omega)} \\
&= Cr^{-(d+\beta)/2}\|u_r\|_U \\
&\le C_a C r^{-(d+\beta)/2}\|L(u_r)\|_F \text{ for all } i,\ 1 \le i \le n,
\end{aligned}$$

under the additional assumption $\mu > d/2$. This result is far from optimal, but it establishes (1.6) and allows us to apply Theorem 1.1 under a stability condition of the form

$$Cs^{m-\mu}r^{-(d+\beta)/2} < \frac{1}{2}. \tag{5.6}$$

This finally implies the following theorem.

THEOREM 5.1. *Assume that the trial kernel K with* (5.4) *is smooth enough to satisfy* (4.8). *Let the trial space consist of quasi-uniform translates on Ω with discretization parameter r, and consider a test discretization on $\overline{\Omega}$ with fill distance s. Then, under the notation* (5.3), *and the additional conditions* (4.3), (5.2), (5.6), *and $\mu > d/2$, the first inequality of* (1.4) *is satisfied with*

$$\|L(u_r)\|_F \quad \le \quad Cs^{\underline{\mu}-\mu}\|\Pi_s L(u_r)\|_{F_s}$$

for all $u_r \in U_r$.

6. Strong convergence in Sobolev spaces. We now assemble what we have in case of our running example with continuous dependence in Sobolev norms. In contrast to the introduction, we proceed here from the user's point of view.

We start with the analytic problem. Consider an operator equation $L(u) = f$ as in (1.1) whose solution u is continuously dependent on the data f. We assume continuous dependence in the sense of (1.2) to hold if we pick spaces $U = W_2^\mu(\Omega)$ and F defined as a Cartesian product of Sobolev trace spaces F^i as in section 2. But note that users have to make sure in each application problem that the a priori inequalities composing (1.2) are actually satisfied. If several choices of μ are possible, the user should know that the final convergence will take place in $U = W_2^\mu(\Omega)$, but large μ have to be paid for by regularity. If convergence of higher-order derivatives is of importance, a sufficiently large μ must be chosen. Since we solve problems in strong form via evaluation of residuals, we have to pick μ large enough to let all data have continuous point evaluations. This is expressed by the requirement (5.1). At this point, the lower bounds for μ will rule out problems with low regularity. Such problems should be tackled with methods using weak data functionals and involving integration. We plan to deal with such methods in the future, in particular with the unsymmetric meshless Petrov–Galerkin method of Atluri and his collaborators [1].

The next step concerns regularity. We assume that the solution should have at least a $U_R \subseteq W_2^m(\Omega)$ regularity with some $m > \mu$. By standard arguments from approximation theory, the difference between m and μ is the driving force for the possible convergence rates. The user has to decide which m is adequate. Larger m will improve the convergence rates, but they may not be justified by the smoothness of the problem.

Then we pick a kernel K which is smooth enough to have its native space N_K contained in $W_2^m(\Omega)$. In view of Table 4.1, this requires (5.4) and (4.8). The solution u must have at least the regularity of $W_2^m(\Omega)$, because it should be in $U_R = N_K \subseteq W_2^m(\Omega)$. The excess regularity of N_K over $W_2^m(\Omega)$ does not pay off later, and thus it is a good idea to stay with a kernel satisfying $\beta + d = 2m$ to have norm equivalence between $W_2^m(\Omega)$ and N_K. Note that the compactly supported radial polynomial kernels of Wendland [18] satisfy this for certain choices of m, β, and d.

Now it is time to pick a meshless trial discretization U_r via a set Y_r of trial centers with fill distance r using the kernel K. Then we can expect an error behavior $\epsilon_r(u) \le Cr^{m-\mu}$ for $u \in N_K$ for a direct interpolant to the regular solution in the points of Y_r. This rate is the ideal goal we want to achieve for our numerical solution of the given operator equation.

The next step is to pick a test discretization via a set X_s of test centers in $\overline{\Omega}$ with fill distance s. The second inequality of (1.4) holds with $c(s)$ independent of s because we assume continuous residuals and corresponding Sobolev embedding theorems. By Theorem 5.1, the first inequality of (1.4) will then hold with $C(r,s) \le Cs^{\mu-\mu}$ using (5.3). But we have to make the test discretization fine enough to satisfy (5.6). As expected, this means that the test discretizations must be somewhat finer than the trial discretizations, and the required relation between s and r is

$$s < c \cdot r^{1+\frac{\mu}{m-\mu}}. \tag{6.1}$$

There is plenty of leeway for small trial and large test spaces.

We are now ready to put everything into Theorem 1.2, while we assume that we solve the discretized problem (1.8) with accuracy $\delta_{r,s}$. With new generic constants

we get

$$\|u - u^*_{r,s}\|_{W_2^\mu(\Omega)} \leq C\left(r^{m-\mu}\left(1 + Cs^{\underline{\mu}-\mu}\right) + \delta_{r,s}\right)\|u\|_K. \tag{6.2}$$

Note that this bound has the proper approximation error of order $r^{m-\mu}$ holding between Sobolev spaces $U = W_2^\mu(\Omega)$ and $U_R = N_K \subseteq W_2^m(\Omega)$, but there also is a counteracting term $s^{\underline{\mu}-\mu}$ which is the price we have to pay for working on discrete residuals in the L_∞ norm while bounding the residual error in the norm on Sobolev trace spaces $W_2^{\mu-\mu_i}(\Omega^i)$ in order to use continuous dependence on Sobolev space data. If we choose $\delta_{r,s}$ properly via (1.9) and s via (5.6), we have solvability of the system and an error bound

$$\|u - u^*_{r,s}\|_{W_2^\mu(\Omega)} \leq Cr^{m-2\mu+\underline{\mu}-\frac{\mu(\mu-\underline{\mu})}{m-\mu}}\|u\|_{W_2^m(\Omega)}.$$

Of course, this is not an optimal bound because μ must be positive and even larger than $d/2$. Consequently, there is quite some future work necessary on this bound, though it is improving when m is much larger than μ. In this context, it is not surprising that most of the practical applications of unsymmetric collocation methods have very regular solutions.

To show the minimum regularity requirements for the results of this section, we should track the possible range of m for the Poisson problem in d dimensions. For operators $L^1 := -\Delta$ and L^2 providing Dirichlet boundary data, we have $\mu_1 = 2$ and $\mu_2 = 1/2$ with $d_1 = d$ and $d_2 = d - 1$. Then (5.2) requires $d/2 < \mu - 2$ while (4.3) leads to $\lfloor m \rfloor > 2 + d$ as the minimum regularity requirement. This clearly needs improvement by future work, but it should be mentioned that the resulting error bound is strong enough to include derivatives up to order μ with $2 + d/2 < \mu < \lfloor m \rfloor - d/2 \geq m - 1 - d/2$.

7. Weak convergence in Sobolev spaces. Analysis of the previous section shows that the term $s^{\underline{\mu}-\mu}$ in (6.2) with some positive μ satisfying (5.1) and $\mu > d/2$ makes the final bound worse than expected. Tracing this back to (4.4) shows that one should better look at another variation which allows $\mu = 0$ at that point without spoiling the assumption that the data are still continuous. In fact, (4.4) also allows

$$\|u\|_{L_\infty(\Omega)} \leq C\left(h^{m-d/2}\|u\|_{W_2^m(\Omega)} + \|u\|_{\infty,Y_h}\right) \text{ for all } u \in W_2^m(\Omega) \tag{7.1}$$

if $d/2 < \lfloor m \rfloor$ holds.

But this does not easily fit into the framework required for continuous dependence. Thus we start anew, defining the data spaces F^i as spaces $C(\Omega^i)$ of continuous functions under the L_∞ norm. To make continuous dependence valid, we use embeddings $C(\Omega^i) \subset L_2(\Omega^i) = W_2^0(\Omega^i) \subseteq W_2^{\mu-\mu_i}(\Omega^i)$ for $\mu := \underline{\mu} = \min \mu_i$. Then we apply the standard continuous dependence relating the standard solution space $W_2^\mu(\Omega)$ to the trace spaces $L^i(W_2^\mu(\Omega)) \subseteq W_2^{\mu-\mu_i}(\Omega^i)$, which fortunately hold for small and even negative μ, if the domain is smooth [12, 3]. This yields a new continuous dependence relation via

$$\begin{aligned}
\|u\|_{W_2^\mu(\Omega)} &\leq C \max_i \|L^i(u)\|_{W_2^{\mu-\mu_i}(\Omega^i)} \\
&\leq C \max_i \|L^i(u)\|_{W_2^0(\Omega^i)} \\
&\leq C \max_i \|L^i(u)\|_{C(\Omega^i)} \\
&= C \max_i \|L^i(u)\|_{F^i} \\
&= C\|L(u)\|_F,
\end{aligned}$$

holding only on the subspace U of functions u in $W_2^\mu(\Omega)$ with continuous data $L(u)$. Note that this will lead to a weak convergence result in $U \subset W_2^\mu(\Omega)$, though the problem formulation is still strong. For instance, a problem with Dirichlet data will lead to $\mu = 1/2$ due to the trace map $W_2^\mu(\Omega) \to W_2^{\mu-1/2}(\partial\Omega)$ if all other trace or differential operators have a larger loss in the order of the respective Sobolev trace spaces.

Thus we now repeat our basic argument for $U \subset W_2^\mu(\Omega)$ with $\mu = \underline{\mu} = \min_i \mu_i$. Our choice of regularity space U_R and the kernel K will be as above. This fixes β and m. To derive the approximation order in (1.3) we apply (4.5) to get

$$\begin{aligned}\|L^i(u - I_r(u))\|_{W_\infty^0(\Omega^i)} &\le Cr^{m-\mu_i-d/2}\|L^i(u - I_r(u))\|_{W_2^{m-\mu_i}(\Omega^i)} \\ &\le Cr^{m-\mu_i-d/2}\|u - I_r(u)\|_{W_2^m(\Omega)} \\ &\le Cr^{m-\overline{\mu}-d/2}\|u\|_{W_2^m(\Omega)},\end{aligned}$$

which requires only

$$d/2 < \lfloor m - \mu_i \rfloor,\ 1 \le i \le n, \tag{7.2}$$

and where we now also need $\overline{\mu}$ from (5.3). Thus we get

$$\epsilon_r(u) \le Cr^{m-\overline{\mu}-d/2}$$

for (1.3).

The discretization of the $F^i = C(\Omega^i)$ spaces is again by pointwise evaluation on a set X_s^i of test centers, taking the discrete L_∞ norm, but we now can skip Sobolev embedding which led to the inequalities (5.1) we want to avoid now. Since every data space is equipped with the L_∞ norm, we have $c(s) = 1$ in the second inequality of (1.4). The proof of the first inequality of (1.4) again proceeds via Theorem 1.1. To prove (1.5) we start with (7.1) on the various data:

$$\|L^i(u)\|_{W_\infty^0(\Omega^i)} \le C\left(s^{m-\mu_i-d_i/2}\|L^i(u)\|_{W_2^{m-\mu_i}(\Omega^i)} + \|L^i(u)\|_{\infty,X_s^i}\right)$$

for all $u \in U_R = N_K$ where we need (7.2) again. We define F_R as in the previous section, and then we have (1.5) in the form

$$\|L(u)\|_F \le C\left(s^{m-\overline{\mu}-d/2}\|L(u)\|_{F_R} + \|\Pi_s L(u)\|_{F_s}\right).$$

Unfortunately, the proof for (1.6) given in section 5 proceeds via $\|.\|_\infty$ and thus needs $\mu > d/2$. Of course there is always some a priori inequality of the form (1.6), but we currently have no explicit upper bounds for $c_3(r)$ in terms of r. Anyway, if we take s small enough to satisfy (1.7) with $c_1(s) = Cs^{m-\overline{\mu}-d/2}$, Theorem 1.1 still is valid and yields the first inequality of (1.4) with $C(r, s)$ independent of r and s provided that s is small enough.

Then we continue as in the proof of Theorem 5.1. Since the discretization scheme is uniformly stable for sufficiently small test discretizations s, Theorem 1.2 now gives the error bound

$$\|u - u_{r,s}^*\|_{W_2^\mu(\Omega)} \le Cr^{m-\overline{\mu}-d/2}\|u\|_{W_2^m(\Omega)} \text{ for all } u \in W_2^m(\Omega)$$

provided that the numerical solution of (1.8) observes (1.9) and the test discretization is fine enough in a way we currently cannot specify explicitly. The left-hand norm

is rather weak here, and the approximation order can probably be improved. For instance, a standard two-dimensional Poisson problem with Dirichlet data would lead to

$$\|u - u^*_{r,s}\|_{W_2^{1/2}(\Omega)} \le C r^{m-3} \|u\|_{W_2^m(\Omega)} \text{ for all } u \in W_2^m(\Omega),$$

but an optimal rate for the Sobolev spaces involved would be $m-1/2$ instead of $m-3$. The minimum regularity in this case is $m = 4$ because of (7.2).

Future work should improve the results of this and the previous sections. This may be done by better choices of spaces and norms, plus better versions of the Markov–Bernstein inequality (1.6) which are currently investigated.

8. Numerical methods. We now look at techniques to solve the discrete problem (1.8). It amounts to solving the n linear problems

$$\Pi_s^i L^i (u - u^*_{r,s}) = 0,\ 1 \le i \le n,$$

approximately, where we discretized the operators L^i on the domains Ω^i by taking only point evaluations. This takes the form of collocation

$$L^i(u)(x_{ji}) = L^i(u^*_{r,s})(x_{ji}) = 0,\ 1 \le i \le n,\ 1 \le j \le N_i,$$

where the points of the test discretization X_s are the union of the sets

$$X_s^i := \{x_{j1}, \dots, x_{jN_i}\},\ 1 \le j \le n,$$

and where we dropped the dependence on s in the notation for the x_{ji} and for N_i. For a shorthand notation, we introduce the functionals

$$\lambda_{ji} \ :\ v \mapsto L^i(v)(x_{ji})$$

and rearrange them into a single-indexed list $\lambda_1, \dots, \lambda_N$ with $N = N_1 + \cdots + N_n$.

Since $u^*_{r,s}$ should be in the trial space U_r generated by translations of the kernel K at trial centers forming $Y_r := \{y_1, \dots, y_M\}$, we arrive at a system

$$\sum_{m=1}^{M} \alpha_m \lambda_i^z K(z, y_m) = \lambda_i^z u(z),\ 1 \le i \le N,$$

with M unknowns and N equations. In case $M = N$ this is exactly the unsymmetric collocation technique dating back to Kansa in 1986 [8]. It has no rigid foundation yet, and it can fail in specially constructed situations [7], though it works fine in many applications. In the first years it was applied to small problems with smooth solutions due to serious condition problems, but recently there have been results on preconditioning [5, 9] that allow a wider range of applicability.

In view of Theorem 5.1 and the two previous sections we know that $N \ge M$ holds and the system has full rank M, provided that our stability conditions are satisfied, calling for a somewhat finer discretization on the test than on the trial side. Thus the system will be unsymmetric and overdetermined, but at least there is no rank loss. Furthermore, we know by (1.3) and (1.9) that there is a good approximate solution to the full system. This means that we can allow any numerical method that produces a solution with similar or less deviation.

Since our convergence analysis worked with the L_∞ norm on the discretized F_s spaces, a first choice would be to go for a best L_∞ approximation of the right-hand side. This means solving a linear optimization problem which minimizes η under the constraints

$$-\eta \le \sum_{m=1}^{M} \alpha_m \lambda_i^z K(z, y_m) - \lambda_i^z u(z) \le \eta,\ 1 \le i \le N, \tag{8.1}$$

where $\alpha_1, \ldots, \alpha_M$ are the other variables. If the revised simplex method is applied to the dual problem, each step has an $\mathcal{O}(M^2)$ complexity. The Kuhn–Tucker conditions ensure that one can work with at most $M+1$ active test conditions at each time. This makes the number $N >> M$ of test centers much less relevant than M, and for nicely chosen low-dimensional trial spaces one can get away with rather small computational complexity, as demonstrated in [10, 11].

But one can also try all other techniques that somehow provide a function $u_{r,s}^* \in U_r$ which by a posteriori inspection leads to a small residual norm $\delta_{r,s}$ in (1.8). This can happen to the original Kansa method when executed on a subset of M test points, or by adaptive bootstrapping techniques like the one in [10, 11] which picks suitable test centers and trial centers one by one. Other alternatives are to use pivoting with row exchange or to go for a least-squares solution first. Anyway, if the resulting residual norm $\delta_{r,s}$ is small, the result of Theorem 1.2 is still valid, proving that one actually has a good approximation to the real solution.

As an aside, we note that a simpler theory is possible if we optimize over a nondiscrete residual norm on F. Section 5 will then be obsolete, but one has to solve semi-infinite optimization problems (if F carries a sup-norm) or apply least-squares methods with integrations (if F carries an inner product). Another strategy to avoid stability problems is to add the numerically accessible quadratic constraint $\|u_r\|_K^2 \le C$ to any method trying to make residuals small. This regularization trick has connections to machine learning [17] and should be investigated in future work.

9. Ill-posed problems. For ill-posed problems, continuous dependence fails, but our method and its analysis will still be useful. We assume that the problem still has the form (1.1), but we now assume that the "true solution" $u \in U$ satisfies only

$$L(u) = f + \rho \in F, \tag{9.1}$$

where F contains the available data f and a small residual ρ. The problem $L(u) = f$ may be unsolvable, and (1.2) is not available. We consider a function $\tilde{u} \in U$ to be acceptable as a "solution" if

$$\|L(\tilde{u}) - L(u)\|_F = \|L(\tilde{u}) - f - \rho\|_F$$

is not much larger than $\|\rho\|_F$. We still assume (1.3) and (1.4), but we have to replace (1.8) by

$$\|\Pi_s(f - L(u_{r,s}^*))\|_{F_s} \le \delta_{r,s}, \tag{9.2}$$

because $L(u)$ now is unknown and does not coincide with f. Furthermore, solvability of the above system now requires

$$c(s)(\|\rho\|_F + \epsilon_r(u)) \le \delta_{r,s} \tag{9.3}$$

instead of (1.9) as a sufficient condition. The proof technique of Theorem 1.2 then still implies the following theorem.

THEOREM 9.1. *If the analytic problem is ill-posed, but solvable by $u \in U_R$ in the sense of* (9.1)*, and if we solve* (9.2) *by some $u^*_{r,s} \in U_r$, then there is a bound*

$$\|L(u - u^*_{r,s})\|_U \le c(s)\|\rho\|_F + \left(\epsilon_r(u)\left(1 + \frac{C(r,s)}{c(s)}\right) + c(s)\delta_{r,s}\right).$$

If the discretization is uniformly stable, then there is a choice of $\delta_{r,s}$ via (9.3) *such that the above residual error behaves asymptotically like the trial approximation error $\epsilon_r(u)$ plus $\|\rho\|_F$.*

Proof. We modify the proof of Theorem 1.2 to get

$$\begin{aligned}
\|L(u - u^*_{r,s})\|_F &\le \|L(u - I_r(u))\|_F + \|L(I_r(u) - u^*_{r,s})\|_F \\
&\le \epsilon_r(u) + c(s)\|\Pi_s L(I_r(u) - u^*_{r,s})\|_{F_s} \\
&\le \epsilon_r(u) + c(s)\|\Pi_s L(I_r(u) - u)\|_{F_s} \\
&\quad + c(s)\|\Pi_s(L(u) - f)\|_{F_s} \\
&\quad + c(s)\|\Pi_s(f - L(u^*_{r,s}))\|_{F_s} \\
&\le \epsilon_r(u) + c(s)\delta_{r,s} + \tfrac{C(r,s)}{c(s)}\|L(I_r(u) - u)\|_F + c(s)\|\Pi_s\|\|\rho\|_F \\
&\le c(s)\|\rho\|_F + \epsilon_r(u)\left(1 + \tfrac{C(r,s)}{c(s)}\right) + c(s)\delta_{r,s}. \qquad \square
\end{aligned}$$

For simplicity of the above presentation, we have replaced the second inequality of (1.4) by

$$c(s)\|\Pi_s g\|_{F_s} \le \|g\|_F \text{ for all } g \in F$$

which is no serious complication. However, we should comment on what happens with Theorems 1.1 and 5.1 if we have no analytic constant C_a for carrying out the proof. We replace C_a by the constant $C_a(r)$ arising in a finite-dimensional version

$$\|u_r\|_U \le C_a(r)\|L(u_r)\|_F \text{ for all } u_r \in U_r$$

of (1.2). This is feasible due to norm equivalence, but we leave it to future research to derive upper bounds for $C_a(r)$.

10. Conclusions. We provided convergence proofs for a generalized nonsquare version of Kansa's collocation method, showing that the convergence rates are determined by approximation results for nonstationary meshless kernel-based trial spaces. The rates improve with the smoothness of the solution, the domain, the differential operator, and the kernel. They hold for large classes of analytic problems, provided that there is continuous dependence on the data, and they result from a fairly general framework that possibly has applications to other unsymmetric methods. On the downside, the results still need improvement by proving better a priori inequalities to plug into the framework.

There are many possibilities for enhancement and extension of these results:

1. Find sufficient conditions for nonsingularity of square Kansa-type collocation matrices.
2. Introduce discretization-dependent weights for different parts of residuals into the theory of this paper in order to align dimension- and order-dependent convergence rates.

3. For important problems of applied analysis, state the continuous dependence of the solution on the data in precise form and derive upper bounds for the analytic constants.
4. Implement algorithms of this paper as local components of a global algorithm using localization features like domain decomposition or partitions of unity and efficiency-enhancing features like preconditioning and iterative solvers.
5. For such a global algorithm, perform large-scale numerical experiments and compare observed convergence rates with the theoretical ones of this paper.
6. Generalize all of this to unsymmetric methods for weak problems like the meshless local Petrov–Galerkin (MLPG) method of Atluri and collaborators [2].

Acknowledgments. The author thanks I. Babuška, C. S. Chen, Y. C. Hon, E. Kansa, L. Ling, G. Lube, and H. Wendland for several very stimulating discussions.

REFERENCES

[1] S. N. Atluri and T. L. Zhu, *A new meshless local Petrov–Galerkin (MLPG) approach in computational mechanics*, Comput. Mech., 22 (1998), pp. 117–127.

[2] S. N. Atluri and T. L. Zhu, *A new meshless local Petrov–Galerkin (MLPG) approach to nonlinear problems in computer modeling and simulation*, Computer Modeling and Simulation in Engineering, 3 (1998), pp. 187–196.

[3] J. P. Aubin, *Approximation of Elliptic Boundary-Value Problems*, Pure Appl. Math., 26, Wiley-Interscience, New York, 1972.

[4] I. Babuška, U. Banerjee, and J. Osborn, *Survey of meshless and generalized finite element methods: A unified approach*, in Acta Numerica 2003, Acta Numer. 12, Cambridge University Press, Cambridge, UK, 2003, pp. 1–215.

[5] D. Brown, L. Ling, E. Kansa, and J. Levesley, *On approximate cardinal preconditioning methods for solving PDEs with radial basis functions*, Engineering Analysis with Boundary Elements, 29 (2005), pp. 343–353.

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

[7] Y. C. Hon and R. Schaback, *On unsymmetric collocation by radial basis functions*, Appl. Math. Comp., 119 (2001), pp. 177–186.

[8] E. J. Kansa, *Application of Hardy's multiquadric interpolation to hydrodynamics*, in Proceedings of the 1986 Annual Simulations Conference, Vol. 4, San Diego, CA, 1986, pp. 111–117.

[9] L. Ling and E. Kansa, *Preconditioning for radial basis functions with domain decomposition methods*, Math. Comput. Modelling, 40 (2005), pp. 1413–1427.

[10] L. L. Ling, R. Opfer, and R. Schaback, *Results on meshless collocation techniques*, Engineering Analysis with Boundary Elements, 30 (2006), pp. 247–253.

[11] L. Ling and R. Schaback, *On adaptive unsymmetric meshless collocation*, in Proceedings of the 2004 Annual International Conference on Computational and Experimental Engineering and Sciences, Advances in Computational and Experimental Engineering and Sciences, S. N. Atluri and A. J. B. Tadeu, eds., Tech Science Press, Forsyth, GA, 2004, CD-Rom edition, paper # 270.

[12] J. L. Lions and E. Magenes, *Problèmes aux limites non homogènes at applications,* Travaux et Recherches Mathématiques 1, Dunod, Paris, 1968.

[13] F. J. Narcowich, J. D. Ward, and H. Wendland, *Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting*, Math. Comp., 74 (2005), pp. 743–763.

[14] F. J. Narcowich, J. D. Ward, and H. Wendland, *Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions*, Constr. Approx., 24 (2006), pp. 175–186.

[15] R. Schaback and H. Wendland, *Inverse and saturation theorems for radial basis function interpolation*, Math. Comp., 71 (1998), pp. 669–681.

[16] R. Schaback and H. Wendland, *Characterization and construction of radial basis functions*, in Multivariate Approximation and Applications, N. Dyn, D. Leviatan, D. Levin, and A. Pinkus, eds., Cambridge University Press, Cambridge, UK, 2001, pp. 1–24.

[17] R. Schaback and H. Wendland, *Kernel techniques: From machine learning to meshless methods*, in Acta Numerica 2006, Acta Numer. 15, Cambridge University Press, Cambridge, UK, 2006, pp. 543–639.

[18] H. Wendland, *Piecewise polynomial, positive definite, and compactly supported radial functions of minimal degree*, Adv. Comput. Math., 4 (1995), pp. 389–396.

[19] H. Wendland, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, UK, 2005.

[20] H. Wendland and C. Rieger, *Approximate interpolation with applications to selecting smoothing parameters*, Numer. Math., 101 (2005), pp. 729–748.