All Well–Posed Problems have Uniformly Stable and Convergent Discretizations

Robert Schaback

Abstract This paper considers a large class of linear operator equations, including linear boundary value problems for partial differential equations, and treats them as linear recovery problems for *functions* from their *data*. *Well-posedness* of the problem means that this recovery is continuous. *Discretization* recovers restricted *trial* functions from restricted *test* data, and it is well-posed or *stable*, if this *restricted* recovery is continuous. After defining a general framework for these notions, this paper proves that all well-posed linear problems have stable and refinable computational discretizations with a stability that is determined by the well-posedness of the problem and independent of the computational discretization, provided that sufficiently many test data are used. The solutions of discretized problems converge when enlarging the trial spaces, and the convergence rate is determined by how well the data of the function solving the analytic problem can be approximated by the data of the trial functions. This allows new and very simple proofs of convergence rates for generalized finite elements, symmetric and unsymmetric Kansa-type collocation, and other meshfree methods like Meshless Local Petrov–Galerkin techniques. It is also shown that for a fixed trial space, weak formulations have a slightly better convergence rate than strong formulations, but at the expense of numerical integration. Since convergence rates are reduced to those coming from Approximation Theory, and since trial spaces are arbitrary, this also covers various spectral and pseudospectral methods. All of this is illustrated by examples.

Keywords Partial Differential Equations · meshless methods · collocation · MLPG · Petrov-Gelerkin methods · discretization · error bounds · well-posedness · stability · convergence · recovery · residual minimization · nodal bases · Kansa method · interpolation · approximation

Robert Schaback
Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestrae 16-18
D-37083 Göttingen
E-mail: schaback@math.uni-goettingen.de

Mathematics Subject Classification (2000) 65M12 · 65M70 · 65N12 · 65N35 · 65M15 · 65M22 · 65J10 · 65J20 · 35D30 · 35D35 · 35B65 · 41A25 · 41A63

1 Introduction and Summary

This paper focuses on mathematical problems that have solutions u in some normed linear space U over $\mathbb{R}$ satisfying infinitely many linear conditions that we write as

$$\lambda(u) = f_\lambda \text{ for all } \lambda \in \Lambda \subseteq U^* \tag{1}$$

with given real numbers f_λ and continuous linear functionals λ on U collected into a set $\Lambda \subset U^*$. We call the real numbers $\{f_\lambda\}_{\lambda \in \Lambda}$ the *data* that hopefully allow to identify the *object* u, which will in most cases be a multivariate function on some domain. Solving (1) for u from given data $\{f_\lambda\}_{\lambda \in \Lambda}$ is a *recovery problem*, and we view it as posed in an abstract mathematical setting that is not directly accessible for computation. It can be called an *analytical* problem in contrast to the *computational* problems that will be introduced in Section 7. The transition from an analytical problem to a computational problem will be called *discretization*, and we deal with this in detail in Section 4.

Typical special cases arise when solving partial differential equations (PDEs). The function u to be recovered is always an element of some space U of real-valued functions on a domain Ω, but weak and strong formulations of PDEs use very different types of *data*, namely either integrals against test functions or derivative values at evaluation points, plus boundary conditions of various forms. If two problems use different data to identify the same function, we consider the problems as different here. In particular, weak and strong PDE formulations for the same equation are *different* problems. For illustration, see Example 1 and sections 13.3 and 13.4 below.

The PDE applications of (1) suggest to view the evaluation of the functionals $\lambda \in \Lambda$ on u as *testing* the *trial* function u. *Discretization* will then fix a finite-dimensional *trial subspace* $U_M \subset U$ and a finite set Λ_N of *test functionals* from Λ. We pursue this distinction between the trial and the test side of (1) throughout this paper.

If a problem in *Mathematical* Analysis is well-posed, it should have a discretization in *Numerical* Analysis that is also well-posed. This requires to derive some sort of numerical stability of well-chosen discretizations from the well-posedness property of the underlying analytical problem. This paper proves the above statement under mild additional assumptions after stating clearly in Section 2 and 4 what is to be understood by *well-posedness* of a *problem* and its *discretization*. It turns out in Section 5 that one can choose refinable discretizations that have stability properties depending only on the well-posedness of the given problem, not on the discretizations chosen. This depends crucially on what we call a *monotone refinable dense (MRD)* discretization strategy in Section 4.

Note that the classical theory concluding convergence from consistency and stability usually assumes solvability of the arising square linear systems. Since this does not hold for meshfree methods in general [23], a new discretization theory is needed that deals with overdetermined linear systems and introduces suitable stability concepts. Otherwise, unsymmetric collocation or the MLPG method cannot

be treated. Our older discretization theory [42, 43, 11] served that purpose. However, error bounds and convergence results depended there on *stability inequalities* that needed complicated proofs [39], while this paper shows that one can always enforce uniform stability by sufficiently thorough testing, provided that the analytic problem is well-posed.

The resulting discretized linear problems will be overdetermined due to this stabilization, and should be solved approximately by minimizing residuals. Section 7 deals with this, and shows that the final error bounds and convergence rates are determined by how well the data of the true solution can be approximated by the data of elements of the trial space. We call this *Trial Space Data Approximation*, and it can be seen as a variation of *consistency*. In particular, error bounds and convergence rates are independent of the details of testing, provided that testing is done sufficiently thoroughly, in the sense of Section 5.

Summarizing, the link of this paper to the classical notions of consistency, stability, and convergence is that here uniform stability is a proven consequence of well-posedness of the underlying analytical problem, and then convergence follows from consistency. The latter is reformulated here in terms of the *Trial Space Data Approximation*.

Section 8 extends the previous results to ill-posed problems and noisy data, while Section 9 and 10 specialize to recovery in Hilbert spaces, where uniformly stable and sometimes optimal discretizations are readily available. These generalize the Ritz-Galerkin technique for weak problems and symmetric collocation for strong problems, as will be explained in Section 13.2 when it comes to examples.

Stability can be spoiled by bad bases. Therefore this paper ignores bases and focuses on spaces instead, with the exception of Section 11 where the influence of bases on the trial and the test side is studied. A very common class of bases are the *nodal* bases used in classical piecewise linear finite elements and various meshfree methods. They parametrize trial spaces of functions entirely in terms of function values at nodes [10]. Many application papers report good stability properties of these bases, and Section 12 provides a fairly general mathematical proof, showing that convergence in the nodes can be derived from convergence of the *Trial Space Data Approximation*.

Reducing the error analysis of PDE solvers to the analysis of approximation errors is not new. The standard situation in finite elements does this via a lemma of "Cea type" that allows to bound the error of a PDE solution by the error of its approximation from the trial space. In various papers on meshfree methods, e.g. [22, 35, 25] the error analysis of PDE solving is done that way, namely considering errors of interpolants or approximations, and stability is not addressed at all or in a different sense [9]. But in general one needs additional arguments to justify working solely with approximation errors, and this paper bridges the gap in quite some generality.

The paper closes with a number of examples that apply the above theory:

1. Polynomial interpolation in Section 13.1 shows that the stabilization results of this paper may imply quite some *overtesting* , i.e. oversampling on the test side to guarantee uniform stability. Furthermore, it points out how spectral methods are covered and why weak formulations yield slightly faster convergence than

    strong formulations, though for weaker norms and at the expense of numerical integration.

2. Section 13.2 deals with the standard setting for finite elements for homogeneous boundary conditions, showing that it fits perfectly into the framework, including extension to other trial spaces and a Petrov–Galerkin treatment.

3. The remaining examples address the standard Poisson problem with Dirichlet boundary conditions, for simplicity.

   (a) Section 13.3 focuses on collocation as a typical strong problem. This covers various kinds of meshfree methods, including Kansa's unsymmetric collocation, and it is shown how to derive specific convergence rates depending on the trial spaces chosen.

   (b) The weak Dirichlet case is handled in Section 13.4, and a comparison of convergence rates for the strong and weak formulations, using the same trial spaces, is provided in Section 13.5.

   (c) Finally, the Meshless Local Petrov Galerkin (MLPG) scheme [5] is treated in Section 13.6. This includes error bounds and convergence rates for different variations of the method, but it was necessary to include a first proof of well-posedness of the local weak problem behind MLPG.

   In the cases (a) and (c) we get a new and much simpler approach to error analysis and convergence, and this approach is applicable to a large class of meshfree methods.

4. In addition, Section 14 provides a numerical example that shows how overtesting improves stability, and two examples checking the theoretical convergence rates for unsymmetric collocation of a Dirichlet problem in strong form.

On the downside, the test strategies guaranteeing uniform stability are only shown to exist, they are not constructed. Future work needs explicit sufficient conditions on specific test strategies to guarantee uniform stability. An *adaptive* improvement of stability can be done by *greedy testing* as touched in earlier papers on adaptivity [47, 24, 33, 45].

    Finally, emphasis so far is only on errors, convergence rates, and stability of algorithms, but not on computational efficiency. It is a major challenge to relate the achievable convergence rates and stability properties to computational efficiency. Anyway, this paper provides a starting point towards these goals.

2 Problems, Data, and Well-Posedness

Behind (1) there is a linear *data map* $D : U \mapsto \mathbb{R}^\Lambda =: V$ that takes each $u \in U$ into the set of values $\{\lambda(u)\}_{\lambda \in \Lambda}$ in the *data space* V. This allows to rewrite (1) as an operator equation

$$D(u) = f \tag{2}$$

for some given f in the data space V. Each operator equation can be formally interpreted this way, e.g. by defining Λ as the set of all functionals $\mu \circ D$ when μ varies in the unit sphere of V^*.

For the rest of this paper, we assume existence of a solution u^* to (2), i.e. the data $f = D(u^*)$ are assumed to be in the range of the data map. We shall relax this assumption in Section 8.

Example 1 As an illustration, consider the standard Dirichlet problem

$$
\begin{aligned}
-\Delta u = f & \quad \text{in } \Omega \subset \mathbb{R}^d \\
u = g & \quad \text{in } \Gamma := \partial\Omega
\end{aligned}
\tag{3}
$$

where f and g are given functions on Ω and Γ. A *strong* formulation writes it in the form (1) with functionals

$$
\begin{aligned}
\lambda_x &: u \mapsto -\Delta u(x), \; x \in \Omega \\
\mu_y &: u \mapsto u(y), \qquad y \in \Gamma
\end{aligned}
$$

on some space U where both types of functionals are continuous. The standard FEM algorithms use *weak* functionals

$$
\lambda_v : u \mapsto (\nabla u, \nabla v)_{L_2(\Omega)} \text{ for all } v \in H_0^1(\Omega)
$$

and add the functionals μ_y for points on the boundary. We postpone further details to Section 13.3 and 13.4, but remark that the data maps differ considerably. Consequently, the strong and weak reformulations of (3) are two *fundamentally different* analytic problems in the sense of this paper. □

We give the data space a norm structure by requiring that

$$
\|Du\|_V := \sup_{\lambda \in \Lambda} |\lambda(u)| \text{ for all } u \in U
\tag{4}
$$

is a norm on $D(U)$ that we assume to be extended to V, if not $V = D(U)$. We shall call this the *data norm*, and note that it leads to a seminorm

$$
\|u\|_D := \|Du\|_V = \sup_{\lambda \in \Lambda} |\lambda(u)| \text{ for all } u \in U
\tag{5}
$$

on the function space U. This is well-defined if all functionals in Λ are uniformly bounded. We assume existence of the data norm from now on, but remind the reader that renormalization of functionals changes the data norm and all issues depending on it, like the well-posedness conditions that we introduce later.

Definition 1 An *analytic problem* in the sense of this paper consists of

1. a linear normed *function* space U,
2. a set Λ of linear continuous functionals on U leading to a *data map D* as in (2)
3. with values in a normed *data space V* such that
4. (4) holds and is a norm on V.

Unique solvability of the problem (1) or (2) requires that $u \in U$ vanishes if all data $\lambda(u)$ for all $\lambda \in \Lambda$ vanish, or that D is injective, or that $\|.\|_D$ is a norm. A somewhat stronger and quantitative notion is *well-posedness*:

Definition 2 An analytic problem in the sense of Definition 1 is *well-posed with respect to a well-posedness norm* $\|.\|_{WP}$ on U if there is a constant C such that a *well-posedness inequality*

$$\|u\|_{WP} \leq C\|Du\|_V = C\|u\|_D \text{ for all } u \in U \tag{6}$$

holds.

This means that D^{-1} is continuous as a map $D(U) \to U$ in the norms $\|.\|_V$ and $\|.\|_{WP}$, and (6) thus reformulates the standard notion of well–posedness as introduced by Hadamard. The well-posedness norm $\|.\|_{WP}$ on U will often be weaker than the norm $\|.\|_U$ on U needed to let the data be well-defined.

Example 2 In the context of Example 1, the strong problem leads to well-posedness with $\|.\|_{WP} = \|.\|_{\infty,\Omega}$ due to the inequality

$$\|u\|_{\infty,\Omega} \leq \|u\|_{\infty,\partial\Omega} + C\|\Delta u\|_{\infty,\Omega} \text{ for all } u \in U := C^2(\Omega) \cap C(\overline{\Omega}), \tag{7}$$

see e.g. [12, (2.3), p. 14]. In contrast to this, the boundary-homogeneous weak problem has $\|.\|_{WP} = \|.\|_{L_2(\Omega)}$ due to the Poincaré-Friedrichs inequality

$$\|u\|_{L_2(\Omega)} \leq C\|\nabla u\|_{L_2(\Omega)} \text{ for all } u \in U := H_0^1(\Omega), \tag{8}$$

see e.g. [12, 1.5, p. 30].

Details will follow in Section 13.3 and 13.4, but we remark here that deriving computationally useful well-posedness inequalities is a serious issue that is not satisfactorily addressed by theoreticians, because they do not use computationally useful norms on the data space. For instance, the continuous dependence of solutions of elliptic problems on the boundary data is often expressed by taking Sobolev trace spaces of fractional order on the boundary, and these spaces are far from being accessible for computation. The examples in Section 13 will shed some light on this issue.

Future research in Applied Mathematics should target practically useful well-posedness results based on norms that are closer to computation.

3 Trial Space Data Approximation

We now perform the first step of *discretization* by choosing a finite-dimensional *trial space* $U_M \subset U$. This allows us to approximate the data $f = D(u^*) \in V$ by data $D(u_M)$ for all trial elements $u_M \in U_M$ in the data norm $\|.\|_V$. From Approximation Theory we require an error bound

$$\|Du^* - Du_M^*\|_V \leq \varepsilon(u^*, U_M, D) \tag{9}$$

for some approximation $u_M^* \in U_M$ that we call a *comparison trial function* in what follows. It may be the best data approximation in the sense

$$\|Du^* - Du_M^*\|_V = \min_{u_M \in U_M} \|Du^* - Du_M\|_V,$$

but in most cases one has a standard approximation process in a norm on U with

$$\|u^* - u_M^*\| \leq \varepsilon(u^*, U_M)$$

and applies the data map afterwards. We shall rely on Approximation Theory to provide the upper bounds $\varepsilon(u^*, U_M, D)$ in (9) and convergence rates for $\varepsilon(u^*, U_M, D) \to 0$ if the spaces U_M get larger and larger. These rates will crucially depend on the smoothness of u^*, the trial spaces U_M, and the data map D. For trial spaces in spectral methods, these convergence rates may be very large, and there may even be exponential convergence. We call (9) the *Trial Space Data Approximation*, but we shall never calculate the comparison trial function u_M^*. We just obtain it from Approximation Theory as an existing good approximation to the true solution u^* including as many derivatives as necessary for the data map, and we later compare our approximate solutions of (2) to it.

Theorem 1 *If the analytic problem is well-posed in the sense of Definition 2, the error bounds and convergence rates of the* Trial Space Data Approximation *immediately carry over to error bounds and convergence rates in the well-posedness norm via*

$$\|u^* - u_M^*\|_{WP} \leq C\|Du^* - Du_M^*\|_V \leq C\varepsilon(u^*, U_M, D),$$

and the constant C is independent of the chosen trial space.

This means that Approximation Theory provides convergence rates for certain approximate solutions of certain well-posed analytic problems, but these approximate solutions are computationally inaccessible.

Example 3 In the context of Example 1, the functions of the trial space have to approximate function values on the boundary in both the strong and the weak case. But for the strong form we have to approximate second derivatives, while the weak form only has to approximate first derivatives. Furthermore, the well-posedness norms are different. This will lead to different convergence rates for strong and weak formulations in Section 13.5.

4 MRD Discretizations

We now show how to discretize the test side of an analytic problem in the sense of Definition 1 in such a way that a uniformly stable and finite computational strategy exists that provides approximations $\tilde{u}_M \in U_M$ with

$$\|u^* - \tilde{u}_M\|_{WP} \leq 2\|u^* - u_M^*\|_{WP} \leq 2C\|Du^* - Du_M^*\|_V \leq 2C\varepsilon(u^*, U_M, D).$$

This implies that Approximation Theory will yield convergence rates for certain *finitely and stably computable* approximate solutions of certain well-posed analytic problems. The convergence will take place in U under the well-posedness norm $\|.\|_{WP}$, and the convergence rate will be the convergence rate of the *Trial Space Data Approximation*. Our main tool will be a *monotonic refinable dense (MRD)* discretization of the data space V that we describe now.

No matter what the data *map* is, the data *space V* should allow some form of *discretization* for computational purposes. We model this by *restriction maps*

$$R_N \, : \, \{f_\lambda\}_{\lambda \in \Lambda} \mapsto \{f_\lambda\}_{\lambda \in \Lambda_N} \in V_N = \mathbb{R}^{|\Lambda_N|}$$

that map V into finite-dimensional *data spaces* V_N over $\mathbb{R}$. The discretizations use *restricted data* belonging to finite subsets Λ_N of Λ, and these data enter practical computation.

On the spaces V_N we introduce the norm

$$\|R_N\{f_\lambda\}_{\lambda \in \Lambda}\|_{V_N} = \|\{f_\lambda\}_{\lambda \in \Lambda_N}\|_{V_N} := \max_{\lambda \in \Lambda_N} |f_\lambda|$$

and we get the *monotonicity* property

$$\|R_M v\|_{V_M} \leq \|R_N v\|_{V_N} \text{ for all } v \in V \text{ and all } \Lambda_M \subseteq \Lambda_N.$$

Refinement of two discretizations defined by sets Λ_M and Λ_N works by taking the union $\Lambda_M \cup \Lambda_N$, and by the monotonicity property this will weakly increase the discrete norms. Finally, we have

$$\|v\|_V := \sup_{R_N, V_N} \|R_N v\|_{V_N} \text{ for all } v \in V, \tag{10}$$

following from (4).

But there are applications where restrictions are not defined by taking *all possible* finite subsets of functionals. They might require background triangulations, e.g. for finite elements, and their refinement does not simply involve taking a union of two finite sets of functionals. We can generalize the above notions by ignoring functionals:

Definition 3 An *MRD discretization* of a data space V consists of a set of pairs (R_N, V_N) called *restrictions* with the properties

1. V_N is a normed linear space with $\dim V_N < \infty$ and norm $\|.\|_{V_N}$,
2. $R_N \, : V \to V_N$ is linear,
3. there is a partially defined *refinement relation* $\preceq$ on the restrictions such that
4. $(R_M, V_M) \preceq (R_N, V_N)$ implies $\|R_M v\|_{V_M} \leq \|R_N v\|_{V_N}$ for all $v \in V$,
5. for each two admissible restrictions (R_M, V_M), (R_N, V_N) there is a restriction (R_P, V_P) such that $(R_M, V_M) \preceq (R_P, V_P)$ and $(R_M, V_M) \preceq (R_P, V_P)$,
6. (10) is a norm, when the sup is taken over all admissible restrictions.

This axiomatic framework is open for further discussion, of course, but we assume it in what follows. We refer to the last three properties as *monotonicity*, *refinement*, and *density*, using the term *MRD discretization* for all six properties. Note that the norm arising in the density property must be the data norm that is used in the well-posedness inequality (6).

The discussion preceding Definition 3 proved

Theorem 2 *Each analytical problem of the form (1) in the sense of Definition 1 has a MRD discretization via taking finite subsets of functionals.* $\qquad\square$

For Example 1, it is clear that one can focus on finitely many functionals when it comes to finite computations, but it is by no means clear which and how many are to be taken to allow a uniformly stable computational method. The refinement in the FEM case is not quite standard, but will satisfy Definition 3, because it still uses finite subsets of functionals.

5 Well-Posedness of Discretized Problems

If we use a MRD restriction (R_N, V_N) on the data together with a chosen trial space U_M, we can pose the *discretized problem* as the linear system

$$R_N D u_M = R_N D u^* \tag{11}$$

to be solved for $u_M \in U_M$, where the computational input data are provided by the restriction $R_N D u^*$ of the data of an exact solution u^*. Such systems will usually be overdetermined.

Since the well-posedness condition (6) also holds on the trial space, the discretized problem (11) is automatically well-posed or *stable* in the sense

$$\|u_M\|_{WP} \leq C(U_M, V_N) \|R_N D u_M\|_{V_N} \text{ for all } u_M \in U_M$$

if we can prove

$$\|D u_M\|_V \leq C(U_M, V_N) \|R_N D u_M\|_{V_N} \text{ for all } u_M \in U_M \tag{12}$$

for some *stability constant* $C(U_M, V_N)$. Similar *stability inequalities* were already treated in [42] for proving convergence of unsymmetric collocation and other methods [39].

We now can state our central result, to be proven later in a somewhat more general form.

Theorem 3 *Assume an analytic problem (2) with an MRD discretization. If U_M is an arbitrary finite-dimensional subspace of U, there always is a restriction (R_N, V_N) such that*

$$\|u_M\|_D \leq 2\|R_N D(u_M)\|_{V_N} \text{ for all } u_M \in U_M. \tag{13}$$

This holds without assuming well-posedness. If the latter is assumed by (6), we have

$$\|u_M\|_{WP} \leq 2C\|R_N D(u_M)\|_{V_N} \text{ for all } u_M \in U_M \tag{14}$$

with the constant C from (6).

In contrast to (12), the constants in (13) and (14) are independent of U_M and V_M, proving a *uniform* well-posedness or stability of the discretized problem for a rather sensible choice of V_N after an arbitrary selection of U_M. Section 13.1 will show that this uniformity may require some hidden amount of *oversampling*, i.e. the dimension of V_N may be much larger than the dimension of U_M. We call this *overtesting*, because it occurs on the test side of the problem. Theorem 3 does not give any practical hints how to care for uniformly stable testing, it just proves existence. The necessary amount of overtesting to achieve uniform stability is left open.

6 Well-Posedness of Data Discretizations

Inspection of (13) shows that the analytic problem and its well-posedness are not relevant for (13), because the actual well-posedness condition (6) enters only into the trivial transition from (13) to (14). In fact, everything follows already from the notion of a *monotone refinable dense* (MRD) discretization. Well-posedness is a later add-on.

Lemma 1 *Consider a data space V and associated MRD restrictions (R_N, V_N) satisfying the assumptions of Section 4. Then for each finite-dimensional subspace W_M of V there always is a restriction (R_N, V_N) such that*

$$\|w_M\|_V \leq 2\|R_N w_M\|_{V_N} \text{ for all } w_M \in W_M.$$

Proof: Define $K \subset W_M$ as the unit sphere of W_M defined via the norm $\|.\|_V$. By compactness, for each $\varepsilon > 0$ we can cover K by finitely many ε-neighborhoods

$$U_\varepsilon(y_j) := \{y \in K \ : \ \|y - y_j\|_V \leq \varepsilon\}, 1 \leq j \leq n$$

with elements $y_1, \ldots, y_n \in K$. By the density property (10) we can find restrictions $R_{N_1}, \ldots, R_{N_n}$ with associated spaces $V_{N_1}, \ldots, V_{N_n}$ such that

$$\|y_j\|_V \leq \|R_{N_j} y_j\|_{V_{N_j}} + \varepsilon, 1 \leq j \leq n$$

and by repeated application of the refinement property we can define R_N and V_N as the "union" of these, and then

$$\|R_{N_j} v\|_{V_{N_j}} \leq \|R_N v\|_{V_N} \text{ for all } v \in V, 1 \leq j \leq n$$

by monotonicity.

We now take an arbitrary $w_M \in K$ and get some j, $1 \leq j \leq n$ with $\|w_M - y_j\|_V \leq \varepsilon$ via the covering. This implies $\|R_N w_M - R_N y_j\|_{V_N} \leq \varepsilon$ by the density property, and then

$$\begin{aligned}
\|R_N w_M\|_{V_N} &\geq \|R_N y_j\|_{V_N} - \varepsilon \\
&\geq \|R_{N_j} y_j\|_{V_{N_j}} - \varepsilon \\
&\geq \|y_j\|_V - 2\varepsilon \\
&\geq \|w_M\|_V - 3\varepsilon \\
&= 1 - 3\varepsilon
\end{aligned}$$

proving

$$\|R_N w_M\|_{V_N} \geq (1 - 3\varepsilon)\|w_M\|_V$$

for all $w_M \in W_M$, and the assertion follows for $\varepsilon = 1/6$.

The proof of Theorem 3 now follows by setting $W_M = D(U_M)$ with an arbitrary data map D. $\qquad \square$

If the space W_M in the proof has dimension M, Lemma 1 says nothing about the dimension of V_N. The proof uses a covering argument, not a dimension argument. It uses Topology, not Linear Algebra. See Section 13.1 for an application where $N \geq M$ ranges from $N = M$ to $N = \mathcal{O}(M^2)$.

7 Solving Discretized Problems

After choosing a trial space U_M and getting a suitable data restriction (R_N, V_N) for Theorem 3, the discretized recovery problem (11) requires computation of some $u_M \in U_M$ from the data $R_N D u^*$, where u^* is the true solution to the analytical problem. This will usually lead to an overdetermined linear system after choosing bases, but we do not want to consider bases unless absolutely necessary.

The simplest basis-free computational method we can propose is to minimize the residual norm $\|R_N D(u^* - u_M)\|_{V_N}$ over all $u_M \in U_M$, which is a finite-dimensional approximation problem. But this would not nicely connect to Approximation Theory. Instead, we refer to the *comparison trial approximation* u_M^* from (9) of the solution u^*, and we are satisfied with computing an element $\tilde{u}_M \in U_M$ with

$$\|R_N D(u^* - \tilde{u}_M)\|_{V_N} \leq C_A \|R_N D(u^* - u_M^*)\|_{V_N}, \tag{15}$$

however it is calculated, with a fixed constant $C_A \geq 1$ that makes computational life easier when chosen not too close to one. Due to the monotonicity and density properties, we also have

$$\|R_N D(u^* - \tilde{u}_M)\|_{V_N} \leq C_A \|R_N D(u^* - u_M^*)\|_{V_N} \leq C_A \|D(u^* - u_M^*)\|_V \leq C_A \varepsilon(u^*, U_M, D),$$

such that the error of the *Trial Space Data Approximation* in (9) is always an upper bound, up to a uniformly bounded factor.

Now (15) implies

$$\begin{aligned}
\|\tilde{u}_M - u_M^*\|_D &\leq 2\|R_N D(\tilde{u}_M - u_M^*)\|_{V_N} \\
&\leq 2\|R_N D(\tilde{u}_M - u^*)\|_{V_N} + 2\|R_N D(u^* - u_M^*)\|_{V_N} \\
&\leq (2C_A + 2)\|R_N D(u^* - u_M^*)\|_{V_N} \\
&\leq (2C_A + 2)\|D(u^* - u_M^*)\|_V \\
&= (2C_A + 2)\varepsilon(u^*, U_M, D)
\end{aligned}$$

and

$$\begin{aligned}
\|\tilde{u}_M - u^*\|_D &\leq \|\tilde{u}_M - u_M^*\|_D + \|u_M^* - u^*\|_D \\
&\leq (2C_A + 3)\varepsilon(u^*, U_M, D),
\end{aligned}$$

proving that the error of the computational solution $\tilde{u}_M$ is up to a factor the same as the error of the comparison trial function u_M^*, evaluated in the data norm.

Theorem 4 *Assume an MRD discretization of an analytic problem along the lines of the previous sections. Then each computational technique to solve the discretized problem approximatively by an element $\tilde{u}_M \in U_M$ such that (15) holds, will guarantee*

$$\|\tilde{u}_M - u^*\|_D \leq (2C_A + 3)\varepsilon(u^*, U_M, D).$$

Corollary 1 *Adding well-posedness to Theorem 4 yields*

$$\|\tilde{u}_M - u^*\|_{WP} \leq C(2C_A + 3)\varepsilon(u^*, U_M, D)$$

proving that convergence rates in the data norm transfer to the same convergence rates in the norm $\|.\|_{WP}$ on U arising in the well-posedness condition (4). $\square$

We summarize what we have so far, for easy reference in the examples.

Theorem 5 *Assume a well-posed analytic problem with an MRD discretization as in Definitions 1, 2, and 3. Then for arbitrary trial spaces one can choose uniformly stable test discretizations to get uniformly stable computational methods based on some form of residual minimization. The convergence rates, measured in the well-posedness norm, are given by the convergence rate of the* Trial Space Data Approximation, *i.e. the rate in which the data of the true solution are approximated by the data of comparison trial functions, measured in the data norm.* □

This will be applied in the following way. First, one assumes additional regularity of the solution function u^* and fixes a well-known approximation process in U that provides good comparison trial functions u_M^* for these trial spaces, and with a very good convergence rate $u_M^* \to u^*$ that may even be spectral in a weak norm like $\|.\|_{WP}$. Then these approximations are used for comparison in the above theory, and the convergence rate in the data norm is calculated from what is known about the approximation process. Then we know that this rate is the one that arises when solving the analytic problem, and it arises in the well-posedness norm. This may even yield spectral convergence, and we shall provide examples. But note that the rate of convergence of our discretized solutions of the analytical problems is only the rate the convergence obtained *after* the data map is applied, and it involves the norm $\|.\|_{WP}$ used in the well-posedness condition.

The above approach applies to a large variety of well-posed analytic problems, and shows that for properly chosen scales of trial spaces U_M and properly chosen test strategies depending on each U_M one gets uniformly stable and convergent computational methods with convergence rates that can be derived from results of Approximation Theory. These rates normally improve with the smoothness of the true solution, but they also depend on the data map and the well-posedness norm. For a given PDE problem like in Example 1, the convergence rates of strong and weak formulations will be different, even if the trial spaces are the same. This is due to the fact that the data maps, data norms, and well-posedness norms are different. Details will follow in Section 13.5.

If the true solution necessarily has certain singularities of a known type, like in elliptic PDE problems on domains with incoming corners, one should always add the correct singular functions to the trial space. Then the approximation quality of the singular solution in the augmented trial space is the same as the approximation quality of a regular solution in the original trial space, and this quality will improve with the smoothness of the regular solution. In this sense, going over to extended trial spaces like in the XFEM or GFEM [7,1,48,8] does not need a change in our theory. But we bypass a complicated stability analysis by requiring an unspecified amount of overtesting here.

Practitioners usually do not care about stability or well-posedness of discretized problems. They go for square stiffness matrices in meshless methods though it is known that those might be singular [23], and all related instability risks are ignored. The upshot of this paper is that sufficient testing brings uniform stability, at the price of non-square overdetermined systems. But it is questionable whether hardcore practitioners will change their hazardous computations.

8 Noisy Data

Corollary 1 showed that ill-posed problems can be treated, if one is satisfied with reproducing the data well. But so far we always have assumed that the input data are given exactly as data of an existing solution. If this is not true, a few changes are necessary. We assume that the data map D is always exact, but the input data for computations are assumed to be polluted by either noise or errors in evaluating the data functionals. This also handles the error committed by numerical integration when the data functionals of weak PDE problems are considered.

The data now consist of a general element v^* of the data space V, and we assume that there is an function $u^* \in U$ such that $\|D(u^*) - v^*\|_V$ is small, and we want to recover this function well, or others with a similarly good data reproduction. We choose a trial space $U_M \subset U$ as before, and we define $W_M := D(U_M) \subset V$ as at the end of Section 6. Clearly, (15) now has to be replaced by

$$\|R_N v^* - R_N D\tilde{u}_M\|_{V_N} \le C_A \|R_N v^* - R_N Du_M^*\|_{V_N}, \tag{16}$$

because there are no other data at hand. Then

Theorem 6 *Assume an MRD discretization of an analytic problem along the lines of the previous sections. Then each computational technique to solve the discretized problem approximatively by an element $\tilde{u}_M \in U_M$ such that (16) holds, will guarantee*

$$\|\tilde{u}_M - u^*\|_D \le (2C_A + 3)\varepsilon(u^*, U_M, D) + (2C_A + 2)\|v^* - Du^*\|_V \tag{17}$$

for any $u^* \in U$.

Proof. We proceed like above, via

$$
\begin{aligned}
\|\tilde{u}_M - u_M^*\|_D &\le 2\|R_N D(\tilde{u}_M - u_M^*)\|_{V_N} \\
&\le 2\|R_N D\tilde{u}_M - R_N v^*\|_{V_N} + 2\|R_N v^* - R_N Du_M^*\|_{V_N} \\
&\le (2C_A + 2)\|R_N v^* - R_N Du_M^*\|_{V_N} \\
&\le (2C_A + 2)\left(\|R_N v^* - R_N Du^*\|_{V_N} + \|R_N Du^* - R_N Du_M^*\|_{V_N}\right) \\
&\le (2C_A + 2)\left(\|v^* - Du^*\|_V + \|D(u^* - u_M^*)\|_V\right) \\
&\le (2C_A + 2)\left(\|v^* - Du^*\|_V + \varepsilon(u^*, U_M, D)\right)
\end{aligned}
$$

and get

$$
\begin{aligned}
\|\tilde{u}_M - u^*\|_D &\le \|\tilde{u}_M - u_M^*\|_D + \|u_M^* - u^*\|_D \\
&\le (2C_A + 3)\varepsilon(u^*, U_M, D) + (2C_A + 2)\|v^* - Du^*\|_V.
\end{aligned}
$$

The inequality (17) shows that errors in the data values, e.g. integration errors for weak data, can spoil the convergence unless they are at least as small as the error committed by the comparison function u_M^* in the data norm. For trial spaces that allow fast convergence, the admissible errors in the data functionals are severely restricted by this observation. This generalizes the well-known fact that high-order solvers of weak problems require highly accurate integration.

One can go into (16) by choosing u_M^* as the minimizer of $\|v^* - D(u_M)\|_V$ over all $u_M \in U_M$. Then

$$\|R_N v^* - R_N D(\tilde{u}_M)\|_{V_N} \le C_A \|R_N v^* - R_N D(u_M^*)\|_{V_N} \le C_A \|v^* - Du_M^*\|_V,$$

and the proof of Theorem 6 yields

$$\begin{aligned}
\|\tilde{u}_M - u_M^*\|_D &\le (2C_A + 2)\|R_N v^* - R_N D u_M^*)\|_{V_N} \\
&\le (2C_A + 2)\|v^* - D u_M^*)\|_V, \\
\|v^* - D\tilde{u}_M\|_D &\le \|v^* - D u_M^*\|_D + \|D u_M^* - D\tilde{u}_M\|_D \\
&\le (2C_A + 3)\|v^* - D u_M^*\|_D.
\end{aligned}$$

Corollary 2 *Assume an analytic problem that has a MRD discretization without being well-posed, and assume that the given data do not necessarily come from some solution u^*. Then there is a uniformly stable computational strategy that provides trial elements that reproduce the given data at the quality of the* Trial Space Data Approximation. *This reduces the error and convergence analysis to an approximation problem for a data element $v^* \in V$ by a data subspace $D(U_M)$ in V under the norm in V.* $\square$

If there is no well-posedness, there still is a *backward error analysis*. Instead of solving the problem with data v^*, which may be unsolvable or ill-posed, one can come up with an element $\tilde{u}_M$ from the trial space which has data that are close to the given data, and roughly as close as possible for the given trial space. For PDE solving, this usually means that one has an exact solution of a PDE with perturbed boundary data and a perturbation in the inhomogeneity of the PDE. If these perturbations are calculated and turn out to be tolerable, the user might be satisfied with $\tilde{u}_M$. Many application papers proceed this way, unfortunately, but users should always keep in mind that there may be very different trial elements that reproduce the data nicely, if there is no well-posedness.

9 Discretization in Hilbert Spaces

We now assume that U is a Hilbert space with inner product $(.,.)_U$ and that the data map is composed of continuous functionals $\lambda \in \Lambda \subset U^*$ like in the beginning of Section 4. The Riesz map allows a transition from functionals to functions, and thus we can fix a finite subset $\Lambda_N = \{\lambda_1, \ldots, \lambda_N\} \subset \Lambda$ and take the Riesz representers $u_1, \ldots, u_N \in U$ of these functionals as trial functions. If linear independence is assumed, we have N-dimensional spaces $L_N \subset U^*$ and $U_N \subset U$ by taking the spans, and the space V_N is $\mathbb{R}^N$ as the range of the restriction R_N with $R_N D u = (\lambda_1(u), \ldots, \lambda_N(u))^T$ which just is the usual projection from $V := \mathbb{R}^\Lambda$ to $V_N := \mathbb{R}^{\Lambda_N} = \mathbb{R}^N$. If orthonormal bases are chosen, we have the 2-norm of coefficients as $\|u\|_U$ for all $u \in U$, but in order to comply with Section 4, we have to take the sup-norm in the range of the data map, which is the identity if discretized in that basis. But then the identity map is not well-posed, due the choice of norms which is not adequate for Hilbert spaces.

We thus have to change the setting, taking the norms in $V_N = \mathbb{R}^N$ as 2-norms, assuming Λ to be countable and total, taking orthonormal bases, and the restrictions as projections focusing on finite subsets of indices in the expansions. For the choice of U_N and V_N as above, we then have (13) and (14) with the constant 1.

This is the standard situation in Rayleigh–Ritz–Galerkin methods. It might be surprising that everything is perfectly well-conditioned here, but this is no miracle

because we focused on spaces, not on bases, and used an optimal basis for the theoretical analysis.

10 Optimal Recovery in Hilbert Spaces

When starting from a finite set Λ_N of functionals providing restricted data $\lambda_j(u^*)$ of a true solution of the analytic problem, the above choice of a trial space as the space spanned by the representers of the functionals is optimal under all other choices of trial spaces. This is a standard result in the theory of Reproducing Kernel Hilbert Spaces, but we include it here in a general form, because of its central importance within the context of studying all possible discretizations. We shall come back to it in Section 13.2.

Theorem 7 *Assume that we have a computational problem posed in a Hilbert space U, and the only available data are of the form $\lambda_1(u^*), \ldots, \lambda_N(u^*)$ for N linearly independent data functionals in U^* and an unknown function $u^* \in U$. Then, for any linear functional $\mu \in U^*$, consider all possible linear computational procedures for calculating good approximations of $\mu(u^*)$ using only the above values. Then there is a unique error-optimal strategy that works as follows:*

1. *Use the representers $u_1, \ldots, u_N \in U$ of the functionals $\lambda_1, \ldots, \lambda_N \in U^*$.*
2. *Calculate the interpolant $\tilde{u}$ to u^* in the span of the representers, i.e. solve the system*

$$\lambda_k(u^*) = \sum_{j=1}^{N} c_j \lambda_k(u_j) = \sum_{j=1}^{N} c_j (u_k, u_j)_U = \sum_{j=1}^{N} c_j (\lambda_k, \lambda_j)_{U^*}, \ 1 \leq k \leq N$$

and define

$$\tilde{u} = \sum_{j=1}^{N} c_j u_j.$$

3. *For each data functional $\mu \in U^*$, use the value*

$$\mu(\tilde{u}) = \sum_{j=1}^{N} c_j \mu(u_j)$$

as an approximation to $\mu(u^*)$.

This approximation has minimal error under all other linear computational procedures using the same data for calculating approximations of $\mu(u^*)$, in the sense that the error functional has minimal norm in U^*. □

This technique is independent of well-posedness and makes optimal use of the available data, error-wise. From the previous sections we conclude that it is uniformly stable when considered in terms of spaces, not bases. If applied to PDE solving, it is realized by *symmetric collocation* in the strong case and by the Rayleigh-Ritz procedure for the weak case [44]. It can also be applied to numerical integration and numerical differentiation, see e.g. [16, 46].

In the context of this paper, the above result shows that the quest for good trial spaces and well-posed discretizations has a simple solution in the Hilbert space situation. We shall come back to this in Section 13.5 when we look at the differences between weak and strong formulations.

11 Bases

We now assume that we have a well-posed analytic problem in the sense of Section 2 with an MRD discretization, and by a proper choice of restrictions (R_N, V_N) according to Theorem 3, we have uniform stability in the form of (14). We specialize here to the case of Theorem 2 where we have functionals $\lambda \in \Lambda$ and restrictions working via subsets $\Lambda_N \subset \Lambda$ selecting finitely many data. We now choose a basis $u_1, \ldots, u_M$ of U_M and take the functionals $\lambda_1, \ldots, \lambda_N$ from the set Λ_N. Then we consider the discretized system

$$\sum_{j=1}^{M} a_j \lambda_k(u_j) \approx f_{\lambda_k} = \lambda_k(u^*), \, 1 \leq k \leq N \qquad (18)$$

that we solve approximatively by residual minimization like in Section 7. Clearly, a bad choice of bases will spoil stability, but we want to study this effect in detail. We quantify the stability of the object basis by norm equivalence

$$c_M \|u_a\|_{WP} \leq \|a\|_M \leq C_M \|u_a\|_{WP} \text{ for all } a \in \mathbb{R}^M,$$

with an unspecified norm $\|.\|_M$ on $\mathbb{R}^M$ that is used in computation. With the $N \times M$ matrix $A = (\lambda_k(u_j))_{1 \leq j \leq M, 1 \leq k \leq N}$. and the basis representation

$$u_a := \sum_{j=1}^{M} a_j u_j$$

with coefficient vectors $a \in \mathbb{R}^M$, we see that $Aa = R_N Du_a$ holds. and get

$$\begin{aligned} \|a\|_M &\leq C_M \|u_a\|_{WP} \\ &\leq CC_M \|u_a\|_D \\ &\leq 2CC_M \|R_N Du_a\|_{V_N} \\ &= 2CC_M \|Aa\|_{V_N} \end{aligned}$$

by (6), and (13).

Theorem 8 *Under the above assumptions, the system (18) has the stability property*

$$\|a\|_M \leq 2CC_M \|Aa\|_{V_N} \text{ for all } a \in \mathbb{R}^M. \qquad \square$$

In Section 7, we minimized $\|R_N(Du^* - Du_M)\|_{V_N}$ over all $u_M \in U_M$. After introducing a basis in U_M, this is the same as minimization of $\|f - Aa\|_{V_N}$ with $f := R_N Du^* = (\lambda_1(u^*), \ldots, \lambda_N(u^*))^T \in \mathbb{R}^N$ over all $a \in \mathbb{R}^M$. We are satisfied with a vector $\tilde{a} \in \mathbb{R}^M$ such that

$$\|f - A\tilde{a}\|_{V_N} \leq \min_{a \in \mathbb{R}^M} \|f - Aa\|_{V_N} \leq C_A \|f - Aa^*\|_{V_N}, \qquad (19)$$

where a^* is a good coefficient vector for the direct approximation of the true solution u^* by elements of the trial space U_M. We use a^* in the way we used u_M^* in Section 7 as a competitor that may come from some special approximation technique. Then we form the elements $\tilde{u}_M := u_{\tilde{a}}$, $u_M^* = u_{a^*} \in U_M$ and see that (15) is satisfied.

This lets us arrive at Theorem 4, implying that the convergence rate is the same as the rate for the *Trial Space Data Approximation*, but this does not yield error bounds in terms of coefficients. However, we can proceed by

$$\begin{aligned}
\frac{1}{2CC_M}\|a^* - \tilde{a}\|_M &\leq \|A(a^* - \tilde{a})\|_{V_N} \\
&\leq \|Aa^* - f\|_{V_N} + \|f - A\tilde{a}\|_{V_N} \\
&\leq (1 + C_A)\|Aa^* - f\|_{V_N}
\end{aligned}$$

and get

$$\|a^* - \tilde{a}\|_M \leq (1 + C_A)(2CC_M)\|Aa^* - f\|_{V_N}.$$

The norm in V_N must be chosen to comply with Section 4, and this works for the discrete sup norm. But if users do not want to minimize $f - Aa$ in the sup norm, an additional norm equivalence comes into play, now on V_N, and this will often depend on $\dim V_N$. In detail, norm equivalence in V_N is assumed as

$$c_N\|R_N v\|_{V_N} \leq \|R_N v\|_N \leq C_N\|R_N v\|_{V_N} \text{ for all } v \in V.$$

and minimization in the new norm $\|.\|_N$ will replace (19) by

$$\|f - A\tilde{a}\|_N \leq \min_{a \in \mathbb{R}^M}\|f - Aa\|_N \leq C_A\|f - Aa^*\|_N,$$

and our above argumentation now yields

$$\|a^* - \tilde{a}\|_M \leq \frac{(1 + C_A)(2CC_M)}{c_N}\|Aa^* - f\|_N. \tag{20}$$

If bases are chosen badly, the quotient C_M/c_N can be extremely large and will spoil the uniformity that we had so far.

Users can check their stiffness matrices A computationally for stability, but Theorem 8 indicates that there may be a strong influence due to a bad choice of the trial basis. Even a calculation of a Singular Value Decomposition will not be completely basis-independent, since it only eliminates *orthogonal* basis transformations in the domain and range of A.

12 Nodal Bases

In *meshfree methods* [10,21,7,31,35,51,48,34,38,30], it is customary to write everything "entirely in terms of nodes" [10], which means that the functions u_M in the trial space U_M are parametrized by their values at certain *nodes* $x_1, \ldots, x_M$, i.e.

$$u_M(x) = \sum_{j=1}^{M} s_j(x)u_M(x_j) \text{ for all } u_M \in U_M$$

with *shape functions* s_j that are usually localized around x_j and have the Lagrange property $s_j(x_k) = \delta_{jk}$, $1 \leq j, k \leq M$. We prefer the term *nodal basis*, because there is nothing meshfree in the above representation, and the standard finite elements, which nobody would call meshfree, are nodal as well in the above sense. Many application papers report experimentally that these bases have favorable stability properties, and we shall now show why.

Theorem 9 *Assume a well-posed problem in the sense of (6), where U is a space of functions on some domain Ω. Furthermore, assume that the point evaluation functionals δ_x are uniformly bounded by $\gamma > 0$ in the norm $\|.\|_{WP}$. Finally, assume that the data space V and the restrictions V_N are normed via supremum norms, as mentioned in Section 4 and Theorem 2 as a special case. Then for each trial space $U_M \subset U$ with a nodal basis $s_1, \ldots, s_M$ using nodes $x_1, \ldots, x_M \in \Omega$ one can find a finite set of functionals $\lambda_1, \ldots, \lambda_N$ such that the $N \times M$ stiffness matrix A with entries $\lambda_j(s_k)$ has the uniform stability property*

$$\|a\|_\infty \leq 2\gamma C \|Aa\|_\infty \text{ for all } a \in \mathbb{R}^M. \tag{21}$$

Proof: We apply Theorem 3. Then

$$|u_M(x_j)| \leq \gamma \|u_M\|_{WP} \leq 2\gamma C \|R_N D(u_M)\|_{V_N} = 2\gamma C \max_{\lambda_k \in \Lambda_N} \left| \sum_{j=1}^{M} \lambda_k(s_j) u_M(x_j) \right|$$

and if we denote the vector of nodal values by $u_X \in \mathbb{R}^M$, we see that

$$\|u_X\|_\infty \leq 2\gamma C \|Au_X\|_\infty$$

with the *stiffness matrix A* with entries $\lambda_k(s_j)$.

This means that all trial spaces with nodal bases can be uniformly stabilized by taking good and large selections of test functionals, but note that the required continuity of the point evaluation functionals restrict the applicability to strong formulations with $\|.\|_{WP} = \|.\|_\infty$.

We can now go over to the error of nodal values. To this end, let a^* be the vector of nodal values of the true solution u^*, and a_M^* be the vector of nodal values of the comparison trial function u_M^*. We then have

$$\begin{aligned}
\|a^* - a_M^*\|_\infty &\leq \gamma \|u^* - u_M^*\|_\infty \\
&= \gamma \|u^* - u_M^*\|_{WP} \\
&\leq 2C\gamma(2C_A + 3)\varepsilon(u^*, U_M, D)
\end{aligned}$$

by Corollary 1.

Corollary 3 *Under the assumptions of Theorem 9, the error of a computational procedure as in Section 7, evaluated on the nodes, is pointwise bounded by the error of the* Trial Space Data Approximation, *i.e. the approximation of the data of the true solution by the data of trial elements, measured in the data norm.* $\qquad\square$

This means that using a nodal basis transfers the results of Theorem 5 directly to a convergence on the nodes. This is a very useful result for many meshfree methods using nodal bases, e.g. when applying Moving Least Squares techniques. Section 7 will provide convergence proofs and convergence rates for such techniques.

Readers might wonder how uniform stability is possible for general trial spaces with nodal bases, though it is well-known that standard stiffness matrices for a mesh-width h and a standard second-order Poisson problem have a condition that blows up like $\mathcal{O}(h^{-2})$ for $h \to 0$. But the point is that this paper does not look at condition. And, since it aims at meshless methods that may have singular "stiffness" matrices if no overtesting is done, the matrices considered here will be non-square and the linear systems will be overdetermined, making the standard notion of condition inadequate. Looking back at (21) shows that we do not consider the "forward" matrix A that maps nodal values to PDE data, but rather the inverse process that maps PDE data to nodal values. And then well–posedness as in (7) or (8) yields the chance to get uniform stability by sufficient overtesting.

To be more specific, consider the trivial discretization that recovers the Laplacian at the origin from the four values in the corners of $[-h, +h]^2$. The forward map, if implemented as the standard five-point star, behaves like $\mathcal{O}(h^{-2})$, but the reverse process just takes a mean with fixed weights 1/4 and does not degenerate for small h. Considering condition would yield $\mathcal{O}(h^{-2})$, but the inverse process is uniformly stable if norms are properly chosen and if the basis is nodal.

13 Analytic Examples

For simplicity, we focus here on interpolation and PDE solving, omitting other linear operator equations that can be handled.

13.1 Interpolation

We start with the rather simple case of recovering a function u on some compact domain $\Omega \subset \mathbb{R}^d$ from data of u that do not involve derivatives. A *strong* formulation takes $\Lambda = \{\delta_x \ : \ x \in \Omega\}$ on a space U on which these functionals are continuous, e.g. $U = C(\Omega)$ under the sup norm. A *weak* formulation uses different data, e.g. functionals

$$\lambda_v(u) := (u, v)_{L_2(\Omega)} \text{ for all } u, v \in U := L_2(\Omega)$$

and

$$\Lambda := \{\lambda_v \ : \ \|v\|_{L_2(\Omega)} = \|\lambda_v\|_{L_2^*(\Omega)} = 1\}.$$

The strong case takes $V := C(\Omega) = U$ under the sup norm, while the weak case uses $V := L_2(\Omega) = U$ under the L_2 norm. In both cases, the data map is the identity, and we have well-posedness in the norms $\|.\|_{WP} = \|.\|_U$ in both cases, but the norms differ.

The restrictions can work by selection of finitely many functionals in both cases, and all axioms of Section 4 are satisfied.

We now fix an arbitrary finite-dimensional trial space $U_M \subset U$, and Theorem 3 yields that there is a restriction that makes the linear system (11) uniformly stable in the sense (14) with $C = 1$.

The computational procedures of Section 7 can use a comparison trial function u_M^* that is the best approximation to the true solution u^* in the norm $\|.\|_U = \|.\|_{WP} = \|.\|_D$, and Theorem 4 then shows that the computational solution $\tilde{u}_M$ has the same convergence rate as the best approximation u_M^*.

The computational solution $\tilde{u}_M$ is obtained via (15) from a stably discretized linear system, and we assume that we perform inexact minimization of the norm $\|R_N(u^* - u_M)\|_{V_N}$.

In the strong case, this is best linear discrete Chebyshev approximation on sufficiently many points, i.e. a linear optimization problem. In the univariate case with U_M being a trial space of polynomials of degree $M - 1$ on an interval I, a discretization on $N \geq M$ test points forming a set P_N will always have a stability inequality

$$\|u\|_{\infty, I} \leq C(M, N) \|u\|_{\infty, P_N}$$

of the form (12), but the stability constant varies. For $M = N$ equidistant points, the constant $C(M, M)$ grows exponentially with M, and for Chebyshev–distributed test points it still grows like $\log M$. Uniform stability holds for $N = \mathscr{O}(M^2)$ equidistant points, as follows from a standard argument going back to the notion of norming sets [26] and using Markov's inequality [56, Ch. 3.3]. Theorem 3 only proves *existence* of a uniformly stable discretization, but this example shows that there may be a considerable amount of oversampling or overtesting behind the scene.

If stability is uniform, nodal bases written in terms of values at M nodes x_j will trivially lead to $|u(x_j)| \leq \|u\|_{\infty, I} \leq C \|u\|_{\infty, P_N}$ for all $u \in U_M$, which is (21).

The weak case discretizes by N well-chosen normalized test functionals λ_{v_j} with normalized Riesz representers $v_j \in L_2(\Omega)$, and the quantity $\|R_N(u^* - u_M)\|_{V_N}$ to be minimized is

$$\max_{1 \leq j \leq N} |(v_j, u^* - u_M)_{L_2(\Omega)}|.$$

Our theory shows that the test functionals can be chosen to render uniform stability, but there is a trivial standard choice via the M functionals represented by an orthonormal basis $v_1, \ldots, v_M$ of U_M. Then the above minimization produces the best approximation u_M^* to u^* from U_M without any oversampling. See [15] for the stability of L_2 recovery for random data points.

In both cases, Theorem 5 is applicable, and we see that the L_∞ or L_2 convergence rates of the non-discrete best approximations carry over to the discrete approximations.

To compare the difference of convergence rates between weak and strong formulations for a given fixed trial space U_M, we see immediately that the L_2 convergence rate is never worse than the L_∞ rate, but it is taken in a weaker norm. If users insist on the best possible convergence rate in L_2, they should take a weak form, at the expense of a sufficiently good numerical integration. But the L_∞ error of their solution will clearly not have a better L_∞ convergence rate than the strong solution.

Both computational approximations, weak or strong, converge like the best approximations in the respective data norm, and this is a fair deal. Convergences can

be spectral in certain cases, e.g. in case of univariate functions on an interval I that have a complex extension that is analytic in a region of the complex plane containing I in its interior. This shows how the theory applies to spectral convergence situations without change.

But, of course, there is the extreme case where the solution is only in L_2 and not in $C(\Omega)$. Then the strong technique is undefined. But then the weak technique is forced to have the weak data given directly, without numerical integration, because the latter is as unfeasible as the strong technique.

All other examples will show a very similar behavior, differing only in their data maps.

13.2 Standard Homogeneous Weak Poisson Problem

We fix a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ and consider the weak Dirichlet problem $-\Delta u = f$ with homogeneous boundary conditions. This works on the Hilbert space $U := H_0^1(\Omega)$ with the inner product

$$(u,v)_1 := \int_\Omega \nabla^T u(x) \nabla v(x) dx \text{ for all } u, v \in U,$$

and the standard (global) weak formulation asks for a function $u \in U$ with

$$(u,v)_1 = (f,v)_{L_2(\Omega)} \text{ for all } v \in U = H_0^1(\Omega).$$

In the sense of this paper, the functionals are

$$\lambda_v \ : \ u \mapsto \lambda_v(u) = (u,v)_1 \text{ for all } u, v \in U$$

and the problem takes the form (1) with

$$\Lambda := \{\lambda_v \ : \ v \in U, \ \|v\|_U = 1\} \subset U^*$$
$$f_{\lambda_v} = (u^*,v)_1 = (f,v)_{L_2(\Omega)} \text{ for all } \lambda_v \in \Lambda$$

where $u^* \in U$ is the true solution.

To check the well-posedness in the sense of Section 2, we get

$$\|Du\|_V = \sup_{\lambda_v \in \Lambda} |\lambda_v(u)| = \sup_{v \in U, \|v\|_U = 1} |(u,v)_U| = \|u\|_U =: \|u\|_{WP}$$

proving well-posedness, and the data map D is an isometry.

We now consider arbitrary trial spaces $U_M \subset U = H_0^1(\Omega)$ to allow standard or extended or generalized finite elements, or even certain spectral methods of Galerkin type. Theorem 5 is applicable, and we see that we get the convergence rate of approximations to the true solution in $U = H_0^1(\Omega)$. This is well-known from finite elements, but it holds in general, provided that MRD testing is done. It applies to Petrov–Galerkin methods and spectral techniques of Galerkin type. The rate mainly depends on the smoothness of the solution and on the trial space chosen.

For the standard finite-element situation with piecewise linear elements, this yields $\mathcal{O}(h)$ convergence in $H_0^1(\Omega)$, as usual for that regularity. To reach $\mathcal{O}(h^2)$ convergence

in $L_2(\Omega)$ under $H^2(\Omega)$ regularity, the Aubin–Nitsche trick is used, and it is an add-on that is not covered by our theory. But it follows from the fact that the best approximation to u^* in $H_0^1(\Omega)$ automatically has $\mathcal{O}(h^2)$ convergence in L_2 under $H^2(\Omega)$ regularity. This is independent of PDE solving, it is a property of Approximation Theory.

But it should be mentioned that *general* solutions on domains with incoming corners may only have $H^1(\Omega)$ regularity, while *special* solutions may have any regularity. Domain properties determine the *minimal* regularity of possible solutions, but not the *maximal* regularity. The theory of this paper, like Approximation Theory, uses the specific regularity of the approximated function or PDE solution to derive convergence rates, without considering properties of the domain.

13.3 Collocation Methods

We now assert that all linear PDE or ODE problems can be numerically solved by collocation in sufficiently many points in such a way that the convergence rate in the well-posedness norm is at least the rate of the *trial space data approximation*. This, of course, includes pseudospectral methods. But we have to add more details to show how it follows from Theorem 5. We only have to show that collocation is an MRD discretization and pick a suitable form of well-posedness.

The space U should be a normed linear space of functions on a domain Ω with boundary Γ, for instance a Sobolev space. To keep things simple, we assume that the analytic problem is posed in strong form by evaluating a linear elliptic second-order differential operator L on points of the domain and a linear boundary operator B on the boundary, i.e.

$$\begin{aligned} Lu(x) &= f(x), \quad \text{for all } x \in \Omega \\ Bu(y) &= g(y), \quad \text{for all } y \in \Gamma \end{aligned} \tag{22}$$

where f and g are given functions on Ω and Γ. Introducing continuous functionals $\lambda_x(u) := Lu(x) = \delta_x \circ L$ and $\mu_y(u) := Bu(y) = \delta_y \circ B$ on U one gets a problem of the form (1) with

$$\Lambda := \{\lambda_x \,:\, x \in \Omega\} \cup \{\mu_y \,:\, y \in \Gamma\} \tag{23}$$

and it should be clear that one can allow more than two operators, and combinations of different boundary conditions.

From here there are different ways to proceed towards well-posedness, but we shall ignore well-posedness for a moment. We normalize all functionals as elements of U^* and pose the problem in the form (1) with $f_\lambda := \lambda(u^*)$ for all $\lambda \in \Lambda$. Then $|\lambda(u)| \leq \|u\|_U$ for all $u \in U$, $\lambda \in \Lambda$, and there is no problem to define the space V and the restrictions via taking suprema. We can apply Theorems 3 and 4 without assuming well-posedness, and we see that we can work on any trial space U_M, but all results only hold in the data norm. Comparing with any existing good approximation u_M^* to u^* from U_M, we get some $\tilde{u}_M \in U_M$ by a discrete computational method such that

$$\sup_{\lambda \in \Lambda} |\lambda(\tilde{u}_M - u^*)| \leq (2C_A + 3) \sup_{\lambda \in \Lambda} |\lambda(u_M^* - u^*)| \leq (2C_A + 3)\|u_M^* - u^*\|_U$$

due to normalization of the functionals, and if we use some u_M^* with small $\|u_M^* - u^*\|_U$, we get the above statement. The backward error analysis of Section 8 will be applicable here.

In the above setting, the most natural well-posedness condition would be of the form

$$\|u\|_{WP} \leq C \max(\|Lu\|_{\infty,\Omega}, \|Bu\|_{\infty,\Gamma}) \qquad (24)$$

for a suitable norm $\|.\|_{WP}$ on U. This holds for $U := C^2(\Omega) \cap C(\overline{\Omega})$ with the sup norm in U [12, (2.3), p. 14] for uniformly elliptic operators L and Dirichlet boundary data. This implies by Theorem 5 that for such problems the convergence rate of the *Trial Space Data Approximation* carries over to the convergence rate of collocation in the sup norm.

As a special example we consider unsymmetric collocation [28] by translates of the kernel K of a Hilbert space U, applied to a Dirichlet problem of the form (22) on a domain $\Omega \subset \mathbb{R}^d$. The trial space U_M is spanned by kernel translates of the form $v_j := K(\cdot, x_j)$, $1 \leq j \leq M$ for nodes $x_1, \ldots, x_M \in \Omega$, but this is not a stable basis. A *nodal basis* in the sense of Section 12 consists of the Lagrange basis $u_1, \ldots, u_M$ spanning the same trial space. Collocation is done via the functionals defined for (23), and to make them continuous we can take a space like $U := H^m(\Omega)$ with some $m > 2 + d/2$. We have well-posedness in the sense of (24) in the sup-norm.

Theorem 10 *Unsymmetric collocation in the sense of E. Kansa [28] has the property that for each possible trial space spanned by kernel translates there is a selection of test functionals such that the stiffness matrix, when written in terms of the nodal basis, has a uniform stability property (21). If solved by residual minimization along the lines of Section 7, error bounds follow from Corollary 1. Convergence rates in the sup norm are obtained from the rate of convergence of second derivatives in the sup norm of interpolants of the true solution by the trial space.* □

This provides many explicit convergence rates via standard results on interpolation by translates of kernels [56, Chapter 11]. For instance, the L_∞ convergence for unsymmetric collocation using the Whittle–Matérn kernel reproducing $H^m(\mathbb{R}^d)$ for $m > 2 + d/2$ is like $\mathcal{O}(h^{m-2-d/2})$ in terms of the *fill distance* $h := \sup_{y \in \Omega} \min_{1 \leq j \leq M} \|y - x\|_2$, while the convergence is exponential for kernels like the Gaussian or multiquadrics [14].

The functionals in (23) are a mixture of two kinds, but Theorem 3 and Lemma 1 do not say how to achieve a uniformly stable balance between testing B on the boundary and testing L in the interior. Future work should address this problem, and Section 13.1 suggests that there might be quite some overtesting needed for uniform stability. Square collocation systems can even be singular [23], such that overtesting is necessary in general.

All of this readily generalizes to plenty of other linear well-posed PDE problems, and readers can use the tools of this paper to assemble what they need. See [32] for a very extensive treatment of collocation and plenty of references.

Note that unsymmetric collocation is a pseudospectral method in the sense of the literature (see e.g. [20, 19, 13]) on spectral methods, and this paper provides a general way to assess convergence of pseudospectral methods. Since we write the analytic

and the computational problems in terms of arbitrary functionals, this approach also covers spectral methods in Tau form.

The interplay between collocation and Galerkin methods, and the way to choose the collocation points to get stability has a long history for ODEs, see e.g. [17,54, 53], but the PDE case still needs a lot of future work.

13.4 Weak Dirichlet Problems

The standard finite element procedures for solving Dirichlet problems for the Laplace operator on bounded domains $\Omega \subset \mathbb{R}^d$ use strong data on the boundary and weak data in the interior. The data functionals are

$$
\begin{aligned}
\Lambda_1 &:= \{\lambda_v \, : \, u \mapsto (\nabla u, \nabla v)_{L_2(\Omega)} \text{ for all } v \in H_0^1(\Omega), \, \|\nabla v\|_{L_2(\Omega)} = 1\}, \\
\Lambda_2 &:= \{\delta_y \, : \, y \in \Gamma := \partial\Omega\}, \\
\Lambda &:= \Lambda_1 \cup \Lambda_2.
\end{aligned}
$$

This leads to the data norm

$$
\|u\|_D = \max(\|u\|_{\infty,\Gamma}, \|\nabla u\|_{L_2(\Omega)})
$$

if we take the sup over all functionals as in Section 4. It is well-defined on the space $U := H^1(\Omega) \cap C(\overline{\Omega})$. Using the Poincaré inequality and the Maximum Principle [27] after splitting u into a harmonic part with boundary conditions and a function in $H_0^1(\Omega)$ satisfying the differential equation, we get a well-posedness inequality

$$
\|u\|_{L_2(\Omega)} \leq C\|u\|_D = C\max(\|u\|_{\infty,\Gamma}, \|\nabla u\|_{L_2(\Omega)}) \text{ for all } u \in U.
$$

Note that the Sobolev inequality forbids to use the sup norm on the left-hand side for space dimension $d > 1$.

Whatever the chosen trial spaces in U are, Theorem 5 shows that the convergence rate in L_2 of uniformly stabilized computational methods will be the convergence rate of the *Trial Space Data Approximation*, i.e. with respect to $\|u\|_{\infty,\Gamma}$ and $\|\nabla u\|_{L_2(\Omega)}$. If the trial space is spanned by translates of the Whittle–Matérn kernel reproducing $H^m(\mathbb{R}^d)$ for $m > 1 + d/2$, the L_2 convergence rate is $\mathscr{O}(h^{m-1})$ in terms of the fill distance h of the trial nodes [56].

For standard finite elements, the above approach yields $\mathscr{O}(h)$ convergence in L_2. This is without the Aubin–Nitsche trick, and it does not use H^2 regularity.

The Aubin–Nitsche trick has nothing to do with finite elements and weak problems. It is a feature of Approximation Theory, doubling a convergence rate for certain nested approximations in Hilbert spaces under additional regularity assumptions. This is well-known from splines [2, 5.10] and kernel-based methods [41]. In the context of this paper, one considers the best approximation to the true solution in $H_0^1(\Omega)$, and it will automatically yield $\mathscr{O}(h^2)$ convergence under H^2 regularity, but only for zero boundary conditions.

13.5 Weak-Strong Comparison

If we compare with what we had in the strong case, the situation for fixed trial spaces is roughly as follows:

1. The weak case has L_2 convergence at the convergence rate for first derivatives,
2. the strong case has L_∞ convergence at the convergence rate for second derivatives.

This usually yields a better rate for the weak case, as we saw when comparing $\mathcal{O}(h^{m-1})$ in L_2 with $\mathcal{O}(h^{m-2-d/2})$ in L_∞ for the trial space spanned by translates of the Whittle–Matérn kernel. On the downside, weak methods usually have to perform numerical integration at an accuracy that complies with the convergence rate, and they converge in a weaker norm. However, numerical examples in Section 14 indicate that the convergence rate of nodal values for unsymmetric collocation solving strong problems might be higher than predicted by Theorem 10 and be comparable to the L_∞ convergence rate predicted for weak methods.

With the exception of Sections 9 and 10, this paper fixed a trial space U_M and then cared for stable testing. If one fixes the available finite data first and then asks for an optimal trial space, Section 10 makes the above discussion about differences between strong and weak methods obsolete, because the input data and the trial spaces are different. The optimal solution is always the one described in Section 10, and it is furnished by taking Riesz representers as trial functions. This is symmetric collocation [44] in the strong case and the Rayleigh-Ritz method in the weak case. Since this paper allows arbitrary evaluation functionals μ in Theorem 7, it is pointwise and L_∞-optimal by taking functionals $\mu = \delta_x$, and L_2-optimal by taking all functionals $\mu = \lambda_v = (.,v)_{L_2(\Omega)}$. See [44] for details, including comparisons of convergence rates between finite elements and various collocation methods, showing the superiority of symmetric collocation, error-wise, under all methods that use strong data.

13.6 MLPG

We stay with the Dirichlet problem for the Laplacian, for simplicity, and describe the standard variation of the *Meshless Local Petrov–Galerkin* [6,5] method. The difference to the standard weak formulation is that the integrals are localized and the boundary integrals are kept. This means that on small subdomains $\Omega_h \subset \Omega$ with boundaries $\Gamma_h \subset \overline{\Omega}$ the strong equation $-\Delta u = f$ is integrated against a test function v_h to define functionals of the form

$$u \mapsto \lambda_{\Omega_h, v_h}(u) = \frac{-1}{vol(\Omega_h)} \int_{\Omega_h} v_h \cdot \Delta u$$

that are continuous on $U := C^2(\Omega) \cap C(\overline{\Omega})$, and the problem (22) takes the form (1) via

$$\begin{aligned}
\lambda_{\Omega_h, v_h}(u) &= \frac{1}{vol(\Omega_h)} \int_{\Omega_h} v_h \cdot f \quad \text{for all } \Omega_h \subset \Omega, \; v_h \in C(\Omega_h) \\
\delta_y(u) &= g(y) \qquad\qquad\qquad \text{for all } y \in \Gamma
\end{aligned} \tag{25}$$

for given continuous functions f on Ω and g on Γ. One can restrict the domains Ω_h and the test functions v_h further, and allow other ways of handling the boundary conditions. Furthermore, the above functionals are usually transformed by integration by parts before they are implemented, but we deal with this later.

This variation of Example 1 defines a *new* analytic problem in the sense of this paper, because the data map is different from what we had in the strong and the global weak case. Therefore we need to prove some form of well-posedness for it, and this seems to be missing completely in the rich literature on the MLPG method. On the space $U = C^2(\Omega) \cap C(\overline{\Omega})$ we know that (24) holds for $\|.\|_{WP} = \|.\|_{\infty,\overline{\Omega}}$, and we assert the well-posedness inequality

$$\|u\|_{\infty,\overline{\Omega}} \leq C \left(\sup_{\Omega_h,v_h} |\lambda_{\Omega_h,v_h}(u)| + \|u\|_{\infty,\Gamma} \right) \text{ for all } u \in U. \qquad (26)$$

But this follows from (24) by setting $f := \Delta u$ in

Lemma 2 *For each continuous function f on some compact domain $\Omega \subset \mathbb{R}^d$ the norms*

$$\|f\|_{\infty,\Omega} = \|f\|_I := \sup_{\Omega_h \subset \Omega} \frac{1}{vol(\Omega_h)} \left| \int_{\Omega_h} f \right|$$

coincide, where the diameter of the admissible sets Ω_h can be bounded above by some arbitrary $r > 0$, if required. One can also restrict the subdomains Ω_h to balls or cubes intersected with Ω.

Proof. Clearly $\|f\|_I \leq \|f\|_{\infty,\Omega}$ holds. To prove $\|f\|_I \geq \|f\|_{\infty,\Omega}$, assume $f \neq 0$ with $\|f\|_{\infty,\Omega} = f(\tilde{x}) > 0$ for some $\tilde{x} \in \Omega$. Then pick an arbitrary $\varepsilon < f(\tilde{x})/2$ and an arbitrary $r > 0$ and choose Ω_h to be a subdomain of the set of points $x \in \Omega$ with

$$0 < (1-\varepsilon)f(\tilde{x}) \leq f(x) \leq f(\tilde{x}), \|x - \tilde{x}\|_2 \leq r.$$

For instance, one can take the intersection of sufficiently small balls or cubes around $\tilde{x}$ with the domain Ω, or if $\tilde{x}$ is on the boundary, one may move slightly into the interior and ensure Ω_h to be in the interior of the domain. Then

$$(1-\varepsilon)\|f\|_{\infty,\Omega} = (1-\varepsilon)f(\tilde{x}) \leq \frac{1}{vol(\Omega_h)} \int_{\Omega_h} f \leq f(\tilde{x}) = \|f\|_{\infty,\Omega}.$$

Note that this proves well-posedness only on $U = C^2(\Omega) \cap C(\overline{\Omega})$, not on a larger space, but for all possible test functions and domain shapes and sizes. The boundary conditions can be rephrased by weak functionals taking means, using Lemma 2 again, now setting $f := g$ and working on the boundary.

Any a-priori renormalization of *all* available functionals will possibly spoil this argument. But as soon as finitely many functionals are selected for computation, one can renormalize for the computational procedure.

If integration by parts is applied to the functionals, they change their computational form without changing their value, and this is used in the known variations of

the MLPG technique. For instance,

$$
\begin{aligned}
\lambda_{\Omega_h,v_h}(u) &= \frac{1}{vol(\Omega_h)} \int_{\Omega_h} \nabla^T v_h \cdot \nabla u - \frac{1}{vol(\Omega_h)} \int_{\Gamma_h} v_h \frac{\partial u}{\partial n} \\
&= \frac{-1}{vol(\Omega_h)} \int_{\Omega_h} \Delta v_h \cdot u + \frac{1}{vol(\Omega_h)} \int_{\Gamma_h} u \frac{\partial v_h}{\partial n} - \frac{1}{vol(\Omega_h)} \int_{\Gamma_h} v_h \frac{\partial u}{\partial n}
\end{aligned}
$$

are two ways to rewrite the functionals, and one can insert different admissible test functions. The basic well-posedness on $U = C^2(\Omega) \cap C(\overline{\Omega})$ will stay as is, because the sup of all these functionals will be bounded above by $\|\Delta u\|_{\infty,\Omega}$, as long as there are no other upper bounds proven.

The special method called MLPG5 uses constant test functions like in Lemma 2. Then the functionals take the extremely simple form

$$
\lambda_{\Omega_h,v_h}(u) = -\frac{1}{vol(\Omega_h)} \int_{\Gamma_h} \frac{\partial u}{\partial n},
$$

i.e. they are only integrals of the normal derivative on subdomain boundaries. Nevertheless, Lemma 2 holds, and there is well-posedness in the sup norm on $U = C^2(\Omega) \cap C(\overline{\Omega})$. It is an open problem to prove other well-posedness inequalities after fixing a special form of the functionals. The above technique via Lemma 2 always goes back to (24), whatever the form of the functionals is after integration by parts. Therefore the convergence theory for given trial spaces will be the same as for the strong collocation methods in Section 13.3.

Theorem 11 *If the Meshless Local Petrov–Galerkin method is carried out*

1. *for a well-posed second-order elliptic problem,*
2. *using sufficiently many well-chosen test functionals (25) along the lines of Theorem 3,*
3. *and applying a residual minimization algorithm as in Section 7 for solving the overdetermined system approximatively,*

the algorithm is convergent with uniform stability, and the convergence rate in the sup norm is the rate of the Trial Space Data Approximation. *This rate is at least as good as for strong collocation using the same trial spaces.* □

Depending on the PDE problem, the smoothness of the true solution, and the trial space chosen, this yields various convergence results, up to spectral convergence. In most applications, the trial functions are shape functions provided by Moving Least Squares, and raising the degree of the local polynomials will increase the convergence rate appropriately [29,55,3,4] if the solution is smooth enough. Readers are encouraged to apply the framework of this paper to derive special convergence results for various trial spaces and different variations of the MLPG technique. In particular, an extension to elasticity problems should be quite useful.

14 Numerical Examples

To demonstrate how additional testing stabilizes unsymmetric collocation of a Poisson problem, consider the Dirichlet problem for $-\Delta$ on the unit square. The trial space is chosen to consist of $M = 25$ regularly placed translates of the radial Whittle-Matérn-Sobolev kernel $r^6 K_6(r)$ generating a Hilbert space that is norm-equivalent to Sobolev space $H^7(\mathbb{R}^2)$. Testing was done on $N = 25, 36, 49, \ldots$ regular points, using function values on the boundary points and Laplacian values on the interior points. The third column of Table 1 shows how the stability improves with overtesting. Note that the $M = N$ case cannot be shown to have a nonsingular matrix [23], but applications have never reported a single singular example, so far.

DOF Trial	DOF Test	$\|A^\dagger\|_\infty$	$\|B^\dagger\|_\infty$
25	25	30.4933	5128.49
25	36	27.8728	7677.07
25	49	5.1670	758.83
25	64	4.6046	573.91
25	81	4.1675	527.35
25	100	4.3014	465.59
25	121	4.2553	476.51

Table 1 Stabilization results

The stability constant for an $N \times M$ matrix A with rank $M \leq N$ can be calculated as the row-sum norm of the pseudoinverse $A^\dagger$, see Table 1. Note that the stability constant was evaluated for the nodal basis. Using the standard non-nodal unsymmetric collocation matrix B instead, the instability is much worse, see the final column in Table 1.

To check the convergence rates for unsymmetric collocation as given in Theorem 10, we take the toy Dirichlet problem for the Laplacian with exact solution $u(x, y) = x^2 + y^2$ on the unit square. The collocation setup is as in the context of Theorem 10, a nodal basis is used, and the error is evaluated at the nodes. The smoothness of the solution is no restriction of the convergence rate, and thus the trial space and the required derivatives determine the convergence rate.

As a first case, we take the radial Whittle-Matérn kernel $K_{m-d/2}(r) r^{m-d/2}$ generating Sobolev space $H^m(\mathbb{R}^d)$ and the trial space spanned by its translates. For $m = 6$ and $d = 2$, the convergence rate in terms of powers of the meshwidth h and proven by Theorem 10 is $m - 2 - d/2 = 6 - 2 - 1 = 3$ in L_∞, but the observed rate in Fig. 1 is about 4.3, leading to the hypothesis that the true order is 4. The weak case as described in Section 13.4 will have a convergence in L_2 with rate $m - 1 = 5$, and this will normally lead to an L_∞ rate of $5 - d/2 = 4$ that is comparable to the strong case.

Choosing the C^6 Wendland kernel $\phi_{3,3}$ from [56] shows in Fig. 2 an observed rate of ≈ 3, while the rate proven by Theorem 10 is $4.5 - 2 - d/2 = 1.5$ in L_∞ because the kernel generates $H^{d/2+k+1/2}(\mathbb{R}^d) = H^5(\mathbb{R}^3)$ with traces in $H^{4.5}(\mathbb{R}^2)$. Again, it can be hypothesized that the true rate is $4.5 - 1 - d/2 = 2.5$, while the expected L_2 rate for

the weak formulation is $4.5 - 1 = 3.5$ which is comparable to a rate of $3.5 - d/2 = 2.5$ in L_∞ as in the strong case.
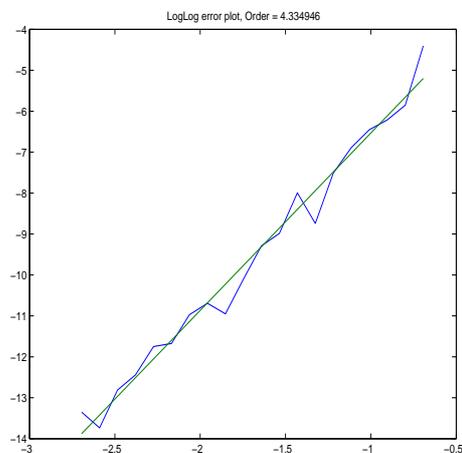

Fig. 1 Errors for Sobolev space H^6, order ≈ 4.3
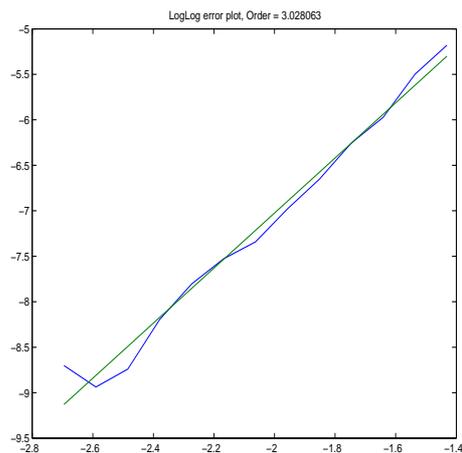

Fig. 2 Errors for C^6 Wendland kernel, order ≈ 3

These results suggest that Theorem 10 does not provide optimal convergence rates, and the rates for L_∞ convergence might be the same for weak and strong problem formulations.

The above convergence rates were evaluated for errors of solution values on the trial nodes. Since the nodes were regularly placed. the rates carry over to the error in

the whole domain, if the nodal values are interpolated using translates of a suitable kernel of finite smoothness. This follows from the fact that nodal Lagrange basis functions are uniformly bounded for regular node distributions [18].

As a numerical example for a meshless method in weak form, we take a case from [37] concerning convergence of the MLPG5 method as described and analyzed in Section 13.6. The three tables of [37] show convergence in L_∞ of the MLPG5 method based on Moving Least Square trial spaces using polynomials of degree $m = 2, 3, 4$, and the estimated convergence rates are roughly 1.7, 2.1, and 3.6, respectively, taken without overtesting, but with no visible instability effect [36]. In [37], the functionals were defined via integrals over normal derives along subdomain boundaries, and thus the rates of trial space approximation are 2, 3, and 4, respectively, for these first-order data. But the theory in Section 13.6 works with a MRD discretization that leads to a data norm involving second derivatives, and thus the rates predicted by this paper are 1, 2, and 3, respectively. Like in the strong case with kernel-based trial spaces, this is not too far from the observations, but needs further investigation.

15 Outlook

Summarizing, the theoretical part of this paper shows that under mild hypotheses

1. all well-posed problems have uniformly stable discretizations made possible by sufficiently extensive testing, and
2. convergence rates for such discretizations can be played back to known convergence rates of *Trial Space Data Approximation*, i.e. the approximation of the data of the true solution by the data of the trial elements. These rates depend on what "data" means and are taken in the norm arising in the well-posedness condition.
3. Weak and strong formulations of a given background problem will have different definitions of "data" and will need different versions of well-posedness, and these differences enter into the previous item and influence the convergence rates, even when trial spaces are the same for both formulations.
4. For a given fixed trial space, it is shown that in standard applications the weak formulation converges slightly faster than the strong formulation.
5. Nodal bases have a stability advantage over other bases.

   But there are also some open problems:

- The paper only proves existence of a uniformly stable discretization via sufficient overtesting, but it is open how much overtesting is actually needed. This is a difficult task to be handled for each specific case. If the dimension of the trial space is M and if $N \geq M$ functionals are used for testing, Section 13.1 shows that one may need $N = \mathcal{O}(M^2)$ for securing uniform stability, if test functionals are chosen naively. In other cases [43], one can get away with $N = \mathcal{O}(M)$, and in case of optimal recovery in the sense of Section 10 by e.g. the Rayleigh-Ritz method, one can work with $N = M$.
- The examples in Section 14 suggest that observed convergence rates for unsymmetric collocation are larger than the proven ones. This definitely needs further investigation.

– The methods of this paper always assume the functionals to be given exactly, not approximately. Only their values can be noisy, as in Section 8. This excludes various interesting applications, namely Finite Differences, the Direct Meshless Local Petrov Galerkin (DMLPG) technique [37], and localized kernel-based methods that provide sparse stiffmess matrices [50, 40, 52, 49, 57].

– Since the convergence rates are determined by how well the PDE data of the solution can be approximated from the trial space, one can adaptively enlarge trial spaces to design adaptive PDE solvers. In Approximation Theory and Image Reconstruction there are plenty of adaptive methods, and they will be applicable to PDE solving, if they are able to approximate derivatives. This opens many possibilites for new research towards adaptive methods.

References

1. Y. Abdelaziz and A. Hamouine. A survey of the extended finite element. *Computers and Structures*, 86:1141–1151, 2008.
2. J.H. Ahlberg, E.N. Nilson, and J.L. Walsh. *The theory of splines and their applications*, volume 38 of *Mathematics in science and engineering*. Academic Press, 1967.
3. M.G. Armentano. Error estimates in Sobolev spaces for moving least square approximations. *SIAM J. Numer. Anal.*, 39(1):38–51, 2001.
4. M.G. Armentano and R.G. Durán. Error estimates for moving least square approximations. *Appl. Numer. Math.*, 37:397–416, 2001.
5. S. N. Atluri. *The meshless method (MLPG) for domain and BIE discretizations*. Tech Science Press, Encino, CA, 2005.
6. S. N. Atluri and T.-L. Zhu. A new meshless local Petrov-Galerkin (MLPG) approach in Computational Mechanics. *Computational Mechanics*, 22:117–127, 1998.
7. I. Babuška, U. Banerjee, and J.E. Osborn. Survey of meshless and generalized finite element methods: a unified approach. *Acta Numer.*, 12:1–125, 2003.
8. T. Belytschko, R. Gracie, and G. Ventura. A review of extended/generalized finite element methods for material modeling. *Modelling and Simulation in Materials Science and Engineering*, 17, 2009.
9. T. Belytschko, Y. Guo, W.K. Liu, and S.P. Xiao. A unified stability analysis of meshless particle methods. *International Journal for Numerical Methods in Engineering*, 48:1359–1400, 2000.
10. T. Belytschko, Y. Krongauz, D.J. Organ, M. Fleming, and P. Krysl. Meshless methods: an overview and recent developments. *Computer Methods in Applied Mechanics and Engineering, special issue*, 139:3–47, 1996.
11. K. Böhmer and R. Schaback. A nonlinear discretization theory. *Journal of Computational and Applied Mathematics*, 254:204–219, 2013.
12. D. Braess. *Finite Elements. Theory, Fast Solvers and Applications in Solid Mechanics*. Cambridge University Press, 2001. Second edition.
13. C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods*. Scientific Computation. Springer, Berlin, 2007.
14. A. H.-D. Cheng, M. A. Golberg, E. J. Kansa, and G. Zammito. Exponential convergence and h-c multiquadric collocation method for partial differential equations. *Numer. Methods Partial Differential Eq.*, 19:571–594, 2003.
15. A. Cohen, M.A. Davenport, and D. Leviatan. On the stability and accuracy of least squares approximations. *Found. Comput. Math.*, 13(5):819–834, 2013.
16. O. Davydov and R. Schaback. Error bounds for kernel-based numerical differentiation. in revision, 2013.
17. C. de Boor and B. Swartz. Collocation at Gaussian points. *SIAM J. Numer. Anal.*, 10:582–606, 1973.
18. St. De Marchi and R. Schaback. Stability of kernel-based interpolation. *Adv. in Comp. Math.*, 32:155–161, 2010.
19. B. Fornberg. *A practical guide to pseudospectral methods*, volume 1 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 1996.

20. B. Fornberg and D.M. Sloan. A review of pseudospectral methods for solving partial differential equations. *Acta Numerica*, pages 203–267, 1994.
21. M. Griebel and M. A. Schweitzer, editors. *Meshfree Methods for Partial Differential Equations I-V*, volume 26, 43, 57, 65, 79 of *Lecture Notes in Computational Science and Engineering*. Springer, 2002-2011.
22. W.M. Han and X.P. Meng. Error analysis of the reproducing kernel particle method. *Computer Methods in Applied Mechanics and Engineering*, 190:6157–6181, 2001.
23. Y. C. Hon and R. Schaback. On unsymmetric collocation by radial basis functions. *Appl. Math. Comput.*, 119:177–186, 2001.
24. Y. C. Hon, R. Schaback, and X. Zhou. An adaptive greedy algorithm for solving large RBF collocation problems. *Numer. Algorithms*, 32(1):13–25, 2003.
25. W. Hu, H.Y. Hu, and J.S. Chen. Error analysis of collocation method based on reproducing kernel approximation. *Numerical Methods for Partial Differential Equations*, 27:554–580, 2011. doi10.1002/num.20539.
26. K. Jetter, J. Stöckler, and J.D. Ward. Error estimates for scattered data interpolation on spheres. *Mathematics of Computation*, 68:733–747, 1999.
27. J. Jost. *Partial differential equations*, volume 214 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.
28. E. J. Kansa. Application of Hardy's multiquadric interpolation to hydrodynamics. In *Proc. 1986 Simul. Conf., Vol. 4*, pages 111–117, 1986.
29. D. Levin. The approximation power of moving least-squares. *Mathematics of Computation*, 67:1517–1531, 1998.
30. H. Li and S.S. Mulay. *Meshless methods and their numerical properties*. CRC Press, Boca Raton, FL, 2013.
31. S.F. Li and W.K. Liu. *Meshfree particle methods*. Springer-Verlag, Berlin, 2004.
32. Z.-C. Li, T.-T. Lu, H.-Y. Hu, and A. H.-D. Cheng. *Trefftz and collocation methods*. WIT Press, Southampton, 2008.
33. L. Ling and R. Schaback. On adaptive unsymmetric meshless collocation. In S.N. Atluri and A.J.B. Tadeu, editors, *Proceedings of the 2004 International Conference on Computational & Experimental Engineering and Sciences*, Advances in Computational & Experimental Engineering & Sciences. Tech Science Press, 2004. paper # 270.
34. G. R. Liu. *Meshless methods*. CRC Press, Boca Raton, FL, second edition, 2010.
35. J.M. Melenk. On approximation in meshless methods. In *Frontiers of Numerical Analysis*, Universitext, pages 65–141. Springer, 2005.
36. D. Mirzaei. Private communication, 2014.
37. D. Mirzaei and R. Schaback. Direct Meshless Local Petrov-Galerkin (DMLPG) method: A generalized MLS approximation. *Applied Numerical Mathematics*, 68:73–82, 2013.
38. V.P. Nguyen, T. Rabczuk, S. Bordas, and M. Duflot. Meshless methods: A review and computer implementation aspects. *Mathematics and Computers in Simulation*, 79:763–813, 2008.
39. C. Rieger, B. Zwicknagl, and R. Schaback. Sampling and stability. In M. Dæhlen, M.S. Floater, T. Lyche, J.-L. Merrien, K. Mørken, and L.L. Schumaker, editors, *Mathematical Methods for Curves and Surfaces*, volume 5862 of *Lecture Notes in Computer Science*, pages 347–369, 2010.
40. B. Šarler. From global to local radial basis function collocation method for transport phenomena. In *Advances in meshfree techniques*, volume 5 of *Comput. Methods Appl. Sci.*, pages 257–282. Springer, Dordrecht, 2007.
41. R. Schaback. Improved error bounds for scattered data interpolation by radial basis functions. *Mathematics of Computation*, 68:201–216, 1999.
42. R. Schaback. Convergence of unsymmetric kernel-based meshless collocation methods. *SIAM J. Numer. Anal.*, 45(1):333–351 (electronic), 2007.
43. R. Schaback. Unsymmetric meshless methods for operator equations. *Numer. Math.*, 114:629–651, 2010.
44. R. Schaback. A computational tool for comparing all linear PDE solvers. to appear in Advances of Computational Mathematics, http://www.num.math.uni-goettingen.de/schaback/research/group.html, 2013.
45. R. Schaback. Direct discretizations with applications to meshless methods for PDEs. *Dolomites Research Notes on Approximation, Proceedings of DWCAA12*, 6:37–51, 2013.
46. R. Schaback. Greedy sparse linear approximations of functionals from nodal data. *Numerical Algorithms*, 67:531–547, 2014.

47. R. Schaback and H. Wendland. Adaptive greedy techniques for approximate solution of large RBF systems. *Numer. Algorithms*, 24(3):239–254, 2000.
48. M.A. Schweitzer. *Meshfree and Generalized Finite Element Methods*. Habilitation, Institute for Numerical Simulation, University of Bonn, 2008.
49. Q. Shen. Local RBF-based differential quadrature collocation method for the boundary layer problems. *Eng. Anal. Bound. Elem.*, 34(3):213–228, 2010.
50. C. Shu, H. Ding, and K. S. Yeo. Computation of incompressible Navier-Stokes equations by local RBF-based differential quadrature method. *CMES Comput. Model. Eng. Sci.*, 7(2):195–205, 2005.
51. J. Sladek and V. Sladek. *Advances in meshless methods*. Tech Science Press, 2006.
52. D. Stevens, H. Power, M. Lees, and H. Morvan. The use of PDE centres in the local RBF Hermitian method for 3D convective-diffusion problems. *J. Comput. Phys.*, 228(12):4606–4624, 2009.
53. B. Swartz. Conditioning collocation. *SIAM J. Numer. Anal.*, 25(1):124–147, 1988.
54. B. Swartz and B. Wendroff. The relation between the Galerkin and collocation methods using smooth splines. *SIAM J. Numer. Anal.*, 11:994–996, 1974.
55. H. Wendland. Local polynomial reproduction and moving least squares approximation. *IMA Journal of Numerical Analysis*, 21:285–300, 2001.
56. H. Wendland. *Scattered Data Approximation*. Cambridge University Press, 2005.
57. G.M. Yao, Siraj ul Islam, and B. Šarler. Assessment of global and local meshless methods based on collocation with radial basis functions for parabolic partial differential equations in three dimensions. *Eng. Anal. Bound. Elem.*, 36(11):1640–1648, 2012.