Analytic regularity and nonlinear approximation of a class of parametric semilinear elliptic PDEs

Analytic regularity and nonlinear approximation of a class of parametric semilinear elliptic PDEs*

M. Hansen and Ch. Schwab

Analytic Regularity and Nonlinear Approximation of a class of parametric semilinear elliptic PDEs

Markus Hansen and Christoph Schwab

May 17, 2011

Abstract

We investigate existence and regularity of a class of semilinear, parametric elliptic PDEs with affine dependence of the principal part of the differential operator on countably many parameters. We establish a-priori estimates and analyticity of the parametric solutions. We establish summability results of coefficient sequences of polynomial chaos type expansions of the parametric solutions in terms of tensorized Taylor-, Legendre- and Chebyshev polynomials on the infinite-dimensional parameter domain. We deduce rates of convergence for N term truncated approximations of expansions of the parametric solution. We also deduce spatial regularity of the solution, and establish convergence rates of N-term discretizations of the parametric solutions with respect to these polynomials in parameter space and with respect to a multilevel hierarchy of Finite Element spaces in the spatial domain of the PDE.

Key Words: Semilinear Elliptic Partial Differential Equations, Infinite Dimensional Spaces, N-term approximation, Analyticity in Infinite Dimensional Spaces, Tensor Product Taylor-, Legendre- and Chebyshev polynomial Approximation.
AMS Mathematics Subject Classification: Primary: 35J61 Secondary: 35B30, 65N30, 41A58

Contents

1 Introduction

The numerical solution of partial differential equations (PDEs) with random inputs has received considerable attention in recent years. This is in particular motivated by questions around *uncertainty quantification*, where PDEs with random inputs are to be solved efficiently *for all values taken by the uncertain input parameters*. Depending on the number of parameters, there arises the question of efficient numerical solution of such PDEs on possibly high-dimensional parameter spaces. In the context of PDEs with spatially inhomogeneous, random coefficients, parametric expansions of such coefficients, such as Karhunen-Loève expansions, give rise to *parametric, deterministic PDEs on infinite dimensional parameter spaces* (the parameters being the coefficients in the Karhunen-Loève expansions). For linear partial differential equations, it has been initiated in [14] and extended in [1, 2, 7, 8] to rather general classes of sparse Karhunen-Loève expansions that the solutions of these parametric, deterministic PDEs admit correspondingly *sparse expansions* in terms of tensorized polynomial systems on the infinite dimensional parameter space, under the provision of sparsity of the random inputs. To extend these results to a class of *nonlinear, elliptic PDEs* is the purpose of the present paper.

The analysis in [2, 7, 8] involved in a crucial way the *analytic dependence* of the solution on the parameters; the extension to the present, nonlinear case involves therefore likewise *analyticity* of the nonlinearity in the PDE, in addition to the usual structural conditions such as convexity resp. monotonicity and polynomial growth conditions at infinity.

The main results of the present paper are as follows: we establish, under the assumption of *affine dependence of the differential operator's principal part on the countably many parameters*, that the unique, parametric solution of nonlinear, elliptic PDE depends holomorphically on the countably many parameters in the input data. We then prove that the unique solution of the parametric, nonlinear elliptic PDE admits representations in unconditionally convergent power series expansions about the origin in parameter space as well as unconditionally convergent "polynomial chaos" type expansions with respect to countable families of tensorized orthogonal polynomials on the parameter space. We consider in particular expansions into tensorized Legendre and in Chebyshev polynomial expansions; the latter representation generalizes, in the linear case, the results in [2, 7, 8] and is useful in connection with recent sampling schemes on the parameter domain which are of "compressed sensing type" (see, e.g. [10]).

The outline of this paper is as follows: in Section 2, we introduce the class of problems of interest, in particular its parametric version depending on countably many parameters, and formulate structural assumptions on the problem data and prove basic well-posedness results for this class of PDEs. All parameters as well as data and solutions in this section are real-valued. Section 3 establishes then the existence, uniqueness and holomorphic dependence of solutions for complex-valued parameter vectors $z \in \mathbb{C}^{\mathbb{N}}$. This holomorphic dependence then is exploited in Section 4 for estimating the Taylor coefficients, and establishing their p-summability as well as absolute and uniform convergence of the respective Taylor series. In Section 5 these results are extended to systems of tensorized Legendre- and Chebyshev-polynomials. Under increased regularity assumptions on the inputs, we show in Section 6 that also the solution possesses higher spatial regularity, which carries over to the respective series expansions. In Section 7 we finally investigate convergence rates for N-term approximation of the solution by Taylor-, Legendre- and Chebyshev-partial sums.

2 A class of semilinear elliptic PDEs

In the following we shall consider the Dirichlet boundary value problem

$$-\nabla\big(a\nabla u\big) + G(u) = f \quad \text{in} \quad D\,, \qquad u = 0 \quad \text{on} \quad \partial D\,, \tag{2.1}$$

where $a = a(x)$ is a given fixed function on D.

We will denote the differential operator by T consisting of a linear part $L = -\nabla(a\nabla\cdot)$ and the (nonlinear) composition operator $T_G = G(\cdot)$. Here we shall not strive for utmost generality of the function G. Since we wish to study parametric problems, which (as an intermediate step) will be extended to complex parameters, our approach limits the possible choices for the function $G : \mathbb{C} \longrightarrow \mathbb{C}$. More specifically, since we aim at proving analyticity of the parameter dependence of the solutions, we shall only consider analytic functions G. Taking into account integrability requirements (see below), we are basically left with studying polynomial functions G.

This leads to studying the problems

$$-\nabla\big(a\nabla u\big) + u^m = f \quad \text{in} \quad D \subset \mathbb{R}^n\,,$$
$$u = 0 \quad \text{on} \quad \partial D\,, \tag{2.2}$$

where the necessary restrictions on the power $m \in \mathbb{N}$ and the dimension n of the domain D will be recalled below.

2.1 Basic properties of some composition operators

If we stick to the space $H_0^1(D)$ for solutions of (2.2), then this automatically yields restrictions on the one hand on the dimension n of the domain $D \subset \mathbb{R}^n$ and on the other hand on the exponent m. The weak formulation of (2.2),

$$b_1(u,v) + b_2(u,v) = l(v)\,,$$

where $u, v \in H_0^1(D)$ and

$$b_1(u,v) = \int_D a(x)\nabla u(x)\nabla v(x)dx\,, \qquad b_{2,m}(u,v) = \int_D u(x)^m v(x)dx\,, \qquad l(v) = \int_D f(x)v(x)dx\,,$$

particularly requires $u \in L_{m+1}(D)$ for $b_{2,m}$ to be well-defined on $H_0^1(D) \times H_0^1(D)$, hence we obtain the following restrictions on the range of m from the Sobolev embedding theorem:

$$n = 1 \text{ or } n = 2: \quad m \in \mathbb{N}$$
$$n = 3 : H_0^1(D) \hookrightarrow L_6(D) \qquad \text{hence } 1 \le m \le 5$$
$$n = 4 : H_0^1(D) \hookrightarrow L_4(D) \qquad \text{hence } 1 \le m \le 3$$
$$n = 5 : H_0^1(D) \hookrightarrow L_{10/3}(D) \qquad \text{hence } 1 \le m \le 2$$
$$n = 6 : H_0^1(D) \hookrightarrow L_3(D) \qquad \text{hence } 1 \le m \le 2$$
$$n \ge 7 : \quad m = 1\,.$$

We will assume the parameter m to be in this range thoughout this paper and denote the corresponding set of parameters (n,m) by $\mathcal{M}$. Thus apart from the possibility of restricting ourselves to smoother functions, i.e. $u \in H^s(D)$ for $s > 1$, there is only a very limited number of possible parameters.

In this subsection we shall be concerned with some basic properties of the composition operators T_G associated with the functions $G_m(t) = t^m$. The following lemma contains some mapping properties needed in our further considerations.

Lemma 2.1. Let $(n, m) \in \mathcal{M}$. Further, let $D \subset \mathbb{R}^n$ be a Lipschitz domain, and put $G(t) = t^m$. Then $T_G : H^1(D) \longrightarrow H^{-1}(D)$ is bounded and locally Lipschitz continuous, i.e. Lipschitz continuous on every bounded subset of $H_0^1(D)$.

Proof. By the Sobolev embedding theorem we know $H_0^1(D) \hookrightarrow L_{m+1}(D)$ with constant ι (depending on m and on D) and we find for $u, v \in H_0^1(D)$ by Hölder's inequality

$$\|G(u)|H^{-1}(D)\| = \sup_{\|v|H_0^1(D)\| \leq 1} \left| \int_D u(x)^m v(x) dx \right|$$

$$\leq \sup_{\|v|H_0^1(D)\| \leq 1} \left(\int_D |u(x)|^{m+1} dx \right)^{m/(m+1)} \left(\int_D |v(x)|^{m+1} dx \right)^{1/(m+1)}$$

$$\leq \iota \|u|L_{m+1}(D)\|^m \leq \iota^{m+1} \|u|H_0^1(D)\|^m .$$

Moreover, we find again by Hölder's inequality

$$\|G(u) - G(v)|H^{-1}(D)\| = \sup_{\|w|H_0^1(D)\| \leq 1} \left| \int_D (u(x)^m - v(x)^m) w(x) dx \right|$$

$$= \sup_{\|w|H_0^1(D)\| \leq 1} \left| \int_D (u - v)(u^{m-1} + u^{m-2}v + \cdots + uv^{m-2} + v^{m-1}) w(x) dx \right|$$

$$\leq \sum_{k=0}^{m-1} \sup_{\|w|H_0^1(D)\| \leq 1} \left| \int_D (u - v)(u^{m-1-k}v^k) w(x) dx \right|$$

$$\leq \sum_{k=0}^{m-1} \sup_{\|w|H_0^1(D)\| \leq 1} \|u - v|L_{m+1}(D)\| \cdot \|w|L_{m+1}(D)\| \cdot \left(\int_D |u|^{(m+1)\frac{m-1-k}{m-1}} |v|^{(m+1)\frac{k}{m-1}} dx \right)^{\frac{m-1}{m+1}}$$

$$\leq \iota \sum_{k=0}^{m-1} \|u - v|L_{m+1}(D)\| \cdot \|u|L_{m+1}(D)\|^{m-1-k} \cdot \|v|L_{m+1}(D)\|^k .$$

Thus if $M \subset H_0^1(D)$ is bounded with bound C_M, then T_G is Lipschitz continuous on M with Lipschitz constant at most $m\iota^{m+1} C_M^{m-1}$. $\square$

Remark 2.1. The elementary fact

$$a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + ab^{n-2} + b^{n-1}) \tag{2.3}$$

has been essential in this argument. A more general version of this result can be found in [12], Theorem 5.3.2/1.

Under the ellipticity assumption

$$0 < r \leq a(x) \leq R < \infty, \qquad x \in D, \tag{2.4}$$

the problem (2.2), restricted to real-valued functions a and G, can be seen to have real-valued solutions u for every real-valued right-hand-side f. Without going into details, this follows from the invertibility of L (by Lax-Milgram-theory) and (Lipschitz-)continuity, coercivity and monotonicity of T by standard results on existence of solutions for nonlinear equations, see e.g. [12], Theorem 6.1/1.

We shall, for purposes of analysis, require solutions of (2.1) or (2.2) for complex-valued functions a. We first note that it is easily seen that the identity (2.3) and Lemma 2.1 remain valid for complex variables and complex-valued functions, respectively. Moreover, under the generalized ellipticity assumption

$$0 < r \leq \Re a(x) \leq |a(x)| \leq R < \infty, \tag{2.5}$$

the Lax-Milgram Theorem shows the bounded invertibility of L also in this case. However, the argument above for the existence of solutions for (2.2) then would fail, since the mentioned theorem in [12] as well as modifications thereof are only valid for (reflexive) *real* Banach spaces. Thus in Section 3 we shall use a different approach based on fixed point iterations for spaces over $\mathbb{C}$.

2.2 The parametric problem

As explained in the introduction, our main interest is the analysis of a parametric version of the problems (2.1) or (2.2). More precisely, we shall investigate the parametric problem

$$-\nabla\big(a(y)\nabla u(y)\big) + G\big(u(y)\big) = f(y) \quad \text{in} \quad D, \qquad u = 0 \quad \text{on} \quad \partial D. \tag{2.6}$$

Herein, the mappings $f : U \longrightarrow H^{-1}(D)$ and $a : U \longrightarrow L_\infty(D)$ are given, real-valued. As the parameter domain U we choose $U = [-1,1]^{\mathbb{N}}$, i.e. this can be interpreted as countably many parameters each taking values in the unit interval $[-1,1]$. We are particularly interested in problems, where the dependence of the mapping a on the parameter vector y is *affine*, i.e. it is of the form

$$a(y) \equiv a(\cdot, y), \qquad a(x,y) = \overline{a}(x) + \sum_{j \geq 0} y_j \psi_j(x), \qquad x \in D, y \in U.$$

The functions $\psi_j \in L_\infty(D)$ are assumed given and real-valued, and additional assumptions will be discussed later on.

We shall extend this problem to complex parameters, i.e. we shall assume that f admits an extension to some $H^{-1}(D)$-valued mapping defined on the complex parameter domain

$$\mathcal{U} = \big\{z \in \mathbb{C}^{\mathbb{N}} : |z_j| \leq 1, j \in \mathbb{N}\big\} \equiv \big\{z \in \mathbb{C} : |z| \leq 1\big\}^{\mathbb{N}}.$$

Clearly, an analytic extension of a is given by

$$a(x,z) = \overline{a}(x) + \sum_{j \geq 0} z_j \psi_j(x), \qquad x \in D, z \in \mathcal{U},$$

where we now permit $\overline{a}$ and ψ_j, $j \in \mathbb{N}$, to be complex-valued. Then our main assumption concerning the function a will be the following counterpart of condition (2.5):

Assumption 1. There exist constants $0 < r \leq R < \infty$, such that we have

$$0 < r \leq \Re a(x,z) \leq |a(x,z)| \leq R < \infty \quad \text{for a.e. } x \in D \quad \text{and all } z \in \mathcal{U}. \tag{2.7}$$

In other words, every function $a(z) = a(\cdot, z)$ satisfies condition (2.5), and the respective upper and lower bounds may be chosen independent of $z \in \mathcal{U}$. In view of this last remark, the following assumption would seem more natural

Assumption 2. For every $z \in \mathbb{C}$ there exist constants $0 < r(z) \leq R(z) < \infty$, such that we have

$$0 < r(z) \leq \Re a(x,z) \leq |a(x,z)| \leq R(z) < \infty \quad \text{for a.e. } x \in D. \tag{2.8}$$

6

The relation between these assumptions is the content of the following lemma.

Lemma 2.2. Assume

$$\sup_{z \in \mathcal{U}} \left\| \left| \sum_{j>J} z_j \psi_j \right| L_\infty(D) \right\| \longrightarrow 0 \quad (J \to \infty).$$

Then the Assumptions 1 and 2 are equivalent.

Proof. Obviously, Assumption 1 implies Assumption 2, hence it remains the reverse implication. The crucial step to see this consists in proving that the mapping $z \mapsto a(z)$ is continuous as a mapping from $\mathcal{U}$, equipped with the product topology, into $L_\infty(D)$.

To prove continuity, let $\varepsilon > 0$ and a convergent sequence $(z^k)_{k \in \mathbb{N}} \subset \mathcal{U}$ be given. In the product topology convergence towards $z \in \mathcal{U}$ means that for every $j \in \mathbb{N}$ we have $z_j^k \to z_j$ as $k \to \infty$. Thus for every $J \in \mathbb{N}$ we can find some $k_0(J, \varepsilon)$ such that

$$\sum_{j=1}^{J} \left| z_j^k - z_j \right| \cdot \left\| \psi_j | L_\infty(D) \right\| \leq \frac{\varepsilon}{3} \quad \text{for all} \quad k \geq k_0.$$

Moreover, we find $J_0(\varepsilon)$ such that

$$\sup_{w \in \mathcal{U}} \left\| \left| \sum_{j>J} w_j \psi_j \right| L_\infty(D) \right\| \leq \frac{\varepsilon}{3} \quad \text{for all} \quad J \geq J_0.$$

Then it holds for all $k \geq k_1(\varepsilon) \equiv k_0(J_0(\varepsilon), \varepsilon)$

$$\|a(z^k) - a(z)|L_\infty(D)\| \leq \left\| \left| a(z^k) - \bar{a} - \sum_{j=1}^{J} z_j^k \psi_j \right| L_\infty(D) \right\| + \left\| \left| \sum_{j=1}^{J} z_j^k \psi_j - \sum_{j=1}^{J} z_j \psi_j \right| L_\infty(D) \right\|$$

$$+ \left\| \left| a(z) - \bar{a} - \sum_{j=1}^{J} z_j \psi_j \right| L_\infty(D) \right\|$$

$$\leq \frac{\varepsilon}{3} + \sum_{j=1}^{J} \left| z_j^k - z_j \right| \cdot \left\| \psi_j | L_\infty(D) \right\| + \frac{\varepsilon}{3} \leq \varepsilon.$$

The result itself now is a consequence of the compactness of $\mathcal{U}$ (which in turn follows from Tychonoff's Theorem). The upper estimate $\sup_{z \in \mathcal{U}} R(z) < \infty$ is immediate, and the lower estimate follows from the observation that due to (2.8) also $z \mapsto \frac{1}{a(z)}$ is a well-defined continuous mapping from $\mathcal{U}$ into $L_\infty(D)$ as well as

$$\sup_{z \in \mathcal{U}} r(z)^{-1} = \left(\inf_{z \in \mathcal{U}} r(z) \right)^{-1}, \quad \text{i.e.} \quad \sup_{z \in \mathcal{U}} \frac{1}{r(z)} < \infty \iff \inf_{z \in \mathcal{U}} r(z) > 0.$$

This completes the proof. □

The extra-condition in the above lemma (uniform approximation by finitely many parameters) will later on automatically be satisfied due to assumptions on the functions ψ_j, cf. Theorem 4.1.

In the linear case, i.e. without the term $G(u)$, in problem (2.6), the ellipticity Assumption 1 already suffices to prove existence and uniqueness of solutions for the respective parametric problem for every fixed parameter $z \in \mathcal{U}$ (Lax-Milgram theory). These arguments fail in the above semi-linear case. Instead in the next section we will use an approach via fixed point assertions.

3 Existence and regularity of solutions for complex parameters

In this section we are going to study existence and uniqueness of solutions $u \in H_0^1(D)$ for the semi-linear problem (2.1), and thus also solutions $u(z) \in H_0^1(D)$ for the parametric problem (2.6) for every fixed parameter $z \in \mathcal{U}$. Afterward we shall investigate the regularity of the mapping $z \longmapsto u(z)$.

3.1 A fixed point iteration

The approach is based on the mentioned well-known fact the linear problem

$$-\nabla\big(a\nabla u\big) = f \quad \text{in} \quad D\,, \qquad u = 0 \quad \text{on} \quad \partial D\,, \tag{3.1}$$

has a uniquely determined solution u_{lin} whenever a satisfies condition (2.5). This is an immediate consequence of the Lax-Milgram-Theorem. Upon denoting the corresponding linear differential operator by $T^{\text{lin}} : H_0^1(D) \longrightarrow H^{-1}(D)$, i.e. $T^{\text{lin}}(u) = -\nabla\big(a\nabla u\big)$, and its bounded linear inverse by $S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)$, we further find that every solution u of the semi-linear problem (2.1) satisfies

$$u = S^{\text{lin}}\big(f - G(u)\big) = u_{\text{lin}} - S^{\text{lin}} \circ G(u)\,.$$

This gives rise to a fixed point problem for the (nonlinear) operator $T : H_0^1(D) \longrightarrow H_0^1(D)$, defined by $Tu = u_{\text{lin}} - S^{\text{lin}} \circ G(u)$.

Since S^{lin} is linear, it immediately follows $Tu - Tv = -S^{\text{lin}}\big(G(u) - G(v)\big)$, and

$$\|Tu - Tv|H_0^1(D)\| \leq \|S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\| \cdot \|G(u) - G(v)|H^{-1}(D)\|\,, \quad u, v \in H_0^1(D)\,.$$

In order to show the operator to be a contraction we have to suppose that the composition operator T_G is locally Lipschitz with appropriate bound on the Lipschitz constant.

Theorem 3.1. Let a satisfy the condition (2.5) with lower bound r, and put

$$\rho = \frac{\|f|H^{-1}(D)\|}{r}\,.$$

Let the function G fulfill $G(0) = 0$, and let the composition operator $T_G : H_0^1(D) \longrightarrow H^{-1}(D)$ be bounded and locally Lipschitz. More precisely, assume

$$\|G(u) - G(v)|H^{-1}(D)\| \leq L(2\rho)\|u - v|H_0^1(D)\|$$

for all $u, v \in B_{2\rho}(0) = \big\{w \in H_0^1(D) : \|w|H_0^1(D)\| \leq 2\rho\big\}$, where the constant $L(2\rho)$ satisfies

$$L(2\rho)\|S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\| \leq \tfrac{1}{2}\,. \tag{3.2}$$

Then the problem (2.1) admits a uniquely determined solution $u \in H_0^1(D)$ satisfying

$$\|u - u_{\text{lin}}|H_0^1(D)\| \leq \rho\,,$$

where u_{lin} is the solution of the linear problem (3.1).

Proof. We want to apply Banach's fixed point theorem. Therefore, we consider the operator T on the set $M = B_\rho(u_{\text{lin}}) = \{v \in H_0^1(D) : \|v - u_{\text{lin}}|H_0^1(D)\| \leq \rho\}$. Due to the well-known a priori estimate $\|u_{\text{lin}}|H_0^1(D)\| \leq \rho$, we find $\|v|H_0^1(D)\| \leq 2\rho$ for all $v \in M$. Moreover, M is a closed subset of a Banach space, hence it can be considered as a complete metric space.

Furthermore, we find for $u \in M$

$$\|Tu - u_{\text{lin}}|H_0^1(D)\| = \|S^{\text{lin}} \circ G(u)|H_0^1(D)\| \leq \|S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\| \cdot \|G(u)|H^{-1}(D)\|$$

$$\leq L(2\rho)\|S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\| \cdot \|u|H_0^1(D)\| \leq \rho.$$

Similarly, the assumed local Lipschitz continuity shows that $T : M \longrightarrow M$ is a contraction. Now an application of Banach's fixed point theorem proves the claim, where u is the corresponding fixed point of T in $B_\rho(u_{\text{lin}})$. □

The last assertion of the Theorem, $u \in B_\rho(u_{\text{lin}}) \subset B_{2\rho}(0)$, can be interpreted as an a-priori estimate, hence we obtain

$$\|u|H_0^1(D)\| \leq 2\rho. \tag{3.3}$$

Moreover, when checking whether the constants $L(\varrho)$ satisfy the assumption (3.2), we can further apply the a-priori estimates for solutions of the linear problem (3.1), which can be formulated as $\|S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\| \leq r^{-1}$. Together with Lemma 2.1 we then find

$$L(2\rho)\|S : H^{-1}(D) \longrightarrow H_0^1(D)\| \leq m\iota(m+1)^{m+1}(2\rho)^{m-1} \cdot r^{-1}.$$

Thus (3.2) turns out to be a smallness assumption for ρ and therefore for $f \in H^{-1}(D)$.

Remark 3.1. Note that without the restriction to "solutions close to the linear ones", i.e. without the additional condition

$$\|u - u_{\text{lin}}|H_0^1(D)\| \leq \frac{\|f|H^{-1}(D)\|}{r},$$

the uniqueness of the solution generally fails. For example the (real) problem for $G(t) = -t|t|^{p-2}$ and $a(x) \equiv 1$ (with certain restrictions on p) even has infinitely many solutions with arbitrarily large norms, see [13], Theorem 7.2 and Remark 7.3.

We now return to the parametric prolem (2.6). Then clearly the solution $u(z)$ can be obtained by applying Theorem 3.1 for every fixed parameter $z \in \mathcal{U}$. It only remains to check whether the condition (3.2) can be fulfilled.

Theorem 3.2. Consider the problem (2.1) with $G(\zeta) = \zeta^m$ for $\zeta \in \mathbb{C}$ and $(n,m) \in \mathcal{M}$. Let a satisfy Assumption 1, and suppose

$$\|f(z)|H^{-1}(D)\|^{m-1} < \frac{r^m}{2^m m\iota(m+1)^{m+1}} \quad \text{for all } z \in \mathcal{U}. \tag{3.4}$$

Then problem (2.1) admits a uniquely determined solution $u(z) \in H_0^1(D)$ for every parameter $z \in \mathcal{U}$, satisfying

$$\|u(z) - u_{\text{lin}}(z)|H_0^1(D)\| \leq \frac{M_f}{r} \quad \text{and} \quad \|u(z)|H_0^1(D)\| \leq \frac{2M_f}{r}, \tag{3.5}$$

where $u_{\text{lin}}(z)$ is the solution of the corresponding linear problem and $M_f = \sup_{z \in \mathcal{U}} \|f(z)|H^{-1}(D)\|$.

Note that if $z \mapsto f(z)$ is continuous on $\mathcal{U}$, then automatically $M_f < \infty$, and the smallness condition then reads as $M_f^{m-1} < \frac{r^m}{2^m m\iota(m+1)^{m+1}}$.

3.2 A fixed point iteration for Helmholtz-type problems.

Before we come to a stability assertion for the problem (2.1), we shall have a look at existence of solutions for Helmholtz-type equations

$$-\nabla\big(a\nabla u\big) + bu = f \quad \text{in} \quad D\,, \qquad u = 0 \quad \text{on} \quad \partial D\,, \tag{3.6}$$

where the function b may be complex-valued. These problems are particularly simple cases of the general problem (2.1), where the (nonlinear) composition operator T_G is replaced by a (linear) multiplication operator T^b, where $T^b u = bu$.

Proposition 3.1. Consider the problem (3.6), where a satisfies condition (2.5). Moreover, assume

$$\|bu|H^{-1}(D)\| \le C_b\,\|u|H^1_0(D)\| \tag{3.7}$$

for every $u \in H^1_0(D)$, where the constant $C_b \equiv \|T^b : H^1_0(D) \longrightarrow H^{-1}(D)\|$ satisfies

$$C_b\|S^{\text{lin}} : H^{-1}(D) \longrightarrow H^1_0(D)\| \le \tfrac{1}{2} \tag{3.8}$$

and where the operator S has the same meaning as in the previous section. Then there exists a uniquely determined solution $u \in H^1_0(D)$ of (3.6) satisfying

$$\|u - u_{\text{lin}}|H^1_0(D)\| \le \rho = \frac{\|f|H^{-1}(D)\|}{r} \qquad \text{and} \qquad \|u|H^1_0(D)\| \le 2\rho\,,$$

where $u_{\text{lin}} = S^{\text{lin}} f$ is the unique solution of the problem (3.1).

Proof. As before, we shall use a fixed point iteration to obtain the desired result from Banach's fixed point theorem. To this end, consider for $u \in H^1_0(D)$ the operator T,

$$Tu = u_{\text{lin}} - S^{\text{lin}} \circ T^b u = u_{\text{lin}} - S^{\text{lin}}\big(bu\big)\,.$$

Then we find by Assumption (3.7)

$$\begin{aligned}
\|Tu - u_{\text{lin}}|H^1_0(D)\| &\le \|S^{\text{lin}} : H^{-1}(D) \longrightarrow H^1_0(D)\| \cdot \|bu|H^{-1}(D)\| \\
&\le C_b\|S^{\text{lin}} : H^{-1}(D) \longrightarrow H^1_0(D)\| \cdot \|u|H^1_0(D)\|\,.
\end{aligned}$$

Hence T is a mapping from $B_\rho(u_{\text{lin}})$ into itself, since $\|u_{\text{lin}}|H^1_0(D)\| \le \rho$ and $\|u|H^1_0(D)\| \le 2\rho$ for all $u \in B_\rho\big(u_{\text{lin}}\big)$. Moreover, we similarly obtain

$$\begin{aligned}
\|Tu - Tv|H^1_0(D)\| &= \|S^{\text{lin}} \circ T^b(u - v)|H^1_0(D)\| \\
&\le \|S^{\text{lin}} : H^{-1}(D) \longrightarrow H^1_0(D)\| \cdot \|b(u - v)|H^{-1}(D)\| \\
&\le C_b\|S^{\text{lin}} : H^{-1}(D) \longrightarrow H^1_0(D)\| \cdot \|u - v|H^1_0(D)\|\,,
\end{aligned}$$

which shows that T is a contraction. An application of Banach's fixed point theorem now shows the existence of precisely one fixed point of T in $B_\rho\big(u_{\text{lin}}\big)$, which is the desired solution of (3.6). $\square$

As opposed to the semilinear problem (2.1) we this time have a smallness condition on the input function b. However, in the following application of this result, this in turn will again be reformulated as a smallness condition for f.

Proposition 3.2. Let u and $\widetilde{u}$ be solutions of (2.1) for $G(\zeta) = \zeta^m$ and $(n, m) \in \mathcal{M}$ with data a and $\widetilde{a}$, both satisfying (2.5) with lower bound r, and the right hand sides $f, \widetilde{f}$. If f and $\widetilde{f}$ satisfy the condition

$$\max\left(\|f|H^{-1}(D)\|, \|\widetilde{f}|H^{-1}(D)\|\right) < r^{\frac{m}{m-1}}\left(2^m m\iota(m+1)^{m+1}\right)^{-\frac{1}{m-1}}, \tag{3.9}$$

then it holds

$$\|u - \widetilde{u}|H_0^1(D)\| \le \frac{2}{r}\left(\|f - \widetilde{f}|H^{-1}(D)\| + 2\|a - \widetilde{a}|L_\infty(D)\| \cdot \frac{\|\widetilde{f}|H^{-1}(D)\|}{r}\right).$$

Proof. Subtracting the corresponding problems for u and $\widetilde{u}$ yields with $w = u - \widetilde{u}$

$$\int_D (f - \widetilde{f})v dx = \int_D a\nabla u \cdot \nabla v dx - \int_D \widetilde{a}\nabla\widetilde{u} \cdot \nabla v dx + \int_D \left(G(u) - G(\widetilde{u})\right)v dx$$

$$= \int_D a\nabla w \cdot \nabla v dx + \int_D (a - \widetilde{a})\nabla\widetilde{u} \cdot \nabla v dx + \int_D vw \sum_{j=0}^{m-1} u^{m-1-j}\widetilde{u}^j dx.$$

Defining

$$b(x) = \sum_{j=0}^{m-1} u^{m-1-j}(x)\widetilde{u}^j(x),$$

then $w = u - \widetilde{u}$ is a weak solution of the Helmholtz-type problem

$$-\nabla(a\nabla w) + bw = \ell \quad \text{in } D, \qquad w\big|_{\partial D} = 0,$$

where

$$\ell(v) = \int_D (f - \widetilde{f})v dx + \int_D (\widetilde{a} - a)\nabla\widetilde{u} \cdot \nabla v dx.$$

To apply the previous proposition we have to check condition (3.7). Since we have the a priori estimates $\|u|H_0^1(D)\| \le 2\rho$ and $\|\widetilde{u}|H_0^1(D)\| \le 2\rho$, where $\rho = r^{-1}\max\left(\|f|H^{-1}(D)\|, \|\widetilde{f}|H^{-1}(D)\|\right)$, we find

$$\|T^b w|H^{-1}(D)\| \le \iota(m+1)^{m+1}\sum_{j=0}^{m-1} \|u|H_0^1(D)\|^{m-1-j} \cdot \|\widetilde{u}|H_0^1(D)\|^j \cdot \|w|H_0^1(D)\|$$

$$\le \iota(m+1)^{m+1}m(2\rho)^{m-1} \cdot \|w|H_0^1(D)\|.$$

In other words, T^b is bounded (and hence globally Lipschitz continuous) with norm at most $C := \iota(m+1)^{m+1}m(2\rho)^{m-1}$. Moreover, we obtain

$$\|\ell|H^{-1}(D)\| \le \|f - \widetilde{f}|H^{-1}(D)\| + \|a - \widetilde{a}|L_\infty(D)\| \cdot \|\widetilde{u}|H_0^1(D)\|$$

$$\le \|f - \widetilde{f}|H^{-1}(D)\| + \|a - \widetilde{a}|L_\infty(D)\| \cdot r^{-1}\|\widetilde{f}|H^{-1}(D)\|.$$

In this case, the counterpart of condition (3.7) reads as

$$C\|S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\| \le \tfrac{1}{2},$$

which can be reformulated as the imposed smallness-condition on f. Applying Proposition 3.1 now yields that there is a uniquely determined solution w_0 of (3.6) in $B_{\widetilde{\rho}}(S^{\text{lin}}\ell)$, where $\widetilde{\rho} = r^{-1}\|\ell|H^{-1}(D)\|$.

11

Thus it remains to show that this solution coincides with w. To this end we note that also w is a fixed point of T (recall $Tu = S^{\text{lin}}\ell - S^{\text{lin}} \circ T^b u$), and that T is a contraction on the whole space $H_0^1(D)$. Then we obtain

$$\|w - w_0|H_0^1(D)\| = \|T(w - w_0)|H_0^1(D)\| \leq C_0 \|S^{\text{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\| \cdot \|w - w_0\|\,,$$

which for $w \neq w_0$ yields a contradiction. Finally, the a priori estimate from Proposition 3.1 yields $\|w|H_0^1(D)\| \leq 2\widetilde{\rho}$, which completes the proof. $\qquad\square$

Note that a priori it is not clear, whether $w \in B_{\widetilde{\rho}}(S\ell)$, hence we used this indirect approach to show that the solutions coincide.

3.3 Partial derivatives

For the arguments used in the later sections it will be necessary to further extend the considered parameter domains. We follow [2] and define for $\delta > 0$ domains

$$\mathcal{A}_\delta = \left\{ z \in \mathbb{C}^{\mathbb{N}} : \delta \leq \Re a(x, z) \leq |a(x, z)| \leq R(z) < \infty \text{ for every } x \in D \right\}.$$

If a fulfills Assumption 1, then $\mathcal{A}_\delta \supset \mathcal{U}$ for every $0 < \delta \leq r$. Moreover, we put $\mathcal{A} = \bigcup_{\delta > 0} \mathcal{A}_\delta$, and assume $z \mapsto f(z)$ to admit an extension from $\mathcal{U}$ to $\mathcal{A}$.

Furthermore, we suppose the data $f(z) \in H^{-1}(D)$ to satisfy the smallness-condition

$$\|f(z)|H^{-1}(D)\| < \left(\frac{\delta^m}{2^m m \iota (m+1)^{m+1}} \right)^{\frac{1}{m-1}} \qquad \text{for all } z \in \mathcal{A}_\delta \text{ and all } 0 < \delta \leq r\,. \qquad (3.10)$$

Then the assumptions of both Theorem 3.2 and Proposition 3.2 are fulfilled. Note that there is no further assumption on the function a or the parameter z, apart from the ellipticity condition (2.5). Hence for every $z \in \mathcal{A}_\delta$ we have a uniquely determined solution $u(z) \in H_0^1(D)$ satisfying the estimate

$$\|u(z)|H_0^1(D)\| \leq \frac{2\|f(z)|H^{-1}(D)\|}{\delta} \leq M(\delta) = 2\delta^{-1} \sup_{z \in \mathcal{A}_\delta} \|f(z)|H^{-1}(D)\|\,, \qquad (3.11)$$

where the finiteness of the supremum follows from (3.10).

With this estimates in hand, we begin with a preparation.

Proposition 3.3. Let $(n, m) \in \mathcal{M}$, and consider $G(\zeta) = \zeta^m$ with $m \geq 2$. Then it holds

$$\lim_{h \to 0} \frac{\left\| G(u(z + he^j)) - G(u(z)) - G'(u(z))(u(z + he^j) - u(z)) \big| H^{-1}(D) \right\|}{|h|} = 0\,.$$

Note that although this expression appears reminiscent of Fréchet or Gâteaux differentiability, it does not imply either (neither is $G'(u(z))(u(z + \cdot e^j) - u(z))$ a (bounded) linear mapping, nor is $u(z + he^j) - u(z)$ a (fixed) direction).

Proof. Fix $z \in \mathcal{A}$, and choose $\delta > 0$ such that $z \in \mathcal{A}_\delta$. First we note that then for $h \in \mathbb{C} \setminus \{0\}$ we have $z + he^j \in \mathcal{A}_{\delta/2}$: With $z \in \mathcal{A}_\delta$ and $|h| \cdot \|\psi_j|L_\infty(D)\| < \delta/2$ we find

$$\Re a(x, z + he^j) \geq \Re a(x, z) - |h| \cdot \|\psi_j|L_\infty(D)\| \geq \delta - \delta/2 = \delta/2$$

as well as

$$|a(x, z + he^j)| \leq |a(x, z)| + |h| \cdot \|\psi_j|L_\infty(D)\| \leq R(z) + \delta/2 < \infty.$$

In particular, this yields $u(z + he^j) \in H_0^1(D)$ to be well-defined. Next we recall that by Taylor's Theorem it holds

$$g(z) - g(z_0) - g'(z_0)(z - z_0) = (z - z_0)^2 \int_0^1 (1 - t)g''\big(z_0 + t(z - z_0)\big)dt$$

for every analytic function g (this version easily follows from the $\mathbb{C}$-valued real-variable version by considering the Taylor polynomial for the mapping $t \mapsto g(z_0 + t(z - z_0))$ around 0, evaluated in $t = 1$). In particular, this can be applied to the polynomial G. Thus we can apply Taylor's formula pointwise (where z corresponds to $u(z + he^j)$ and z_0 corresponds to $u(z)$) to obtain

$$G\big(u(z + he^j)\big) - G\big(u(z)\big) - G'\big(u(z)\big)\big(u(z + he^j) - u(z)\big)$$
$$= \big(u(z + he^j) - u(z)\big)^2 \int_0^1 (1 - t)G''\big(u(z) + t(u(z + he^j) - u(z))\big)dt.$$

Multiplying with $v \in H_0^1(D)$ and integrating over D yields

$$\int_D \left|\Big(G(u(z + he^j)) - G(u(z)) - G'(u(z))(u(z + he^j) - u(z))\Big)v(x)\right|dx$$
$$= \int_D \left|u(z + he^j) - u(z)\right|^2 |v(x)| \int_0^1 \left|(1 - t)G''\big(u(z) + t(u(z + he^j) - u(z))\big)\right|dt dx$$
$$\leq \|u(z + he^j) - u(z)|L_{m+1}(D)\|^2 \cdot \|v(x)|L_{m+1}(D)\|$$
$$\times \left\|\left|\int_0^1 \left|(1 - t)G''\big(u(z) + t(u(z + he^j) - u(y))\big)\right|dt\right|L_{\frac{m+1}{m-2}}(D)\right\|,$$

where in the last line we used Hölder's inequality (clearly the case $m = 2$ is trivial). The last factor can be further estimated using the (generalized) Minkowski inequality. Then we obtain

$$\left\|\left|\int_0^1 \left|(1 - t)G''\big(u(z) + t(u(z + he^j) - u(z))\big)\right|dt\right|L_{\frac{m+1}{m-2}}(D)\right\|$$
$$\leq \int_0^1 \left\|\left|(1 - t)G''\big(u(z) + t(u(z + he^j) - u(z))\big)\right|L_{\frac{m+1}{m-2}}(D)\right\|dt$$
$$= \int_0^1 (1 - t)m(m - 1)\left\|\left|u(z) + t(u(z + he^j) - u(z))\right|^{m-2}\right|L_{\frac{m+1}{m-2}}(D)\right\|dt$$
$$\leq \int_0^1 (1 - t)m(m - 1)\Big((1 - t)\|u(z)|L_{m+1}(D)\| + t\|u(z + he^j)|L_{m+1}(D)\|\Big)^{m-2}dt$$
$$\leq m(m - 1)\int_0^1 (1 - t)dt\Big(\iota(m + 1)M(\delta)\Big)^{m-2} = \frac{1}{2}m(m - 1)\Big(\iota(m + 1)M(\delta)\Big)^{m-2},$$

where at the end we used the a priori estimate (Theorems 3.1 and 3.2), and $\iota(m+1)$ is the embedding constant for $H_0^1(D) \hookrightarrow L_{m+1}(D)$. Combining these two estimates yields together with the stability estimate (Proposition 3.2)

$$\left\|G(u(z + he^j)) - G(u(z)) - G'(u(z))(u(z + he^j) - u(z))|H^{-1}(D)\right\|$$

$$\leq \tfrac{1}{2}m(m-1)\iota(m+1)^{m+1}\|u(z+he^j)-u(z)|H_0^1(D)\|^2 \cdot M(\delta)^{m-2}$$
$$\leq \tfrac{1}{2}m(m-1)\iota(m+1)^{m+1}M(\delta)^{m-2}4\delta^{-2}$$
$$\times \left(\|f(z+he^j)-f(z)|H^{-1}(D)\| + \|a(\cdot,z+h)-a(\cdot,z)|L_\infty(D)\|M(\delta) \right)^2$$
$$\leq 2m(m-1)\iota(m+1)^{m+1}M(\delta)^{m-2}\delta^{-2}\left(\mathcal{O}(|h|) + |h|\cdot\|\psi_j|L_\infty(D)\|M(\delta) \right)^2 = \mathcal{O}(|h|^2)\,.$$

This proves the claim. $\qquad\square$

We now aim at proving existence of partial derivatives for the mapping $z \mapsto u(z)$. To do so we require an additional assumption on f.

Assumption 3. We assume the mapping $z \mapsto f(z)$ to admit partial derivatives $\partial_{z_j}f(z) \in H^{-1}(D)$ at every point $z \in \mathcal{A}$ with respect to every variable z_j, $j \in \mathbb{N}$.

Then the desired regularity result for $z \mapsto u(z)$ reads as follows.

Theorem 3.3. Let $z \mapsto f(z) \in H^{-1}(D)$ satisfy the condition (3.10) and Assumption 3. Then at any point $z \in \mathcal{A}$, the function $z \mapsto u(z)$ admits partial derivatives $\partial_{z_j}u(z) \in H_0^1(D)$ with respect to each complex variable z_j. For each $j \in \mathbb{N}$, $\partial_{z_j}u$ is the unique solution in $H_0^1(D)$ of the linear problem:

$$\int_D a(x,z)\nabla w(x) \cdot \nabla v(x)dx + \int_D G'\big(u(x,z)\big)w(x)v(x)dx$$
$$= \int_D \partial_{z_j}f(x,z)v(x,z)dx - \int_D \psi_j(x)\nabla u(x,z) \cdot \nabla v(x)dx \qquad \text{for all } v \in H_0^1(D)\,.$$

Proof. Fix again $z \in \mathcal{A}$, and choose $\delta > 0$ such that $z \in \mathcal{A}_\delta$. As in the proof of Proposition 3.3 we have $z + he^j \in \mathcal{A}_{\delta/2}$ for $|h|$ sufficiently small. For such h, the difference quotient

$$w_h(z) = \frac{u(z+he^j)-u(z)}{h} \in H_0^1(D)$$

is well-defined and we find for arbitrary $v \in H_0^1(D)$ and such h that

$$h\int_D \frac{f(x,z+he^j)-f(x,z)-h\partial_{z_j}f(x,z)}{h}v(x)dx + h\int_D \partial_{z_j}f(x,z)v(x)dx$$
$$= \int_D a(x,z+he^j)\nabla u(x,z+he^j)\cdot\nabla v(x)dx - \int_D a(x,z)\nabla u(x,z)\cdot\nabla v(x)dx$$
$$\quad + \int_D \Big(G(u(x,z+he^j))-G(u(x,z))\Big)v(x)dx$$
$$= h\int_D a(x,z)\nabla w_h(x,z)\cdot\nabla v(x)dx + \int_D \big(a(x,z+he^j)-a(x,z)\big)\nabla u(x,z+he^j)\cdot\nabla v(x)dx$$
$$\quad + \int_D \Big(G(u(x,z+he^j))-G(u(x,z))\Big)v(x)dx$$
$$= h\int_D a(x,z)\nabla w_h(x,z)\cdot\nabla v(x)dx + h\int_D \psi_j(x)\nabla u(x,z+he^j)\cdot\nabla v(x)dx$$
$$\quad + h\int_D \frac{G(u(x,z+he^j))-G(u(x,z))-G'(u(x,z))(u(x,z+he^j)-u(x,z))}{h}v(x)dx$$

14

$$+ h \int_D G'(u(x,z))w_h(x,z)v(x)dx \,.$$

By Proposition 3.3 the third integral in the last line converges to 0 as $h \to 0$, and the first integral on the left hand side converges to 0 as well due to Assumption 3.

If we define $\ell_h(v) = \int_D \psi_j(x)\nabla u(x, z + he^j) \cdot \nabla v(x)dx$ and $\ell_0(v) = \int_D \psi_j(x)\nabla u(x,z) \cdot \nabla v(x)dx$, then this yields bounded linear functionals on $H_0^1(D)$, and together with (3.11) we find

$$
\begin{aligned}
|\ell_h(v) - \ell_0(v)| &= \left| \int_D \psi_j\big(\nabla u(z + he^j) - \nabla u(z)\big) \cdot \nabla v\, dx \right| \\
&\leq \|\psi_j|L_\infty(D)\| \cdot \|u(z + he^j) - u(z)|H_0^1(D)\| \cdot \|v|H_0^1(D)\| \\
&\leq \frac{1}{\delta}\|\psi_j|L_\infty(D)\|\Big(\|f(z + he^j) - f(z)|H^{-1}(D)\| + |h| \cdot \|\psi_j|L_\infty(D)\|M(\delta)\Big)\|v|H_0^1(D)\| \,.
\end{aligned}
$$

In the last line we applied the stability estimate (Proposition 3.2). We conclude $\ell_h \to \ell_0$ in $H^{-1}(D)$ as $h \to 0$. Furthermore, defining

$$
\begin{aligned}
L_h(v) = &\int_D \frac{f(x, z + he^j) - f(x,z) - h\partial_{z_j}f(x,z)}{h}v(x)dx - \int_D \psi_j(x)\nabla u(x, z + he^j) \cdot \nabla v(x)dx \\
&- \int_D \frac{G(u(x, z + he^j)) - G(u(x,z)) - G'(u(x,z))(u(x, z + he^j) - u(x,z))}{h}v(x)dx
\end{aligned}
$$

this yields another bounded linear functional on $H_0^1(D)$, and we find $L_h \to \ell_0$ in $H^{-1}(D)$ as $h \to 0$. The next step consists in an application of Proposition 3.1 with $b(x,z) = G'(u(x,z))$, hence we have to check condition (3.7). Similar to Lemma 2.1, applying Hölder's inquality yields

$$\|G'(u(z))v|H^{-1}(D)\| \leq m\iota(m+1)^{m+1}\|u(z)|H_0^1(D)\|^{m-1} \cdot \|v|H_0^1(D)\| \,, \qquad v \in H_0^1(D) \,,$$

hence (3.7) is fulfilled with $C_b \mathrel{\widehat{=}} (2\rho(z))^{m-1}m\iota(m+1)^{m+1}$, where $\rho(z) = \delta^{-1}\|f(z)|H^{-1}(D)\|$. Thus condition (3.8) reduces once more to a smallness-condition for $f(z)$,

$$\|f(z)|H^{-1}(D)\| \leq \delta\big(2^m m\iota(m+1)^{m+1}\|S_z^{\mathrm{lin}} : H^{-1}(D) \longrightarrow H_0^1(D)\|\big)^{-1/(m-1)} \,,$$

which in turn is satisfied due to assumption (3.10). This implies that the function w_h is the unique weak solution of the problem

$$\int_D a(x,z)\nabla w_h(x,z) \cdot \nabla v(x)dx + \int_D G'(u(x,z))w_h(x,z)v(x)dx = \int_D \partial_{z_j}f(x,z)v(x)dx + L_h(v) \,.$$

Similarly, w_0 is the unique weak solution of the problem

$$\int_D a(x,z)\nabla w_0(x,z) \cdot \nabla v(x)dx + \int_D G'(u(x,z))w_0(x,z)v(x)dx = \int_D \partial_{z_j}f(x,z)v(x)dx + \ell_0(v) \,.$$

Combining these two problems, we once more obtain from Proposition 3.1 together with Theorem 3.1 the estimate

$$\|w_h(z) - w_0(z)|H_0^1(D)\| \leq \frac{\|L_h - \ell_0|H^{-1}(D)\|}{r} \to 0 \quad (h \to 0) \,.$$

Thus $\partial_{z_j}u(z) = w_0(z) = \lim_{h\to 0} w_h(z)$ exists in $H_0^1(D)$. $\qquad\square$

4   Estimates of Taylor coefficents

From now on we assume the functions a and f to satisfy Assumption 1, the smallness-condition (3.10) and Assumption 3, respectively.

4.1 δ-admissible sequences

We remind on the domains

$$\mathcal{A}_\delta = \left\{ z \in \mathbb{C}^\mathbb{N} : \delta \leq \Re a(x,z) \leq |a(x,z)| \leq R(z) < \infty \text{ for every } x \in D \right\}.$$

where $\delta > 0$, and $\mathcal{U} \subset \mathcal{A}_\delta$ for $\delta \leq r$ as a consequence of Assumption 1. Following [2] we call a sequence $\rho = (\rho_j)_{j \in \mathbb{N}}$ of positive real numbers δ-admissible, if

$$\sum_{j=1}^\infty \rho_j |\psi_j(x)| \leq \Re\, \overline{a}(x) - \delta$$

for almost every $x \in D$. This particularly implies that the poly-disc

$$\mathcal{U}_\rho = \prod_{j \in \mathbb{N}} \left\{ \zeta \in \mathbb{C} : |\zeta| \leq \rho_j \right\}$$

is contained in $\mathcal{A}_\delta$. Indeed, we find for $z \in \mathcal{U}_\rho$ and almost every $x \in D$

$$\Re a(x,z) \geq \Re \overline{a}(x) - \sum_{j=1}^\infty |z_j| \cdot |\psi_j(x)| \geq \Re \overline{a}(x) - \sum_{j=1}^\infty \rho_j |\psi_j(x)| \geq \delta$$

as well as

$$|a(x,z)| \leq |\overline{a}(x)| + \sum_{j=1}^\infty |z_j| \cdot |\psi_j(x)| \leq |\,\overline{a}(x)| + \Re \overline{a}(x) - \delta \leq 2|\overline{a}(x)| \leq 2R.$$

Similarly, we put

$$\mathcal{U}_{\rho,E} = \prod_{j \in E} \left\{ \zeta \in \mathbb{C} : |\zeta| \leq \rho_j \right\}$$

for every (finite) set $E \subset \mathbb{N}$.

4.2   A preliminary estimate for partial derivatives of u

We first prove a preliminary result for estimates for partial derivatives of $u(y)$ (and hence for its Taylor coefficients) for general δ-admissible sequences ρ.

Before we come to the next proposition, we shall need some more notation. We denote by

$$\mathcal{F} = \left\{ \nu = (\nu_j)_{j \in \mathbb{N}} \in \mathbb{N}_0^\mathbb{N} : |\operatorname{supp} \nu| < \infty \right\}$$

the set of all sequences of non-negative integers with at most finitely many non-vanishing components. Here

$$\operatorname{supp} \nu = \left\{ j \in \mathbb{N} : \nu_j \neq 0 \right\}$$

denotes the support of ν, and we put $|\nu| = \sum_{j \in \mathbb{N}} \nu_j$, which is finite if, and only if, $|\operatorname{supp} \nu| < \infty$.

Proposition 4.1. Let $f \in H^{-1}(D)$ satisfy the condition (3.10). Let a δ-admissible sequence $\rho = (\rho_j)_{j \in \mathbb{N}}$ be given. Then for every $\nu \in \mathcal{F}$ we have the estimate

$$\|\partial^\nu u(y)|H_0^1(D)\| \leq \nu! M(\delta) \prod_{j \in \operatorname{supp} \nu} \rho_j^{-\nu_j}.$$

Proof. The proof is based on the argument in [2, Lemma 2.4].

Put $E = \operatorname{supp} \nu$ and $J = |E|$. Writing $z = (z_E, z_{E^c})$, i.e. $z_E \in \mathbb{C}^J$ contains the components corresponding to inidces $j \in E$, the δ-admissibility of ρ then mplies the estimate

$$\|u(z_E, 0)|H_0^1(D)\| \leq \frac{2\|f(z_E, 0)|H^{-1}(D)\|}{\delta} \leq M(\delta) \tag{4.1}$$

for every $z_E \in \mathcal{U}_{\rho, E}$. W.l.o.g. we assume $E = \{1, \ldots, J\}$, which always may be achieved by renumerating the variables. If we further define the sequence $\widetilde{\rho}$ by

$$\widetilde{\rho}_j = \rho_j + \varepsilon, \quad j \in E, \quad \varepsilon = \frac{\delta}{2 \sum_{j \in E} \|\psi_j|L_\infty(D)\|}, \quad \widetilde{\rho}_j = \rho_j, \quad j \notin E,$$

it is easily checked that $\widetilde{\rho}$ is $\delta/2$-admissible. In particular, u_E is analytic in an open neighbourhood of $\mathcal{U}_{\rho, E}$, where writing $u_E(z_1, \ldots, z_J) = u_E(z_E) \equiv u(z_E, 0)$.

We may thus apply Cauchy's integral formula (see, e.g., [6]) in each variable z_j, $j \in E$, to obtain

$$u_E(z_1, \ldots, z_J) = (2\pi i)^{-J} \oint_{\Gamma_1(z_1)} \cdots \oint_{\Gamma_J(z_J)} \frac{u(z_E', 0)}{(z_1' - z_1) \cdots (z_J' - z_J)} dz_1' \cdots dz_J'$$

$$= (2\pi i)^{-J} \oint_{\Gamma_1} \cdots \oint_{\Gamma_J} \frac{u(z_1' + z_1, \ldots, z_J' + z_J)}{z_1' \cdots z_J'} dz_1' \cdots dz_J'.$$

where $\Gamma_j(z_j)$ denotes the circle with radius ρ_j and center z_j, and $\Gamma_j \equiv \Gamma_J(0)$. Differentiation (or directly applying the formula for derivatives) then yields

$$\partial^\nu u(0) = \frac{\partial^{|\nu|} u}{\partial z_1^{\nu_1} \cdots \partial z_J^{\nu_J}}(0) = \nu! (2\pi i)^{-J} \oint_{\Gamma_1} \cdots \oint_{\Gamma_J} \frac{u(z_1, \ldots, z_J)}{z_1^{\nu_1 + 1} \cdots z_J^{\nu_J + 1}} dz_1 \cdots dz_J.$$

Eventually, together with (4.1) we conclude

$$\|\partial^\nu u(0)|H_0^1(D)\| \leq \nu! M(\delta) \prod_{j \in E} \rho_j^{-\nu_j}.$$

This proves the claim. □

We put emphasis on the observation that Proposition 4.1 is valid for *arbitrary* δ-admissible sequences. Hence the resulting estimate may be improved to

$$\|\partial^\nu u(0)|H_0^1(D)\| \leq \nu! M(\delta) \inf_{\rho \ \delta\text{-admissible}} \prod_{j \in \operatorname{supp} \nu} \rho_j^{-\nu_j}.$$

Moreover, note that the optimal sequence ρ (if it exists) may be a different one for every $\nu \in \mathcal{F}$. We eventually aim at proving ℓ_p-summability results for the Taylor coefficients. For this purpose it will be sufficient to construct "suitable" δ-admissible sequences $\rho \equiv \rho(\nu)$, upon which the respective summability result will be based.

4.3 Construction of $\delta/2$-admissible sequences

The construction is essentially based on corresponding arguments in [2]. The starting point for our construction will be a given fixed δ-admissible sequence ϱ. For given multiindex $\nu \in \mathcal{F}$ we will now construct a $\delta/2$-admissible sequence ρ. We shall use the abbreviation $\gamma_j = \|\psi_j|L_\infty(D)\|$. The basic assumption then reads as

$$\big(\varrho_j\|\psi_j|L_\infty(D)\|\big)_{j\in\mathbb{N}} \in \ell_1(\mathbb{N})\,, \qquad \text{i.e.} \quad \sum_{j\geq 1} \varrho_j\gamma_j < \infty\,. \tag{4.2}$$

To begin with choose $M \in \mathbb{N}$ such that

$$\sum_{j>M} \varrho_j\gamma_j \leq \frac{\delta}{12}\,,$$

which exists due to the choice of ϱ. Without loss of generality, we assume that the indexing of the parameters y_j is chosen such that the sequence $\big(\varrho_j\gamma_j\big)_{j\in\mathbb{N}}$ is non-increasing. Then we partition $\mathbb{N}$ into two sets $E = \{1, \ldots, M\}$ and $F = \mathbb{N} \setminus E$. We further choose $\kappa > 1$ such that

$$(\kappa - 1) \sum_{j\leq M} \varrho_j\gamma_j \leq \frac{\delta}{4}\,.$$

Finally, we define our sequence ρ by

$$\rho_j = \kappa\varrho_j\,, \quad j \in E\,; \qquad \rho_j = \max\Big(\varrho_j, \frac{\delta\nu_j}{4|\nu_F|\gamma_j}\Big)\,, \quad j \in F\,,$$

where ν_E denotes the restriction of ν to the index set E, and $|\nu_F|$ denotes the ℓ_1-norm of the multiindex, i.e. $|\nu_F| = \sum_{j>M} \nu_j$. (with the convention $\frac{\nu_j}{|\nu_F|} = 0$ if $|\nu_F| = 0$).

Now we first check that this sequence ρ indeed is $\delta/2$-admissible. To do so, we estimate

$$\begin{aligned}
\sum_{j\geq 1} \rho_j|\psi_j(x)| &= \kappa \sum_{j\leq J} \varrho_j|\psi_j(x)| + \sum_{j>J} \max\Big(\varrho_j, \frac{\delta\nu_j}{4|\nu_F|\gamma_j}\Big)|\psi_j(x)| \\
&\leq (\kappa - 1) \sum_{j\leq J} \varrho_j\gamma_j + \sum_{j\leq J} \varrho_j|\psi_j(x)| + \sum_{j>J} \Big(\varrho_j + \frac{\delta\nu_j}{4|\nu_F|\gamma_j}\Big)|\psi_j(x)| \\
&\leq \frac{\delta}{4} + \sum_{j\in\mathbb{N}} \varrho_j|\psi_j(x)| + \frac{\delta}{4}\,,
\end{aligned}$$

and hence by choice of ϱ

$$\sum_{j\geq 1} \rho_j|\psi_j(x)| \leq \sum_{j\in\mathbb{N}} \varrho_j|\psi_j(x)| + \frac{\delta}{2} \leq \Re\overline{a}(x) - \frac{\delta}{2}\,.$$

Then we obtain by Proposition 4.1 the estimate

$$\|t_\nu|H_0^1(D)\| \leq M(\delta/2)\rho^{-\nu} \leq M(\delta/2)\Big(\prod_{j\in E} \eta^{\nu_j}\varrho_j^{-\nu_j}\Big)\Big(\prod_{j\in F} \Big(\frac{|\nu_F|d_j}{\nu_j}\Big)^{\nu_j}\varrho_j^{-\nu_j}\Big)\,,$$

18

where $\eta = \frac{1}{\kappa} < 1$ and $d_j = \frac{4\varrho_j \gamma_j}{\delta}$. Moreover, factors with exponent $\nu_j = 0$ are understood to be 1. This estimate can be equivalently written as

$$\varrho^\nu \|t_\nu|H_0^1(D)\| \leq M(\delta/2)\left(\prod_{j \in E} \eta^{\nu_j}\right)\left(\prod_{j \in F} \left(\frac{|\nu_F| d_j}{\nu_j}\right)^{\nu_j}\right). \tag{4.3}$$

Finally, we note that the choice of M implies

$$\|d|\ell_1(F)\| = \sum_{j > M} d_j \leq \frac{1}{3}. \tag{4.4}$$

4.4 Summability of Taylor coefficients

The construction in the previous section are the basis for the following theorem. We shall follow closely the argument given in [2] (Sections 2.2, 3.2–3.3 and 4.4). Before stating the main theorem, we mention the following basic result of [1].

Proposition 4.2. Given $0 < p < 1$, it holds $\left(\frac{|\nu|!}{\nu!} b^\nu\right)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$ if, and only if, $\sum_{j \in \mathbb{N}} b_j < 1$ and $(b_j)_{j \in \mathbb{N}} \in \ell_p(\mathbb{N})$.

In the sequel we shall use the following notion: We say that a sequence $(\Lambda_N)_{N \in \mathbb{N}}$ of finite subsets of $\mathcal{F}$ *exhausts* $\mathcal{F}$, if every finite set $\Lambda \subset \mathcal{F}$ is contained in every set Λ_N for all $N \geq N_0$ (with $N_0 = N_0(\Lambda)$ chosen sufficiently large).

Theorem 4.1. Suppose f satisfies condition (3.10). Moreover, let $\rho = (\rho_j)_{j \in \mathbb{N}}$ be an arbitrary δ-admissible sequence, and assume

$$\left(\|\psi_j|L_\infty(D)\|\right)_{j \in \mathbb{N}} \in \ell_1(\mathbb{N}) \qquad \text{and} \qquad \left(\rho_j \|\psi_j|L_\infty(D)\|\right)_{j \in \mathbb{N}} \in \ell_p(\mathbb{N}) \quad \text{for some } p < 1. \tag{4.5}$$

Then the Taylor coefficients t_ν of the solution u of (2.1) satisfy $\left(\rho^\nu \|t_\nu|H_0^1(D)\|\right)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$.
If additionally

$$\sup_{z \in \mathcal{U}_\rho} \|f(z) - f(z_1, \ldots, z_J, 0, \ldots,)|H^{-1}(D)\| \longrightarrow 0 \quad (J \longrightarrow \infty), \tag{4.6}$$

then it holds

$$\sum_{\nu \in \mathcal{F}} t_\nu z^\nu = u(z), \qquad z \in \mathcal{U}_\rho,$$

with uniform and unconditional convergence, which has to be understood in the following sense: If $(\Lambda_N)_{N \geq 1}$ exhausts $\mathcal{F}$, then the partial sums $S_{\Lambda_N} u(z) = \sum_{\nu \in \Lambda_N} t_\nu z^\nu$ satisfy

$$\lim_{N \to \infty} \sup_{z \in \mathcal{U}_\rho} \|u(z) - S_{\Lambda_N} u(z)|H_0^1(D)\| = 0.$$

Proof. **Step 1: Proof of the ℓ_p-summability.**
We first note that every δ-admissible sequence may serve as the fixed reference sequence ϱ in the constructions of the last section. Hence, we conclude from (4.3)

$$\sum_{\nu \in \mathcal{F}} \left(\rho^\nu \|t_\nu|H_0^1(D)\|\right)^p \leq M(\delta/2)^p \sum_{\nu \in \mathcal{F}} \left(\prod_{j \in E} \eta^{\nu_j}\right)^p \left(\prod_{j \in F} \left(\frac{|\nu_F| d_j}{\nu_j}\right)^{\nu_j}\right)^p$$

19

$$\equiv M(\delta/2)^p \bigg(\sum_{\nu \in \mathcal{F}_E} \alpha(\nu)^p \bigg) \bigg(\sum_{\nu \in \mathcal{F}_F} \beta(\nu)^p \bigg) \equiv M(\delta/2)^p A_E A_F \,,$$

where $\mathcal{F}_E = \{\nu \in \mathcal{F} : \operatorname{supp}\nu \subset E\}$ and $\mathcal{F}_F = \{\nu \in \mathcal{F} : \operatorname{supp}\nu \subset F\}$. Then it further follows

$$A_E = \bigg(\sum_{\nu \in \mathcal{F}_E} \alpha(\nu)^p \bigg) = \sum_{\nu \in \mathcal{F}_E} \prod_{j \in E} \eta^{\nu_j p} = \prod_{j \in E} \sum_{n=0}^{\infty} \eta^{np} = \bigg(\frac{1}{1 - \eta^p} \bigg)^M \,,$$

recall $\eta < 1$ (since E is finite and $|E| = M$ the index set $\mathcal{F}_E$ may be identified with $\mathbb{N}_0^M$). Now we turn to showing $A_F < \infty$. Using as before the convention $0^0 = 1$ and $d^{\nu_F} = \prod_{j \in F} d_j^{\nu_j}$ we find

$$\beta(\nu) = \prod_{j \in F} \Big(\frac{|\nu_F| d_j}{\nu_j} \Big)^{\nu_j} = \frac{|\nu_F|^{|\nu_F|}}{\prod_{j \in F} \nu_j^{\nu_j}} d^{\nu_F} \,, \qquad \nu \in \mathcal{F}_F \,. \tag{4.7}$$

Applying the Stirling inequalities

$$\frac{n! e^n}{e \sqrt{n}} \leq n^n \leq \frac{n! e^n}{\sqrt{2\pi n}} \,, \qquad n \geq 1 \,,$$

we can further estimate the numerator and denominator in (4.7),

$$|\nu_F|^{|\nu_F|} \leq |\nu_F|! e^{|\nu_F|} \qquad \text{and} \qquad \prod_{j \in F} \nu_j^{\nu_j} \geq \frac{\nu_F! e^{|\nu_F|}}{\prod_{j \in F} \max(1, e\sqrt{\nu_j})} \geq \frac{\nu_F! e^{|\nu_F|}}{\prod_{j \in F} e^{\nu_j}} \,,$$

where at the end we used the bound $\max(1, e\sqrt{n}) \leq e^n$. Altogether we then obtain from (4.7)

$$\beta(\nu) \leq \frac{|\nu_F|!}{\nu_F!} e^{|\nu_F|} d^{\nu_F} = \frac{|\nu_F|!}{\nu_F!} \prod_{j \in F} (e d_j)^{\nu_j} \,.$$

We next apply Proposition 4.2 to the sequence $(e d_j)_{j \in \mathbb{N}}$. The assumptions of Proposition 4.2 are satisfied due to (4.4), $e < 3$, and condition (4.5). Eventually, this yields $\sum_{\nu \in \mathcal{F}_F} \beta(\nu)^p < \infty$, and thus the asserted summability of the Taylor coefficients.

Step 2: Convergence of the Taylor series.
We have just shown $\big(\rho^\nu \|t_\nu | H_0^1(D)\| \big)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$, which particularly implies $\big(\rho^\nu \|t_\nu | H_0^1(D)\| \big)_{\nu \in \mathcal{F}} \in \ell_1(\mathcal{F})$ and thus the absolute convergence of the Taylor series $\sum_{\nu \in \mathcal{F}} t_\nu z^\nu$ on $\mathcal{U}_\rho$. We now show its convergence towards $u(z)$ in two substeps: First we reduce it to the case of a finite number of parameters z_j, and then prove the convergence in that special case.

Substep 2.1: Reduction to a finite number of parameters.
If we put

$$a_J(\cdot, z_1, \ldots, z_J) = \overline{a}(\cdot) + \sum_{1 \leq j \leq J} z_j \psi_j(\cdot) \,,$$

we obtain

$$\big\| a(\cdot, z) - a_J(\cdot, z_1, \ldots, z_J) \big| L_\infty(D) \big\| = \bigg\| \sum_{j > J} z_j \psi_j \Big| L_\infty(D) \bigg\| \leq \sum_{j > J} |z_j| \cdot \big\| \psi_j | L_\infty(D) \big\|$$

and thus

$$\sup_{z \in \mathcal{U}_\rho} \left\| a(\cdot, z) - a_J(\cdot, z_1, \dots, z_J) \big| L_\infty(D) \right\| \le \sum_{j>J} \rho_j \|\psi_j | L_\infty(D)\| \longrightarrow 0 \quad (J \longrightarrow \infty) \tag{4.8}$$

due to condition (4.5). If we denote by u_J the restriction of u to J parameters, i.e.

$$u_J(z_1, \dots, z_J) = u(z_1, \dots, z_J, 0, \dots) \equiv u(z_{(J)}),$$

then it follows from Proposition 3.2 and δ-admissibility of ρ

$$\left\| u(z) - u_J(z_1, \dots, z_J) \big| H_0^1(D) \right\| \le \frac{2}{\delta} \Big(\left\| f(z) - f(z_1, \dots, z_J, 0, \dots,) \big| H^{-1}(D) \right\|$$
$$+ M(\delta) \left\| a(\cdot, z) - a_J(\cdot, z_1, \dots, z_J) \big| L_\infty(D) \right\| \Big). \tag{4.9}$$

We finally conclude from (4.8) and (4.9) together with the assumption (4.6)

$$\sup_{z \in \mathcal{U}_\rho} \left\| u(z) - u_J(z_1, \dots, z_J) \big| H_0^1(D) \right\| \longrightarrow 0 \quad (J \longrightarrow \infty).$$

In particular, we have

$$\sup_{z \in \mathcal{U}_\rho} \left\| u(z) - u_J(z_1, \dots, z_J) \big| H_0^1(D) \right\| \le \frac{1}{2N}$$

for every $J \ge J_0(N)$ and $N \in \mathbb{N}$.

Substep 2.2: Convergence for finitely many parameters.
As mentioned before, the mapping $z_{(J)} \mapsto u(z_{(J)})$ is holomorphic in an open neighbourhood of the polydisc $\mathcal{U}_{\rho,J} = \bigotimes_{1 \le j \le J} \{\zeta \in \mathbb{C} : |\zeta| \le \rho_j\}$, see the proof of Proposition 4.1. By standard results on Banach space-valued holomorphic functions (cf. [4, Proposition 3.5] or [6, Theorem 2.1.2]) this implies its analyticity and therefore the uniform convergence of its Taylor series on this polydisc. In particular, for the index sets

$$\Lambda_K^* = \left\{ \mu \in \mathbb{N}_0^J : |\mu| \le K \right\}$$

we have

$$\sup_{z = (z_1, \dots, z_J) \in \mathcal{U}_{\rho,J}} \left\| u_J(z) - S_{\Lambda_K^*} u_J(z) \big| H_0^1(D) \right\| \le \frac{1}{2N}$$

for all $K \ge K_0(N)$. W.l.o.g. we choose $K_0(N) \ge N$ and $J_0(N) \ge N$ to be increasing. Moreover, the index set $\Lambda_K^* \subset \mathbb{N}_0^J$ can be identified with an index set $\overline{\Lambda}_K^* \subset \mathcal{F}$ in an obvious way (the contained indices being supported in $\{1, \dots, J\}$). In the same spirit we identify the partial sums $S_{\Lambda_K^*} u_J(z_1, \dots, z_J)$ and $S_{\overline{\Lambda}_K^*} u(z_{(J)}) \equiv S_{\overline{\Lambda}_K^*} u(z)$ for $z \in \mathcal{U}_\rho$. Finally, putting $\Lambda_N^* = \Lambda_{K_0(N)} \subset \mathbb{N}_0^{J_0(N)}$, we end up with

$$\sup_{z \in \mathcal{U}_\rho} \left\| u(z) - S_{\overline{\Lambda}_N^*} u(z) \big| H_0^1(D) \right\| \le \frac{1}{N},$$

which proves the (uniform) convergence of the Taylor series for this particular order of summation. Clearly the sequence of sets $(\overline{\Lambda}_N^*)_{N \in \mathbb{N}}$ exhausts $\mathcal{F}$.

Substep 3: Unconditional summability.
Now let an arbitrary exhausting sequence $(\Lambda_N)_{N \in \mathbb{N}}$ of finite subsets of $\mathcal{F}$ be given. Then for any fixed $\varepsilon > 0$ we can find some $M = M(\varepsilon)$, such that

$$\sup_{z \in \mathcal{U}_\rho} \left\| u(z) - S_{\overline{\Lambda}_M^*} u(z) \big| H_0^1(D) \right\| \le \frac{\varepsilon}{2}.$$

21

Moreover, we may assume

$$\sum_{\nu \notin \overline{\Lambda}_M^*} \rho^\nu \|t_\nu | H_0^1(D)\| \le \frac{\varepsilon}{2}$$

since $(\overline{\Lambda}_N^*)_{N \in \mathbb{N}}$ exhausts $\mathcal{F}$. Because $(\Lambda_N)_{N \in \mathbb{N}}$ exhausts $\mathcal{F}$ there exists some $N_0(\varepsilon)$ such that $\overline{\Lambda}_M^* \subset \Lambda_N$ for all $N \ge N_0$. This finally implies

$$\begin{aligned}
\sup_{z \in \mathcal{U}_\rho} \big\|u(z) - S_{\Lambda_N} u(z) \big| H_0^1(D)\big\| &\le \sup_{z \in \mathcal{U}_\rho} \big\|u(z) - S_{\overline{\Lambda}_M^*} u(z) \big| H_0^1(D)\big\| + \sup_{z \in \mathcal{U}_\rho} \big\|S_{\Lambda_N \setminus \overline{\Lambda}_M^*} u(z) \big| H_0^1(D)\big\| \\
&\le \frac{\varepsilon}{2} + \sup_{z \in \mathcal{U}_\rho} \sum_{\nu \in \Lambda_N \setminus \overline{\Lambda}_M^*} |z^\nu| \cdot \|t_\nu | H_0^1(D)\| \\
&\le \frac{\varepsilon}{2} + \sum_{\nu \notin \overline{\Lambda}_M^*} \rho^\nu \|t_\nu | H_0^1(D)\| \le \varepsilon \, .
\end{aligned}$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

5 Series expansions in tensorized polynomial bases

5.1 Legendre Series

In this section we shall consider expansions of the mapping $z \mapsto u(z)$ into series of tensorized Legendre polynomials. These expansions and related N-term approximations are better suited than Taylor polynomials, if the error is measured in the least squares sense, i.e. it results in better decay estimates.

When talking about Legendre polynomials there are different versions according to the chosen normalization. In the univariate case, we define the system $(P_n)_{n \ge 0}$ to be Legendre polynomials with the L_∞-normalization $\|P_n | L_\infty([-1,1])\| = P_n(1) = 1$, and we denote by $L_n = \sqrt{2n+1} P_n$ their L_2-normalized version, i.e.

$$\int_{-1}^1 |L_n(t)|^2 \frac{dt}{2} = 1 \, .$$

We further put $P_0 = L_0 = 1$. For $\nu \in \mathcal{F}$, we define

$$P_\nu(z) = \prod_{j \ge 1} P_{\nu_j}(z_j) \qquad \text{and} \qquad L_\nu(z) = \prod_{j \ge 1} L_{\nu_j}(z_j) \, .$$

We note that the choice $P_0 = L_0 = 1$ renders $P_\nu(z)$ well-defined, since due to the finite support of ν, only finitely many factors are different from 1.

As a direct consequence we note that $(L_\nu)_{\nu \in \mathcal{F}}$ is an orthonormal basis in $L_2(U, d\mu)$, where $d\mu$ is the countable product of the probability measures $\frac{dy_j}{2}$ on $[-1,1]$. The spaces $L_p(U, d\mu)$, $0 < p \le \infty$, as well as the vector-valued spaces $L_p(U, d\mu; X)$ are to be understood similarly.

From condition (3.4) and the a priori estimate (3.5) we immediately conclude $u \in L_\infty(U, d\mu; H_0^1(D)) \hookrightarrow L_2(U, d\mu; H_0^1(D))$, thus we obtain the unique expansions

$$u(y) = \sum_{\nu \in \mathcal{F}} u_\nu P_\nu(y) = \sum_{\nu \in \mathcal{F}} v_\nu L_\nu(y)$$

with convergence in $L_2(U, d\mu; H_0^1(D))$, where the $H_0^1(D)$-valued coefficients u_ν and v_ν are given by

$$v_\nu = \int_U u(y) L_\nu(y) d\mu(y) \qquad \text{and} \qquad u_\nu = \left(\prod_{j \geq 1} (1 + 2\nu_j) \right)^{1/2} v_\nu \,.$$

We then find the following analog of Theorem 4.1 for tensorized Legendre expansions.

Theorem 5.1. Let a fulfill Assumption 1 for some $0 < r \leq R < \infty$, and suppose f satisfies condition (3.10) and Assumption 3. Moreover, let $\rho = (\rho_j)_{j \in \mathbb{N}}$ be a δ-admissible sequence with $\rho_j \geq 1$, $j \in \mathbb{N}$, and assume

$$\left(\|\psi_j | L_\infty(D)\| \right)_{j \in \mathbb{N}} \in \ell_1(\mathbb{N}) \qquad \text{and} \qquad \left(\rho_j \|\psi_j | L_\infty(D)\| \right)_{j \in \mathbb{N}} \in \ell_p(\mathbb{N}) \quad \text{for some } p < 1. \qquad (5.1)$$

Then the Legendre coefficients u_ν and v_ν of the solution u of (2.1) satisfy $\left(\rho^\nu \|u_\nu | H_0^1(D)\| \right)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$ and $\left(\rho^\nu \|v_\nu | H_0^1(D)\| \right)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$. If additionally

$$\sup_{z \in \mathcal{U}} \|f(z) - f(z_1, \ldots, z_J, 0, \ldots,) | H^{-1}(D)\| \longrightarrow 0 \quad (J \longrightarrow \infty) \,, \qquad (5.2)$$

then it holds

$$u(z) = \sum_{\nu \in \mathcal{F}} u_\nu P_\nu(z) = \sum_{\nu \in \mathcal{F}} v_\nu L_\nu(z) \,, \qquad z \in \mathcal{U} \,,$$

with uniform and unconditional convergence, which has to be understood in the following sense: If $(\Lambda_N)_{N \geq 1}$ is a sequence of subsets of $\mathcal{F}$ which exhausts $\mathcal{F}$, then the partial sums $S_{\Lambda_N} u(z) = \sum_{\nu \in \Lambda_N} u_\nu P_\nu(z) = \sum_{\nu \in \Lambda_N} v_\nu L_\nu(z)$ satisfy

$$\lim_{N \to \infty} \sup_{z \in \mathcal{U}} \|u(z) - S_{\Lambda_N} u(z) | H_0^1(D)\| = 0 \,.$$

As for the Taylor series, we start with a preliminary estimate.

Proposition 5.1. Let the coefficient a satisfy Assumption 1. For any $\nu \in \mathcal{F}$ let ρ denote a ν dependent δ-admissible sequence which is chosen in the following fashion: $\rho_j \geq 1$ for all $j \in \mathbb{N}$ and $\rho_j > 1$ for all $j \in \operatorname{supp} \nu$. Moreover, suppose f fulfills condition (3.10) and Assumption 3. Then we have the estimate

$$\|u_\nu | H_0^1(D)\| \leq M(\delta) \prod_{j \in \operatorname{supp} \nu} \frac{\pi \rho_j}{2(\rho_j - 1)} (2\nu_j + 1) \rho_j^{-\nu_j} \,.$$

The proof is exactly the same as the corresponding one in [2, Lemma 4.2], based on the a priori estimate (3.11). Equipped with this estimate, the proof of Theorem 5.1 follows by exactly the same argument as in [2], Sections 4.3 and 4.4. Specifically, this means a modification of the argument for Taylor coefficients and Taylor series (i.e. construction of suitable $\delta/2$-admissible sequences $\rho(\nu)$, reformulation of the estimate resulting from Proposition 5.1, and finally applying Proposition 4.2). We omit the details.

5.2 Chebyshev Series

As another example of orthogonal polynomials defined on the interval $[-1, 1]$ we consider the Chebyshev polynomials $T_n(t) = \cos(n \arccos(t))$, $n \in \mathbb{N}$, and $T_0(t) = 1$. They satisfy

$$\|T_n | L_\infty([-1, 1])\| = 1 \,, \quad n \geq 0 \,, \qquad \text{and} \qquad \int_{[-1,1]} |T_n(t)|^2 \frac{dt}{\sqrt{1 - t^2}} = \frac{\pi}{2} \,, \quad n \in \mathbb{N} \,.$$

23

As before we consider the set U equipped with the Borel σ-algebra $\mathcal{B}(U)$ and the probability measure

$$d\eta = \bigotimes_{j \in \mathbb{N}} \frac{dt}{\pi\sqrt{1-t^2}} \,.$$

As in the Legendre-case condition (3.4) and the a priori estimate (3.5) imply $u \in L_\infty(U, d\eta; H_0^1(D)) \hookrightarrow L_2(U, d\eta; H_0^1(D))$. Moreover, the system of tensorized polynomials $(T_\nu)_{\nu \in \mathcal{F}}$, where for $\nu \in \mathcal{F}$ we put $T_\nu(y) = \prod_{j \in \mathbb{N}} T_{\nu_j}(y_j)$, constitutes an orthogonal basis. Note that they are not orthonormal with respect to the measure η, but it holds

$$\int_U |T_\nu(y)|^2 d\eta(y) = \prod_{j \in \operatorname{supp}\nu} \frac{1}{2} = 2^{-|\operatorname{supp}\nu|} \,.$$

Nevertheless, this yields the unique expansions

$$u(y) = \sum_{\nu \in \mathcal{F}} w_\nu T_\nu(y) \,, \quad y \in U \,, \qquad w_\nu = 2^{|\operatorname{supp}\nu|} \int_U u(y) T_\nu(y) d\eta(y) \,,$$

with convergence in $L_2(U, d\eta; H_0^1(D))$.

We again aim at a summability result as in Theorems 4.1 and 5.1, and once more we start with an estimate in terms of arbitrary δ-admissible sequences.

Proposition 5.2. Under the assumptions of Proposition 5.1 it holds the estimate

$$\|w_\nu | H_0^1(D)\| \leq M(\delta) \prod_{j \in \operatorname{supp}\nu} 2\rho_j^{-\nu_j} \equiv M(\delta) 2^{|\operatorname{supp}\nu|} \rho^{-\nu} \,.$$

Proof. We will only discuss the case $\nu = ne^1$, $n \in \mathbb{N}$, the general case then is a straightforward modification by applying the one-variable case to every variable z_j with $j \in \operatorname{supp}\nu$.

Similar to the proof of Proposition 4.1 we write $z = (z_1, z') \in \mathcal{A}_\delta$ and put $u_1(z_1) = u(z_1, z')$ for some arbitrary $z' \in \mathcal{U}' = \prod_{j \geq 2}\{\zeta \in \mathbb{C} : |\zeta| \leq 1\}$ (note that due to the assumption on ρ we have $\mathcal{U} \subset \mathcal{U}_\rho$). Then we have $\|u_1(z_1)|H_0^1(D)\| \leq M(\delta)$, and we further find

$$\int_U T_\nu(y) u(y)\, d\eta(y) = \int_{U'} \int_{[-1,1]} T_n(t) u(t, y') \frac{dt}{\pi\sqrt{1-t^2}} d\eta(y') \equiv \int_{U'} \mathcal{I}_n(y')\, d\eta(y') \,.$$

It clearly suffices to bound $\|\mathcal{I}_n(y')|H_0^1(D)\|$ independently of y'. At first we find

$$\pi \mathcal{I}_n(y') = \int_0^\pi \cos(n\theta) u_1(\cos\theta) d\theta = \frac{1}{2} \int_{-\pi}^\pi u_1(\cos\theta) \cos(n\theta) d\theta$$

$$= \frac{1}{2i} \int_{|\zeta|=1} u_1\left(\frac{\zeta + \zeta^{-1}}{2}\right) \left(\frac{\zeta^n + \zeta^{-n}}{2}\right) \frac{d\zeta}{\zeta} \equiv \frac{1}{2i} \int_{|\zeta|=1} u_1\big(\mathcal{J}(\zeta)\big)\big(\mathcal{J}(\zeta^n)\big) \frac{d\zeta}{\zeta} \,.$$

The last step is verified by substituting the standard parametrization $\zeta(\theta) = e^{i\theta}$ and by the Joukowsky-transform $\mathcal{J}(\zeta) = (\zeta + \zeta^{-1})/2$. Vice versa it is well-known that $\mathcal{J}$ maps the unit circle onto the interval $[-1, 1]$ (traversed twice), since with $|\zeta| = 1$ it follows $\mathcal{J}(\zeta) = \Re(\zeta)$. Then we can estimate

$$4\pi\|\mathcal{I}_n(y')|H_0^1(D)\| \leq \left\|\int_{|\zeta|=1} u_1\big(\mathcal{J}(\zeta)\big)\zeta^n \frac{d\zeta}{\zeta}\Big|H_0^1(D)\right\| + \left\|\int_{|\zeta|=1} u_1\big(\mathcal{J}(\zeta)\big)\zeta^{-n} \frac{d\zeta}{\zeta}\Big|H_0^1(D)\right\|$$

24

$$= \left\| \int_{|\zeta|=\rho_1} u_1\big(\mathcal{J}(\zeta)\big)\zeta^{n-1}d\zeta \Big| H_0^1(D) \right\| + \left\| \int_{|\zeta|=\rho_1^{-1}} u_1\big(\mathcal{J}(\zeta)\big)\zeta^{-n-1}d\zeta \Big| H_0^1(D) \right\|$$

$$\leq M(\delta)\rho_1^{n-1} \cdot 2\pi\rho_1 + M(\delta)\rho_1^{-n-1} \cdot 2\pi\rho_1^{-1} = 4\pi M(\delta)\rho_1^{-n}.$$

Here we used Cauchy's Theorem, since $|\mathcal{J}(\zeta)| \leq \rho_1$ for $|\zeta| = \rho_1$ or $|\zeta| = \rho_1^{-1}$, and u_1 is analytic in an open neighbourhood of the disc $\{\zeta \in \mathbb{C} : |\zeta| \leq \rho_1\}$, see the proof of Proposition 4.1. $\qquad\square$

The above classical argument can essentially be found in Section 3 of [11].

By a modification of the construction and summation argument from Section 4 similar to the Legendre-case, we then obtain an analogous summability result for coefficients in the Chebyshev expansion.

Theorem 5.2. Theorem 5.1 remains true upon replacing the Legendre polynomials L_ν by the Chebyshev polynomials T_ν as well as the Legendre coefficients v_ν by the Chebyshev coefficients w_ν.

6 Higher spatial regularity

In connection with numerical approximations of the problem (2.6) in the domain D, it is of interest to establish sufficient conditions on the data which imply additional regularity of the parametric solutions in D, *while maintainig the summability of their series expansions.* In this section, we assume $f \in L_2(D) \hookrightarrow H^{-1}(D)$. Then we have

$$\|f|H^{-1}(D)\| \leq C_P \|f|L_2(D)\|,$$

where C_P denotes the Poincaré constant of D. Here, the smoothness space $W(D)$ is defined as space of all solutions to the Dirichlet problem

$$-\Delta u = f \quad \text{in } D, \qquad u|_{\partial D} = 0,$$

with $f \in L_2(D)$, i.e.

$$W(D) = \left\{ u \in H_0^1(D) : \Delta u \in L_2(D) \right\}.$$

The space $W(D)$ can be semi-normed and normed, respectively, by

$$|v|_W = \|\Delta v|L_2(D)\|, \qquad \|v|W(D)\| = \|v|H_0^1(D)\| + |v|_W.$$

We remark that for polygonal resp. polyhedral domains D, this space is contained in certain weighted $H^2(D)$ spaces, with weight functions which vanish at corners resp. corners and vertices of D (see, e.g., [5, 9] and the references there).

6.1 W-analyticity of u

Before we return to our parametric problem we shall first prove an a priori estimate and a stability assertion for the problem (2.1), now with respect to the W-norm. If we additionally assume $a \in W_\infty^1(D)$, i.e. $\nabla a \in L_\infty(D)$, then we may reformulate the problem (2.1) as

$$-\Delta u = \frac{1}{a}\big(f + \nabla a \cdot \nabla u - G(u)\big). \tag{6.1}$$

Assuming as before the validity of the ellipticity condition (2.5) we need to assure $G(u) \in L_2(D)$ in order to conclude $\Delta u \in L_2(D)$. Restricting ourselves once more to functions $G(t) = t^m$, this further restricts the choices of possible pairs (n, m). More precisely, we define

$$\mathcal{M}' = \left\{ (n, m) \in \mathcal{M} : H_0^1(D) \hookrightarrow L_{2m}(D) \right\}.$$

Under these conditions we can conclude: If $u \in H_0^1(D)$ solves the problem (2.2), then we already have $\Delta u \in L_2(D)$, and from (6.1) we obtain the estimate

$$|u|_W = \|\Delta u|L_2(D)\| \leq \frac{1}{r} \left(\|f|L_2(D)\| + \|\nabla a|L_\infty(D)\| \cdot \|u|H_0^1(D)\| + \|G(u)|L_2(D)\| \right)$$

$$\leq \frac{1}{r} \left(\|f|L_2(D)\| + \|\nabla a|L_\infty(D)\| \cdot \frac{2\|f|H^{-1}(D)\|}{r} + \|u|L_{2m}(D)\| \right), \qquad (6.2)$$

and hence

$$\|u|W(D)\| \leq \frac{1}{r} \left(1 + C_P \left(2 + \frac{2\|\nabla a|L_\infty(D)\|}{r} + \frac{2\iota(2m)}{r} \right) \right) \|f|L_2(D)\|. \qquad (6.3)$$

We continue with the announced stability assertion.

Lemma 6.1. Let u and $\widetilde{u}$ be solutions of (2.2) with $f, \widetilde{f} \in L_2(D)$ and coefficients $a, \widetilde{a} \in W_\infty^1(D)$, satisfying condition (2.5) for the same constants r and R. Then there holds the estimate

$$|u - \widetilde{u}|_W \leq \frac{1}{r} \Bigg(\|f - \widetilde{f}|L_2(D)\| + \|a - \widetilde{a}|L_\infty(D)\| \cdot |u|_W$$

$$+ \|\nabla(a - \widetilde{a})|L_\infty(D)\| \cdot \|u|H_0^1(D)\| + \|\nabla\widetilde{a}|L_\infty(D)\| \cdot \|u - \widetilde{u}|H_0^1(D)\|$$

$$+ \iota(2m)^{2m} C_P^{2(m-1)} \|u - \widetilde{u}|H_0^1(D)\|^2 \left(\frac{2\max\left(\|f|L_2(D)\|, \|\widetilde{f}|L_2(D)\|\right)}{r} \right)^{2(m-1)} \Bigg).$$

Proof. We already know from the above considerations $u, \widetilde{u} \in W(D)$. Moreover, from

$$-a\Delta u = f + \nabla a \cdot \nabla u - G(u) \qquad \text{and} \qquad -\widetilde{a}\Delta\widetilde{u} = \widetilde{f} + \nabla\widetilde{a} \cdot \nabla\widetilde{u} - G(\widetilde{u})$$

we conclude

$$\widetilde{a}\Delta(\widetilde{u} - u) = (a - \widetilde{a})\Delta u + f - \widetilde{f} + \nabla(a - \widetilde{a}) \cdot \nabla u + \nabla\widetilde{a} \cdot \nabla(u - \widetilde{u}) + G(\widetilde{u}) - G(u).$$

Taking into account condition (2.5) for $\widetilde{a}$, the claim now follows by applying the $L_2(D)$-norm. However, we need to have a closer look on the term $G(\widetilde{u}) - G(u)$. Applying Hölder's inequality we obtain

$$\|G(u) - G(\widetilde{u})|L_2(D)\| = \int_D |u(x)^m - \widetilde{u}(x)^m|^2 dx$$

$$= \int_D |u(x)^m - \widetilde{u}(x)^m| \cdot |u(x) - \widetilde{u}(x)| \cdot \left| \sum_{j=0}^{m-1} u(x)^{m-1-j}\widetilde{u}(x)^j \right| dx$$

$$\leq \|u - \widetilde{u}|L_{2m}(D)\| \cdot \|u^m - \widetilde{u}^m|L_{2m}(D)\| \sum_{j=0}^{m-1} \|u|L_{2m}(D)\|^{m-1-j}\|\widetilde{u}|L_{2m}(D)\|^j$$

$$\leq \|u-\widetilde{u}|L_{2m}(D)\|^2 \left(\sum_{j=0}^{m-1} \|u|L_{2m}(D)\|^{m-1-j}\|\widetilde{u}|L_{2m}(D)\|^j\right)^2$$

$$\leq \iota(2m)^{2m}\|u-\widetilde{u}|H_0^1(D)\|^2 \left(\sum_{j=0}^{m-1} \|u|H_0^1(D)\|^{m-1-j}\|\widetilde{u}|H_0^1(D)\|^j\right)^2$$

$$\leq \iota(2m)^{2m}C_P^{2(m-1)}\|u-\widetilde{u}|H_0^1(D)\|^2 \left(\frac{2\max\bigl(\|f|L_2(D)\|,\|\widetilde{f}|L_2(D)\|\bigr)}{r}\right)^{2(m-1)}.$$

This finally completes the proof. $\qquad\square$

We now return to the parametric problem. In order to derive a counterpart of Theorem 3.3, we shall henceforth assume $\psi_j \in W_\infty^1(D)$ for all j, and we define domains

$$\mathcal{A}_{\delta,B} := \bigl\{z \in \mathbb{C}^{\mathbb{N}} : 0 < \delta \leq \Re(a(x,z)) \leq |a(x,z)| \leq 2R \text{ and } |\nabla a(x,z)| \leq B \text{ for a.e. } x \in D\bigr\} \subset \mathcal{A}_\delta.$$

Clearly, choosing B sufficiently large implies that $\mathcal{A}_{\delta,B}$ is nonempty. Moreover, the estimate (6.3) implies the uniform bound

$$\|u(z)|W(D)\| \leq M(\delta,B) = \frac{1}{\delta}\left(1 + 2C_P\Bigl(1 + \frac{B}{\delta} + \frac{\iota(2m)}{\delta}\Bigr)\right) \sup_{z\in\mathcal{A}_{\delta,B}} \|f(z)|L_2(D)\|. \qquad (6.4)$$

Finally, we put $\widetilde{\mathcal{A}} = \bigcup_{\delta>0,B>0} \mathcal{A}_{\delta,B}$.

Proposition 6.1. Let $(n,m) \in \mathcal{M}'$, and consider $G(\zeta) = \zeta^m$ with $m \geq 2$. Then it holds

$$\lim_{h\to 0} \frac{\bigl\|G(u(z+he^j)) - G(u(z)) - G'(u(z))(u(z+he^j)-u(z))\bigl|L_2(D)\bigr\|}{|h|} = 0, \quad z \in \widetilde{\mathcal{A}}.$$

Proof. Fix $z \in \mathcal{A}_{\delta,B}$. Then similar to the proof of Proposition 6.1 we find $z + he^j \in \mathcal{A}_{\delta/2,2B}$ for $|h|$ sufficiently small, i.e. $u(z+he^j) \in H_0^1(D)$ is well-defined. As in the proof of Proposition 3.3 we apply Taylor's Theorem to obtain

$$\left\|G\bigl(u(z+he^j)\bigr) - G\bigl(u(z)\bigr) - G'\bigl(u(z)\bigr)\bigl(u(z+he^j)-u(z)\bigr)\Bigl|L_2(D)\right\|^2$$

$$= \int_D \bigl|u(z+he^j)-u(z)\bigr|^4 \left|\int_0^1 (1-t)G''\bigl(u(z)+t(u(z+he^j)-u(z))\bigr)dt\right|^2 dx$$

$$\leq \|u(z+he^j)-u(z)|L_{2m}(D)\|^4 \cdot \left\|\int_0^1 (1-t)G''\bigl(u(z)+t(u(z+he^j)-u(y))\bigr)dt\Bigl|L_{\frac{2m}{m-2}}(D)\right\|^2,$$

where in the last line we used Hölder's inequality (again the case $m = 2$ is trivial). As before the last factor can be further estimated using the (generalized) Minkowski inequality to find

$$\left\|\int_0^1 (1-t)G''\bigl(u(z)+t(u(z+he^j)-u(z))\bigr)dt\Bigl|L_{\frac{2m}{m-2}}(D)\right\|$$

$$\leq \int_0^1 (1-t)m(m-1)\left\||u(z)+t(u(z+he^j)-u(z))|^{m-2}\Bigl|L_{\frac{2m}{m-2}}(D)\right\|dt$$

27

$$\leq \frac{m(m-1)}{2}\big(\iota(2m)M(\delta,B)\big)^{m-2}\,.$$

Combining these two estimates finally yields

$$\big\|G(u(z+he^j)) - G(u(z)) - G'(u(z))(u(z+he^j)-u(z))\big|L_2(D)\big\|$$
$$\leq \tfrac{1}{2}m(m-1)\iota(2m)^m\|u(z+he^j)-u(z)|H_0^1(D)\|^2 \cdot M(\delta,B)^{m-2} = \mathcal{O}(|h|^2)\,,$$

cf. either Proposition 3.2 or Theorem 3.3. This proves the claim. $\qquad\square$

As before, we still need an assumption on the regularity of $z \mapsto f(z)$.

Assumption 4. We assume the mapping $z \mapsto f(z)$ to admit partial derivatives $\partial_{z_j} f(z) \in L_2(D)$ at every point $z \in \widetilde{\mathcal{A}}$ with respect to every variable z_j, $j \in \mathbb{N}$.

Theorem 6.1. For every $z \in \widetilde{\mathcal{A}}$ the mapping $z \mapsto u(z)$ admits complex partial derivatives $\partial_{z_j} u(z) \in W(D)$ with respect to each variable z_j.

Proof. From the proof of Theorem 3.3 we know that the difference quotient $w_h(z)$ is the unique weak solution of

$$-\nabla\big(a(\cdot,z)\nabla w\big) + G'\big(u(z)\big)w = \partial_{z_j} f(z) + L_h\,,$$

where

$$L_h = \frac{f(z+he^j) - f(z) - h\partial_{z_j} f(z)}{h} + \nabla\big(\psi_j \nabla u(z+he^j)\big)$$
$$- \frac{G(u(z+he^j)) - G(u(z)) - G'(u(z))(u(z+he^j)-u(z))}{h}\,.$$

The assumptions $\nabla\psi_j \in L_\infty(D)$ and $(n,m) \in \mathcal{M}'$ then further yield

$$L_h = \frac{f(z+he^j) - f(z) - h\partial_{z_j} f(z)}{h} + \nabla\psi_j \cdot \nabla u(z+he^j) + \psi_j \Delta u(z+he^j)$$
$$- \frac{G(u(z+he^j)) - G(u(z)) - G'(u(z))(u(z+he^j)-u(z))}{h} \in L_2(D)\,.$$

As above this further implies $w_h \in W(D)$. Similarly, we know from Theorem 3.3 that $w_h(z)$ converges in $H_0^1(D)$ towards $w_0(z) = \partial_{z_j} u(z)$, which is the solution of

$$-\nabla\big(a(\cdot,z)\nabla w\big) + G'\big(u(z)\big)w = \partial_{z_j} f(z) + \nabla\psi_j \cdot \nabla u(z) + \psi_j \Delta u(z)\,.$$

Again we conclude $w_0 \in W(D)$, and we find by subtracting these two problems

$$-a(\cdot,z)\Delta(w_h - w_0) = -G'\big(u(z)\big)(w_h - w_0) + \nabla\psi_j \cdot \nabla\big(u(z+he^j)-u(z)\big)$$
$$+ \psi_j \Delta\big(u(z+he^j)-u(z)\big) + \frac{f(z+he^j) - f(z) - h\partial_{z_j} f(z)}{h}$$
$$- \frac{G(u(z+he^j)) - G(u(z)) - G'(u(z))(u(z+he^j)-u(z))}{h}\,.$$

By dividing by $a(\cdot,z)$ and applying the $L_2(D)$-norm we then obtain

$$|w_h - w_0|_W \longrightarrow 0 \quad (h \longrightarrow 0)\,,$$

since all the terms on the right hand side converge to 0 separately (this follows from Theorem 3.3, Proposition 3.2, Lemma 6.1, Assumption 4, and Proposition 6.1, in this order). This completes the proof. $\qquad\square$

6.2 Summability of Taylor, Legendre and Chebyshev coefficient sequences

We begin with the counterpart of Propositions 4.1, 5.1 and 5.2.

Lemma 6.2. Under the assumptions of Theorem 6.1, we have the estimate

$$\|t_\nu|W(D)\| \leq M(\delta, B)\rho^{-\nu}$$

for every $\nu \in \mathcal{F}$ and every (δ, B)-admissible sequence ρ. If additionally we have $\rho_j \geq 1$ for all $j \in \mathbb{N}$ and $\rho_j > 1$ for $j \in \text{supp}\,\nu$, we also find

$$\|u_\nu|W(D)\| \leq M(\delta, B) \prod_{j \in \text{supp}\,\nu} \frac{\pi \rho_j}{2(\rho_j - 1)} (2\nu_j + 1)\rho_j^{-\nu_j} \quad \text{and} \quad \|w_\nu|W(D)\| \leq M(\delta, B) \prod_{j \in \text{supp}\,\nu} 2\rho_j^{-\nu_j}.$$

Proof. The proof of the first part is identical to the one of Proposition 4.1, using now the regularity result from Theorem 6.1, and replacing the estimate (4.1) by its counterpart based upon (6.4). In a similar way, we obtain the second part by the same argument as in the proofs of Propositions 5.1 and 5.2. $\qquad \square$

Theorem 6.2. Let a satisfy Assumption 1 for some $0 < r \leq R < \infty$. Suppose f fulfills

$$\|f(z)|H^{-1}(D)\| < \left(\frac{\delta^m}{2^{2m}m\iota(m+1)^{m+1}} \right)^{\frac{1}{m-1}}, \quad z \in \mathcal{A}_{\delta/2,B}, \tag{6.5}$$

and Assumption 4. Moreover, let $\rho = (\rho_j)_{j \in \mathbb{N}}$ be an arbitrary (δ, B)-admissible sequence with B sufficiently large and $\rho_j \geq 1$, $j \in \mathbb{N}$, and assume

$$\left(\|\psi_j|W_\infty^1(D)\| \right)_{j \in \mathbb{N}} \in \ell_1(\mathbb{N}) \qquad \text{and} \qquad \left(\rho_j \|\psi_j|W_\infty^1(D)\| \right)_{j \in \mathbb{N}} \in \ell_p(\mathbb{N}) \quad \text{for some } p < 1. \tag{6.6}$$

Then the Taylor coefficients t_ν and the Legendre coefficients u_ν and v_ν of the solution u of (2.2) satisfy $\left(\rho^\nu \|t_\nu|W(D)\| \right)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$, $\left(\rho^\nu \|u_\nu|W(D)\| \right)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$ and $\left(\rho^\nu \|v_\nu|W(D)\| \right)_{\nu \in \mathcal{F}} \in \ell_p(\mathcal{F})$. If additionally

$$\sup_{z \in \mathcal{U}_\rho} \big\| f(z) - f(z_1, \ldots, z_J, 0, \ldots,) | L_2(D) \big\| \longrightarrow 0 \quad (J \longrightarrow \infty), \tag{6.7}$$

then it holds

$$u(z) = \sum_{\nu \in \mathcal{F}} t_\nu z^\nu, \quad z \in \mathcal{U}_\rho, \qquad \text{and} \qquad u(z) = \sum_{\nu \in \mathcal{F}} u_\nu P_\nu(z) = \sum_{\nu \in \mathcal{F}} v_\nu L_\nu(z) = \sum_{\nu \in \mathcal{F}} w_\nu T_\nu(z), \quad z \in \mathcal{U},$$

with uniform and unconditional convergence in $W(D)$, which is to be understood as in Theorems 4.1, 5.1 and 5.2.

Proof. The proof of the summability assertion is exactly the same as the one for Theorems 4.1, 5.1 and 5.2 upon replacing $\gamma_j = \|\psi_j|L_\infty(D)\|$ by $\widetilde{\gamma}_j = \|\psi_j|W_\infty^1(D)\| = \|\psi_j|L_\infty(D)\| + \|\nabla\psi_j|L_\infty(D)\|$, see also [2, Theorem 5.1]. Moreover, also the convergence proof remains the same upon now using the stability assertion for the W-semi-norm (Lemma 6.1). $\qquad \square$

7 Best N-term approximation rates

The convergence results from Theorems 4.1 and 5.1 may be re-interpreted in the following way: when approximating the mapping $y \mapsto u(y)$, i.e. approximately solving the parametric problem (2.6), instead of solving single instances of the problem (i.e. solving the equation for single fixed choices $y^{(j)} \in U$, $j = 1, \ldots, N$) and afterward interpolating these points $u(y^{(j)})$, we can try to simultaneously approximate the solution for all $y \in U$ by approximating u by appropriate partial sums of the Taylor, Legendre or Chebyshev expansions. The main task consists in bounding truncation errors of best N-term approximation by monomials y^ν or polynomials P_ν, L_ν or T_ν, respectively. Moreover, we can utilize the summability results from Theorems 4.1, 5.1 and 5.2 to estimate the rate of convergence. To do so, we note that by Assumption 1 the constant sequence $\rho_j = 1$, $j \in \mathbb{N}$, is δ-admissible for all $0 < \delta \le r$. Then

$$\sup_{y \in U} \left\| u(y) - S_\Lambda^{(T)} u(y) \big| H_0^1(D) \right\| = \sup_{y \in U} \left\| u(y) - \sum_{\nu \in \Lambda} t_\nu y^\nu \Big| H_0^1(D) \right\| \le \sum_{\nu \notin \Lambda} \| t_\nu | H_0^1(D) \| .$$

Similarly, we obtain by Parseval's identity

$$\left\| u - S_\Lambda^{(L)} u \big| L_2\big(U; d\mu; H_0^1(D)\big) \right\|^2 = \left\| u - \sum_{\nu \in \Lambda} v_\nu L_\nu(y) \Big| L_2\big(U; d\mu; H_0^1(D)\big) \right\|^2 = \sum_{\nu \notin \Lambda} \| v_\nu | H_0^1(D) \|^2 .$$

These observations are to be combined with Stechkin's Lemma, which in our case can be formulated as follows: Given a sequence $\gamma = (\gamma_\nu)_{\nu \in \mathcal{F}} \subset H_0^1(D)$, and denoting by $\Lambda_N \subset \mathcal{F}$ an index set corresponding to the N largest values of $\| \gamma_\nu | H_0^1(D) \|$, it holds for $0 < p \le q \le \infty$

$$\left(\sum_{\nu \notin \Lambda_N} \| \gamma_\nu | H_0^1(D) \|^q \right)^{1/q} \le N^{-1/p + 1/q} \left(\sum_{\nu \in \mathcal{F}} \| \gamma_\nu | H_0^1(D) \|^p \right)^{1/p}$$

and this choice of Λ_N yields a best N-term approximation of γ. We have shown

Theorem 7.1. Let a and f satisfy the assumptions of Theorems 4.1, 5.1 and 5.2. If we denote by $\Lambda_N^{(T)} \subset \mathcal{F}$ an index set corresponding to the N largest values of $\| t_\nu | H_0^1(D) \|$, then it holds

$$\sup_{y \in U} \left\| u(y) - S_{\Lambda_N^{(T)}}^{(T)} u(y) \big| H_0^1(D) \right\| \le \left\| \big(\| t_\nu | H_0^1(D) \| \big)_{\nu \in \mathcal{F}} \Big| \ell_p(\mathcal{F}) \right\| N^{-s} , \qquad s = \frac{1}{p} - 1 .$$

This remains true for the Legendre Polynomials P_ν and the Chebyshev polynomials T_ν with the respective coefficients, partial sums and index sets for largest coefficients.
Moreover, if $\Lambda_N^{(L)} \subset \mathcal{F}$ is a set corresponding to the N largest values of $\| v_\nu | H_0^1(D) \|$, then it holds

$$\left\| u - S_{\Lambda_N^{(L)}}^{(L)} u \big| L_2\big(U, d\mu; H_0^1(D)\big) \right\| \le \left\| \big(\| v_\nu | H_0^1(D) \| \big)_{\nu \in \mathcal{F}} \Big| \ell_p(\mathcal{F}) \right\| N^{-s} , \qquad s = \frac{1}{p} - \frac{1}{2} ,$$

and once more this remains true for Chebyshev polynomials T_ν with respect to the measure η.
Finally, if a and f satisfy the assumptions of Theorem 6.2, then all the estimates remain true upon replacing the $H_0^1(D)$-norm by the $W(D)$-norm.

If it comes to an actual computation of these approximating partial sums, one has to take into consideration that also the coefficients itself must be approximated, e.g. for Taylor coefficients by approximately solving linear problems as in Theorem 3.3.

Here the increased spatial regularity from Section 6 comes into play. These results enable us to solve the corresponding linear problems for the Taylor coefficients using Finite element techniques. If we denote by t the rate of convergence for the Finite element method (depending on the method itself as well as the shape of the domain D), then the error of computing $t_\nu \in W(D)$, measured in the $H_0^1(D)$-norm, can be estimated by CM_ν^{-t}, where M_ν is the number of degrees of freedom used in the computation of t_ν. The total number of degrees of freedom required for approximating the partial sum $\widetilde{S}_\Lambda u$ is then given by

$$N_{\text{dof}} = \sum_{\nu \in \Lambda} M_\nu \,.$$

With this notation the approximation error in both, the parametric space as well as in the physical domain D, satisfies the estimate

$$\sup_{y \in U} \big\| u(y) - \widetilde{S}_{\Lambda_N^{(T)}}^{(T)} u(y) \big| H_0^1(D) \big\| \leq C \, N_{\text{dof}}^{-\min\{s,t\}} \,,$$

where the occurring constants can be estimated in dependence of $\big\| (\|t_\nu|H_0^1(D)\|)_{\nu \in \mathcal{F}} \big| \ell_p(\mathcal{F}) \big\|$ and $\big\| (\|t_\nu|W(D)\|)_{\nu \in \mathcal{F}} \big| \ell_p(\mathcal{F}) \big\|$. Similar considerations are true for Legendre or Chebyshev partial sums (see [2], Section 5).

References

[1] A. COHEN, R. A. DEVORE AND CH. SCHWAB, Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. Journ. Found. Comp. Math. Volume **10**, Number 6, December 2010, pp. 615-646.

[2] A. COHEN, R. A. DEVORE AND CH. SCHWAB, Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. Analysis and Applications **9**, 1-37 (2011).

[3] G. DA PRATO, *An introduction to infinite-dimensional analysis*. Universitext, Springer-Verlag, Berlin, 2006. Revised and extended from the 2001 original by Da Prato.

[4] S. DINEEN, *Complex Analysis on Infinite Dimensional Spaces*, Springer Monographs in Mathematics, Springer Verlag, Berlin, 1999.

[5] P. GRISVARD, *Elliptic Problems on Nonsmooth Domains* Pitman, Boston 1985.

[6] M. HERVÉ, *Analyticity in infinite dimensional spaces.* De Gruyter, Berlin, 1989.

[7] V.H. HOANG AND CH. SCHWAB, Sparse tensor Galerkin discretizations for parametric and random parabolic PDEs. I: Analytic regularity and gpc-approximation Report 2010-11, Seminar for Applied Mathematics, ETH Zürich (in review).

[8] V.H. HOANG AND CH. SCHWAB, Analytic regularity and gpc approximation for parametric and random 2nd order hyperbolic PDEs Report 2010-19, Seminar for Applied Mathematics, ETH Zürich (to appear in Analysis and Applications (2011)).

[9] V.G. MAZ'YA AND J. ROSSMANN, Elliptic Equations in Polyhedral Domains, Mathematical Surveys and Monographs **162** (2010), AMS Publishing House.

[10] H. Rauhut and R. Ward, *Sparse Legendre expansions via ℓ_1-minimization*, preprint, Bonn, 2010.

[11] T.J. Rivlin, *The Chebyshev Polynomials.* Ser. Pure & Applied Mathematics, Wiley, New York, 1974.

[12] T. Runst, W. Sickel, *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations.* De Gruyter series in nonlinear analysis and applications, De Gruyter, Berlin, 1996.

[13] M. Struwe, *Variational methods.* A series of modern surveys in mathematics, 4th Ed., Springer, Berlin, Heidelberg, 2008.

[14] R.A. Todor, *Convergence Rates of Sparse Chaos Approximations of Elliptic Problems with stochastic coefficients*, IMA Journ. Numer. Anal. **44**(2007) 232-261.

Research Reports