Electronic Journal of Differential Equations, Vol. 2004(2004), No. 75, pp. 1–26.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu (login: ftp)

CONSTRUCTIVE SOBOLEV GRADIENT PRECONDITIONING FOR SEMILINEAR ELLIPTIC SYSTEMS

JÁNOS KARÁTSON

Abstract. We present a Sobolev gradient type preconditioning for iterative methods used in solving second order semilinear elliptic systems; the n-tuple of independent Laplacians acts as a preconditioning operator in Sobolev spaces. The theoretical iteration is done at the continuous level, providing a linearization approach that reduces the original problem to a system of linear Poisson equations. The method obtained preserves linear convergence when a polynomial growth of the lower order reaction type terms is involved. For the proof of linear convergence for systems with mixed boundary conditions, we use suitable energy spaces. We use Sobolev embedding estimates in the construction of the exact algorithm. The numerical implementation has focus on a direct and elementary realization, for which a detailed discussion and some examples are given.

1. Introduction

The numerical solution of elliptic problems is a topic of basic importance in numerical mathematics. It has been a subject of extensive investigation in the past decades, thus having vast literature (cf. [5, 16, 23, 26] and the references there). The most widespread way of finding numerical solutions is first discretizing the elliptic problem, then solving the arising system of algebraic equations by a solver which is generally some iterative method. In the case of nonlinear problems, most often Newton's method is used. However, when the work of compiling the Jacobians exceeds the advantage of quadratic convergence, one may prefer gradient type iterations including steepest descent or conjugate gradients (see e.g. [4, 9]). An important example in this respect is the Sobolev gradient technique, which represents a general approach relying on descent methods and has provided various efficient numerical results [28, 29, 30]. In the context of gradient type iterations the crucial point is most often preconditioning. Namely, the condition number of the Jacobians of the discretized systems tends to infinity when discretization is refined, hence suitable nonlinear preconditioning technique has to be used to achieve a convenient condition number [2, 3]. The Sobolev gradient technique presents a

2000 *Mathematics Subject Classification.* 35J65, 49M10.
Key words and phrases. Sobolev gradient, semilinear elliptic systems, numerical solution, preconditioning.
©2004 Texas State University - San Marcos.
Submitted March 18, 2004. Published May 21, 2004.
Partially supported by the Hungarian National Research Fund OTKA.

general efficient preconditioning approach where the preconditioners are derived from the representation of the Sobolev inner product.

The Sobolev gradient idea does in fact opposite to that which first discretizes the problem. Namely, the iteration may be theoretically defined in Sobolev spaces for the boundary-value problem (i.e. at the continuous level), reducing the nonlinear problem to auxiliary linear Poisson equations. Then discretization may be used for these auxiliary problems. This approach is based on the various infinite-dimensional generalizations of iterative methods, beginning with Kantorovich. For recent and earlier results see [15, 22, 29, 30]. The author's investigations include the development of the preconditioning operator idea as shown in [12]. Some recent numerical results are given in [20, 21] which are closely related to the Sobolev gradient idea. Namely, a suitable representation of the gradient yields a preconditioning elliptic operator; for Dirichlet problems the usual Sobolev inner product leads to the Laplacian as preconditioner. For systems one may define independent Laplacians as preconditioners, see [17] for an earlier treatment for uniformly elliptic Dirichlet problems using the strong form of the operators. We note that the constructive representation of the Sobolev gradients with Laplacians in [17, 21] is due to a suitable regularity property.

In this context the Sobolev gradient can be regarded as infinite dimensional preconditioning by the Laplacian. It yields that the speed of linear convergence is determined by the ellipticity properties of the original problem instead of the discretized equation, i.e., the ratio of convergence is explicitly obtained from the coefficients of the boundary-value problem. Therefore, it is independent of the numerical method used for the auxiliary problems. Another favourable property is the reduction of computational issues to those arising for the linear auxiliary problems. These advantages appear in the finite element methods (FEM) realization [13]. In [13, 17, 21] Dirichlet problems are considered for uniformly elliptic equations.

The aim of this paper is to develop the above described realization of Sobolev gradients for semilinear elliptic systems, including the treatment of non-uniformity (caused by polynomial growth of the lower order reaction-type terms) such that linear convergence is preserved, and considering general mixed boundary conditions which need suitable energy spaces. The studied class of problems includes elliptic reaction-diffusion systems related to reactions of autocatalytic type.

The paper first gives a general development of the method: after a brief Hilbert space background, the theoretical iteration is constructed in Sobolev space and convergence is proved. Linear convergence is obtained in the norm of the corresponding energy space, equivalent to the Sobolev norm. An excursion to Sobolev embeddings is enclosed, which is necessary for determining the required descent parameter (and the corresponding convergence quotient). Then numerical realization is considered with focus on direct elementary realization. A detailed discussion is devoted to the latter, giving a general extension of the ideas of [19, 21]. Also numerical examples are presented.

2. General construction and convergence

2.1. **The gradient method in Hilbert space.** In this subsection the Hilbert space result of [17] on non-differentiable operators is suitably modified for our purposes.

First we quote the theorem on differentiable operators this result relies on. The classical theorem, mentioned already by Kantorovich [22], is given in the form needed for our setting, including suitable conditions and stepsize.

Theorem 2.1. *Let H be a real Hilbert space and $F : H \to H$ have the following properties:*

(i) *F is Gâteaux differentiable;*
(ii) *for any $u, k, w, h \in H$ the mapping $s, t \mapsto F'(u + sk + tw)h$ is continuous from $\mathbf{R}^2$ to H;*
(iii) *for any $u \in H$ the operator $F'(u)$ is self-adjoint;*
(iv) *there are constants $M \geq m > 0$ such that for all $u \in H$*

$$m\|h\|^2 \leq \langle F'(u)h, h\rangle \leq M\|h\|^2 \quad (h \in H).$$

Then for any $b \in H$ the equation $F(u) = b$ has a unique solution $u^* \in H$, and for any $u^0 \in H$ the sequence

$$u^{k+1} = u^k - \frac{2}{M+m}(F(u^k) - b) \quad (k \in \mathbb{N})$$

converges linearly to u^*, namely,

$$\|u^k - u^*\| \leq \frac{1}{m}\|F(u^0) - b\|\big(\frac{M-m}{M+m}\big)^k \quad (k \in \mathbb{N})\,. \tag{2.1}$$

A short proof of the theorem (cf. [27]) is based on proving the estimate

$$\|F(u^k) - b\| \leq \big(\frac{M-m}{M+m}\big)^k\|F(u^0) - b\| \quad (k \in \mathbb{N}) \tag{2.2}$$

(to which end one verifies that $J(u) \equiv u - \frac{2}{M+m}(F(u^k) - b)$ possesses contraction constant $\frac{M-m}{M+m}$). Then (2.2) yields (2.1) since assumption (iv) implies

$$m\|u - v\| \leq \|F(u) - F(v)\| \quad (u, v \in H).$$

Proposition 2.2. *Under the assumptions of Theorem 2.1 we have*

$$\|u^k - u^0\| < \frac{1}{m}\|F(u^0) - b\| \quad (k \in \mathbb{N}).$$

Proof.

$$\begin{aligned}
\|u^k - u^0\| &\leq \sum_{i=0}^{k-1}\|u_{i+1} - u_i\| \\
&= \frac{2}{M+m}\sum_{i=0}^{k-1}\|F(u^i) - b\| \\
&\leq \frac{2}{M+m}\|F(u^0) - b\|\sum_{i=0}^{k-1}\big(\frac{M-m}{M+m}\big)^i \\
&\leq \|F(u^0) - b\|\frac{2}{(M+m)(1-\frac{M-m}{M+m})} = \frac{1}{m}\|F(u^0) - b\|.
\end{aligned}$$

□

Proposition 2.2 allows localization of the ellipticity assumption (cf. [15]):

Corollary 2.3. *Let $u^0 \in H$, $r_0 = \frac{1}{m}\|F(u^0) - b\|$, $B(u^0, r_0) = \{u \in H : \|u - u^0\| \leq r_0\}$. Then in Theorem 2.1 it suffices to assume that*

(iv)' *there exist* $M \geq m > 0$ *such that for all* $u \in H$ $\langle F'(u)h, h\rangle \geq m\|h\|^2 \quad (h \in H)$, *and for all* $u \in B(u^0, r_0)$ $\langle F'(u)h, h\rangle \leq M\|h\|^2 \quad (h \in H)$,

instead of assumption (iv), in order that the theorem holds.

Proof. The lower bound in (iv)' ensures that F is uniformly monotone, which yields existence and uniqueness as before. Owing to Proposition 2.2 the upper bound M is only exploited in $B(u^0, r_0)$ to produce the convergence result. □

Turning to non-differentiable operators, we now quote the corresponding result in [17]. First the necessary notations are given.

Definition 2.4. Let $B : D \to H$ be a strictly positive symmetric linear operator. Then the energy space of B, i.e. the completion of D with respect to the scalar product

$$\langle x, y\rangle_B \equiv \langle Bx, y\rangle \quad (x, y \in D)$$

is denoted by H_B. The corresponding norm is denoted by $\|\cdot\|_B$.

For any $r \in \mathbb{N}^+$ we denote by $H^r \equiv H \times H \times \cdots \times H$ (r times) the product space. The corresponding norm is denoted by

$$[[u]] \equiv \Big(\sum_{i=1}^{r} \|u_i\|^2\Big)^{1/2} \quad (u \in H^r).$$

The obvious analogous notation is used for the products of H_B.

Theorem 2.5 ([17])**.** *Let H be a real Hilbert space, $D \subset H$. Let $T_i : D^r \to H \quad (i = 1, \dots, r)$ be non-differentiable operators. We consider the system*

$$T_i(u_1, \dots, u_r) = g_i \quad (i = 1, \dots, r) \tag{2.3}$$

with $g = (g_1, \dots, g_r) \in H^r$. Let $B : D \to H$ be a symmetric linear operator with lower bound $\lambda > 0$. Assume that the following conditions hold:

(i) $R(B) \supset R(T_i) \quad (i = 1, \dots, r)$;

(ii) *for any i=1,...,r the operators $B^{-1}T_i$ have Gâteaux differentiable extensions $F_i : H_B^r \to H_B$, respectively;*

(iii) *for any $u, k, w, h \in H_B^r$ the mappings $s, t \mapsto F_i'(u+sk+tw)h$ are continuous from $\mathbf{R}^2$ to H_B;*

(iv) *for any $u, h, k \in H_B^r$*

$$\sum_{i=1}^{r} \langle F_i'(u)h, k_i\rangle_B = \sum_{i=1}^{r} \langle h_i, F_i'(u)k\rangle_B;$$

(v) *there are constants $M \geq m > 0$ such that for all $u, h \in H_B^r$*

$$m \sum_{i=1}^{r} \|h_i\|_B^2 \leq \sum_{i=1}^{r} \langle F_i'(u)h, h_i\rangle_B \leq M \sum_{i=1}^{r} \|h_i\|_B^2 .$$

Let $g_i \in R(B) \quad (i = 1, \dots, r)$. Then

(1) *system (2.3) has a unique weak solution $u^* = (u_1^*, \dots, u_r^*) \in H_B^r$, i.e. which satisfies*

$$\langle F_i(u^*), v\rangle_B = \langle g_i, v\rangle \quad (v \in H_B,\ i = 1, \dots, r);$$

(2) *for any* $u^0 \in D^r$ *the sequence* $u^k = (u_1^k, \dots, u_r^k)_{k\in\mathbb{N}}$, *given by the coordinate sequences*

$$u_i^{k+1} \equiv u_i^k - \frac{2}{M+m} B^{-1}(T_i(u^k) - g_i) \quad (i = 1, \dots, r; k \in \mathbb{N}),$$

converges linearly to u^*. *Namely,*

$$[[u^k - u^*]]_B \le \frac{1}{m\sqrt{\lambda}} [[T(u^0) - g]] \Big(\frac{M-m}{M+m}\Big)^k \quad (k \in \mathbb{N}).$$

The proof of this theorem in [17] is done by applying Theorem 2.1 to the operator $F = (F_1, \dots, F_r)$ and the right-side $b = \{b_i\}_{i=1}^r = \{B^{-1}g_i\}_{i=1}^r$ in the space H_B^r. This implies that the assumption can be weakened in the same way as in Corollary 2.3. That is, we have

Corollary 2.6. *Let* $u^0 \in D$, $b_i = B^{-1}g_i$ $(i = 1, \dots, r)$, $r_0 = \frac{1}{m}[[F(u^0) - b]]_B$, $B(u^0, r_0) = \{u \in H_B^r : [[u - u^0]]_B \le r_0\}$. *Then in Theorem 2.2 it suffices to assume that*

(v)' *there exist* $M \ge m > 0$ *such that for all* $u \in H_B$ $\sum_{i=1}^r \langle F_i'(u)h, h_i\rangle_B \ge m[[h]]_B^2$ $(h \in H_B)$, *and for all* $u \in B(u^0, r_0)$ $\sum_{i=1}^r \langle F_i'(u)h, h_i\rangle_B \le M[[h]]_B^2$ $(h \in H_B)$,

instead of assumption (v), in order that the theorem holds.

Finally we remark that the *conjugate gradient method (CGM)* in Hilbert space is formulated in [9] for differentiable operators under fairly similar conditions as for Corollary 2.3, and this result is extended to non-differentiable operators in [18] similarly to Corollary 2.6. Compared to the gradient method, the CGM improves the above convergence ratio to $(\sqrt{M}-\sqrt{m})(\sqrt{M}+\sqrt{m})$, on the other hand, the extra work is the similar construction of two simultaneous sequences together with the calculation of required inner products and numerical root finding for the stepsize.

2.2. **The gradient method in Sobolev space.** We consider the system of boundary value problems

$$\begin{aligned} T_i(u_1, \dots, u_r) \equiv & -\operatorname{div}(a_i(x)\nabla u_i) + f_i(x, u_1, \dots, u_r) = g_i(x) \quad \text{in } \Omega \\ & Qu_i \equiv (\alpha(x)u_i + \beta(x)\partial_\nu u_i)\big|_{\partial\Omega} = 0 \end{aligned} \tag{2.4}$$

$(i = 1, \dots, r)$ on a bounded domain $\Omega \subset \mathbb{R}^N$ with the following conditions:

(C1) $\partial\Omega \in C^2$, $a_i \in C^1(\overline{\Omega})$, $f_i \in C^1(\overline{\Omega} \times \mathbb{R}^r)$, $g_i \in L^2(\Omega)$.

(C2) $\alpha, \beta \in C^1(\partial\Omega)$, $\alpha, \beta \ge 0$, $\alpha^2 + \beta^2 > 0$ almost everywhere on $\partial\Omega$.

(C3) There are constants $m, m' > 0$ such that $0 < m \le a_i(x) \le m'$ $(x \in \overline{\Omega})$, further, $\eta \equiv \sup_{\Gamma_\beta} \frac{\alpha}{\beta} < +\infty$ where

$$\Gamma_\beta \equiv \{x \in \partial\Omega : \beta(x) > 0\}.$$

(C4) Let $2 \le p \le \frac{2N}{N-2}$ (if $N > 2$), $2 \le p$ (if $N = 2$). There exist constants $\kappa' \ge \kappa \ge 0$ and $\gamma \ge 0$ such that for any $(x, \xi) \in \overline{\Omega} \times \mathbb{R}^r$ the Jacobians $\partial_\xi f(x, \xi) = \{\partial_{\xi_k} f_j(x, \xi_1, \dots, \xi_r)\}_{j,k=1}^r \in \mathbb{R}^{r\times r}$ are symmetric and their eigenvalues μ fulfil

$$\kappa \le \mu \le \kappa' + \gamma \sum_{j=1}^r |\xi_j|^{p-2}.$$

Moreover, in the case $\alpha \equiv 0$ we assume $\kappa > 0$, otherwise $\kappa = 0$.

Let

$$D_Q \equiv \{u \in H^2(\Omega) : Qu\big|_{\partial\Omega} = 0 \text{ in trace sense}\}. \tag{2.5}$$

We define

$$D(T_i) = D_Q^r$$

as the domain of T_i $(i = 1, \dots, r)$.

An essential special case of (2.4) is that with polynomial nonlinearity

$$f_i(x, u_1, \dots, u_r) = \sum_{|j| \le s_i} c^{(i)}_{j_1,\dots,j_r}(x) u_1^{j_1} \dots u_r^{j_r} \tag{2.6}$$

that fulfils condition (C4), where $s_i \in \mathbb{N}^+$, $s_i \le p-1$, $c^{(i)}_{j_1,\dots,j_r} \in C(\overline{\Omega})$ and $|j| \equiv j_1 + \dots j_r$ for $j = (j_1, \dots, j_r) \in \mathbb{N}^r$. This occurs in steady states or in time discretization of autocatalytic reaction-diffusion systems. (Then a_i and $c^{(i)}_{j_1,\dots,j_r}$ are generally constant).

(a) Energy spaces. The construction of the gradient method requires the introduction and the study of some properties of energy spaces of the Laplacian.

Definition 2.7. Let

$$Bu \equiv -\Delta u + cu,$$

defined for $u \in D(B) = D_Q$ (see (2.5)), where $c = \frac{\kappa}{m}$ (≥ 0) with m and κ from conditions (C3)-(C4) (i.e. $B = -\Delta$ except the case $\alpha \equiv 0$).

Remark 2.8. It can be seen in the usual way that B is symmetric and strictly positive in the real Hilbert space $L^2(\Omega)$.

Corollary 2.9. (a) *The eigenvalues λ_i $(i \in \mathbb{N}^+)$ of B are positive.*

(b) *We have $\int_\Omega (Bu)u \ge \lambda_1 \int_\Omega u^2$ $(u \in D_Q)$ where $\lambda_1 > 0$ is the smallest eigenvalue of B.*

Definition 2.10. Denote by $H^1_Q(\Omega)$ the *energy space* of B, i.e. $H^1_Q(\Omega) = H_B$ (cf. Definition 2.4). Due to the divergence theorem we have

$$\langle u, v\rangle_{H^1_Q} \equiv \langle u, v\rangle_B = \int_\Omega \big(\nabla u \cdot \nabla v + cuv\big)\, dx + \int_{\Gamma_\beta} \frac{\alpha}{\beta} uv\, d\sigma \quad (u, v \in D). \tag{2.7}$$

Remark 2.11. Using Corollary 2.9, we can deduce the following properties:

(a) $(1 + \lambda^{-1})\|u\|^2_{H^1_Q} \ge \|u\|_{H^1(\Omega)} \equiv \int_\Omega \big(|\nabla u|^2 + u^2\big)\, dx$ $(u \in H^1_Q(\Omega))$.

(b) $H^1_Q(\Omega) \subset H^1(\Omega)$.

(c) (2.7) holds for all $u, v \in H^1_Q(\Omega)$.

Remark 2.12. Remark 2.11 (b) implies that the Sobolev embedding theorem [1] holds for $H^1_Q(\Omega)$ in the place of $H^1(\Omega)$. Namely, for any $p \ge 2$ if $N = 2$, and for $2 \le p \le \frac{2N}{N-2}$ if $N > 2$, there exists $K_{p,\Omega} > 0$ such that

$$H^1_Q(\Omega) \subset L^p(\Omega), \quad \|u\|_{L^p(\Omega)} \le K_{p,\Omega}\|u\|_{H^1_Q} \quad (u \in H^1_Q(\Omega)). \tag{2.8}$$

Definition 2.13. The product spaces $L^2(\Omega)^r$ and $H^1_Q(\Omega)^r$ are endowed with the norms

$$\|u\|_{L^2(\Omega)^r} \equiv \big(\sum_{i=1}^r \|u_i\|^2_{L^2(\Omega)}\big)^{1/2} \quad \text{and} \quad \|u\|_{H^1_Q(\Omega)^r} \equiv \big(\sum_{i=1}^r \|u_i\|^2_{H^1_Q}\big)^{1/2},$$

respectively, where $u = (u_1, \dots, u_r)$ and $\|\cdot\|_{H^1_Q} = \|\cdot\|_{H^1_Q(\Omega)}$ for brevity as in Def. 2.3.

(b) The convergence result.

Theorem 2.14. *Under the conditions (C1)-(C4) the following results hold.*
(1) *The system (2.4) has a unique weak solution* $u^* = (u_1^*, \dots, u_r^*) \in H_Q^1(\Omega)^r$.
(2) *Let* $u_i^0 \in D_Q$ $(i = 1, \dots, r)$ *and*

$$\begin{aligned} M &= \max\{1, \eta\} m' + \kappa' \lambda_1^{-1} \\ &\quad + \gamma K_{p,\Omega}^p \mu_p \big(\|u^0\|_{H_Q^1(\Omega)^r} + m^{-1} \lambda_1^{-1/2} \|T(u^0) - g\|_{L^2(\Omega)^r} \big)^{p-2} \end{aligned} \tag{2.9}$$

(with m, m', η *from condition (C3),* p, κ', γ *from (C4),* $K_{p,\Omega}$ *from Remark 2.12,* λ_1 *from Corollary 2.9 and* $\mu_p = \max\{1, r^{(4-p)/2}\}$*).*
Let

$$u_i^{k+1} = u_i^k - \frac{2}{M+m} z_i^k \quad (k \in \mathbb{N},\, i = 1, \dots, r) \tag{2.10}$$

where

$$g_i^k = T_i(u^k) - g_i \quad (k \in \mathbb{N},\, i = 1, \dots, r) \tag{2.11}$$

and $z_i^k \in D_Q$ *is the solution of the auxiliary problem*

$$\begin{aligned} (-\Delta + c) z_i^k &= g_i^k \\ \alpha(x) z_i^k + \beta(x) \partial_\nu z_i^k \big|_{\partial\Omega} &= 0\,. \end{aligned} \tag{2.12}$$

(We solve Poisson equations $-\Delta z_i^k = g_i^k$ *, owing to* $c = 0$*, except the case of the Neumann problem.)*
Then the sequence $(u^k) = (u_1^k, \dots, u_r^k) \subset D_Q^r$ *converges to* u^* *according to the linear estimate*

$$\|u^k - u^*\|_{H_Q^1(\Omega)^r} \le \frac{1}{m\sqrt{\lambda_1}} \|T(u^0) - g\|_{L^2(\Omega)^r} \big(\frac{M-m}{M+m} \big)^k \quad (k \in \mathbb{N}^+).$$

(Owing to Remark 2.11 this also means convergence in the usual $H^1(\Omega)$ *norm.)*

Proof. (a) First we remark the following facts: condition (C4) implies that for all $i, k = 1, \dots, r$ and $(x, \xi) \in \overline{\Omega} \times \mathbb{R}^r$

$$|\partial_{\xi_k} f_i(x, \xi)| \le \kappa' + \gamma \sum_{j=1}^r |\xi_j|^{p-2}\,.$$

Hence from Lagrange's inequality we have for all $i = 1, \dots, r$, $(x, \xi) \in \overline{\Omega} \times \mathbf{R}^r$

$$|f_i(x, \xi)| \le |f_i(x, 0)| + \Big(\kappa' + \gamma \sum_{j=1}^r |\xi_j|^{p-2} \Big) \sum_{k=1}^r |\xi_k| \le \kappa'' + \gamma' \sum_{j=1}^r |\xi_j|^{p-1} \tag{2.13}$$

with suitable constants $\kappa'', \gamma' > 0$.
(b) To prove our theorem, we have to check conditions (i)-(iv) of Theorem 2.2 and (v)' of Corollary 2.6 in our setting in the real Hilbert space $H = L^2(\Omega)$.

(i) For any $u \in D_Q^r$ we have

$$|T_i(u)| \le \sum_{k=1}^N \big(|\partial_k a_i\, \partial_k u_i| + |a_i\, \partial_k^2 u_i| \big) + |f_i(x, u_1, \dots, u_r)|\,.$$

Here $\partial_k a_i$ and a_i are in $C(\overline{\Omega})$, $\partial_k u_i$ and $\partial_k^2 u_i$ are in $L^2(\Omega)$, hence the sum term is in $L^2(\Omega)$. Further, assumption (C4) implies $2p - 2 < \frac{2N}{N-4}$, hence $H^2(\Omega) \subset L^{2p-2}(\Omega)$ [1]. Thus (2.13) yields $|f_i(x, u_1, \dots, u_r)| \le \kappa'' + \gamma' \sum_{j=1}^r |u_j|^{p-1} \in L^2(\Omega)$. That is, T_i maps indeed into $L^2(\Omega)$. Further,

assumption s (C1)-(C2) imply that for any $g \in L^2(\Omega)$ the weak solution of $-\Delta z + cz = g$ with $\alpha z + \beta \partial_\nu z\big|_{\partial\Omega} = 0$ is in $H^2(\Omega)$ [11], i.e. $R(B) = L^2(\Omega)$. Hence $R(B) \supset R(T_i)$ holds.

(ii) For any $u \in D_Q^r$, $v \in D_Q$ and $i = 1, \dots, r$ the divergence theorem yields (similarly to (2.7))

$$\begin{aligned} \langle B^{-1} T_i(u), v \rangle_{H_Q^1} &= \int_\Omega T_i(u) v \\ &= \int_\Omega \big(a_i \nabla u_i \cdot \nabla v + f_i(x,u) v \big)\, dx + \int_{\Gamma_\beta} a_i \frac{\alpha}{\beta} u_i v \, d\sigma \,. \end{aligned} \tag{2.14}$$

Let us put arbitrary $u \in H_Q^1(\Omega)^r, v \in H_Q^1(\Omega)$ in (2.14). Setting $Q(u) \equiv \gamma' \sum_{j=1}^r |u_j|^{p-1} = \gamma' \sum_{j=1}^r |u_j|^{p/q}$, where $p^{-1} + q^{-1} = 1$, we have $|f_i(x,u)| \le \kappa'' + Q(u)$ from (2.13) where $Q(u) \in L^q(\Omega)$. Then (2.14) can be estimated by the expression

$$\begin{aligned} &\max_{\overline{\Omega}} a_i \, \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + \kappa'' |\Omega|^{1/2} \|v\|_{L^2(\Omega)} \\ &+ \|Q(u)\|_{L^q(\Omega)} \|v\|_{L^p(\Omega)} + \eta \max_{\Gamma_\beta} a_i \, \|u\|_{L^2(\partial\Omega)} \|v\|_{L^2(\partial\Omega)}, \end{aligned}$$

where $|\Omega|$ is the measure of Ω. Using (2.8) and the continuity of the trace operator, we see that for any fixed $u \in H_Q^1(\Omega)^r$ the expression (2.14) defines a bounded linear functional on $H_Q^1(\Omega)^r$. Hence (using Riesz's theorem) we define the operator $F_i : H_Q^1(\Omega)^r \to H_Q^1(\Omega)$ by the formula

$$\langle F_i(u), v \rangle_{H_Q^1} = \int_\Omega \big(a_i \nabla u_i \cdot \nabla v + f_i(x,u) v \big)\, dx + \int_{\Gamma_\beta} a_i \frac{\alpha}{\beta} u_i v \, d\sigma \,,$$

$u \in H_Q^1(\Omega)^r, v \in H_Q^1(\Omega)$.

For $u \in H_Q^1(\Omega)^r$ let $S_i(u) \in B\left(H_Q^1(\Omega)^r, H_Q^1(\Omega)\right)$ be the bounded linear operator defined by

$$\langle S_i(u) h, v \rangle_{H_Q^1} = \int_\Omega \Big(a_i \nabla h_i \cdot \nabla v + \sum_{k=1}^r \partial_{\xi_k} f_i(x,u) h_k v \Big)\, dx + \int_{\Gamma_\beta} a_i \frac{\alpha}{\beta} h_i v \, d\sigma \tag{2.15}$$

$u \in H_Q^1(\Omega)^r, v \in H_Q^1(\Omega)$. The existence of $S_i(u)$ is provided by Riesz's theorem similarly as above. Now having the estimate

$$\begin{aligned} &\int_\Omega |\partial_{\xi_k} f_i(x,u) h_k v| \, dx \\ &\le \kappa' \|h_k\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \gamma \Big\| \sum_{k=1}^r |u_j|^{p-2} \Big\|_{L^{\frac{p}{p-2}}(\Omega)} \|h_k\|_{L^p(\Omega)} \|v\|_{L^p(\Omega)} \end{aligned}$$

for the terms with f_i, using $(\frac{p}{p-2})^{-1} + p^{-1} + p^{-1} = 1$. We will prove that F_i is Gâteaux differentiable ($i = 1, \dots, r$), namely,

$$F_i'(u) = S_i(u) \quad \left(u \in H_Q^1(\Omega)^r\right) .$$

Let $u, h \in H^1_Q(\Omega)^r$, further, $\mathcal{E} \equiv \{v \in H^1_Q(\Omega) : \|v\|_{H^1_Q(\Omega)} = 1\}$ and

$$\begin{aligned}\delta^i_{u,h}(t) &\equiv \frac{1}{t}\|F_i(u+th) - F_i(u) - tS_i(u)h\|_{H^1_Q(\Omega)} \\ &= \frac{1}{t}\sup_{v\in\mathcal{E}}\langle F_i(u+th) - F_i(u) - tS_i(u)h, v\rangle_{H^1_Q(\Omega)}\,.\end{aligned}$$

Then, using linearity, we have

$$\begin{aligned}\delta^i_{u,h}(t) &= \frac{1}{t}\sup_{v\in\mathcal{E}}\int_\Omega \Big(f_i(x,u+th) - f_i(x,u) - t\sum_{k=1}^r \partial_{\xi_k} f_i(x,u)h_k\Big)v\,dx \\ &= \sup_{v\in\mathcal{E}}\int_\Omega \sum_{k=1}^r \left(\partial_{\xi_k} f_i(x,u+\theta th) - \partial_{\xi_k} f_i(x,u)\right)h_k v\,dx \\ &\le \sup_{v\in\mathcal{E}}\sum_{k=1}^r \Big(\int_\Omega |\partial_{\xi_k} f_i(x,u+\theta th) - \partial_{\xi_k} f_i(x,u)|^{\frac{p}{p-2}}dx\Big)^{\frac{p-2}{p}} \\ &\quad\times \|h_k\|_{L^p(\Omega)}\|v\|_{L^p(\Omega)}\,.\end{aligned}$$

Here $\|v\|_{L^p(\Omega)} \le K_{p,\Omega}$ from (2.8), further, $|\theta th| \le |th| \to 0$ a.e. on Ω, hence the continuity of $\partial_{\xi_k} f_i$ implies that the integrands converge a.e. to 0 when $t \to 0$. The integrands are majorized for $t \le t_0$ by

$$\Big|2(\kappa' + \gamma\sum_{j=1}^r |u_j + t_0 h_j|^{p-2})\Big|^{\frac{p}{p-2}} \le \tilde\kappa + \tilde\gamma\sum_{j=1}^r |u_j + t_0 h_j|^p$$

in $L^1(\Omega)$, hence, by Lebesgue's theorem, the obtained expression tends to 0 when $t \to 0$. Thus

$$\lim_{t\to 0}\delta^i_{u,h}(t) = 0\,.$$

(iii) It is proved similarly to (ii) that for fixed functions $u, k, w \in H^1_Q(\Omega)^r, h \in H^1_Q(\Omega)$ the mapping $s,t \mapsto F_i'(u+sk+tw)h$ is continuous from $\mathbf{R}^2$ to $H^1_Q(\Omega)$. Namely,

$$\begin{aligned}\omega_{u,k,w,h}(s,t) &\equiv \|F_i'(u+sk+tw)h - F_i'(u)h\|_{H^1_Q(\Omega)} \\ &= \sup_{v\in\mathcal{E}}\langle F_i'(u+sk+tw)h - F_i'(u)h, v\rangle_{H^1_Q(\Omega)} \\ &= \sup_{v\in\mathcal{E}}\int_\Omega \sum_{k=1}^r \left(\partial_{\xi_k} f_i(x,u+sk+tw) - \partial_{\xi_k} f_i(x,u)\right)h_k v\,dx\,.\end{aligned}$$

Then we obtain just as above (from the continuity of $\partial_{\xi_k} f_i$ and Lebesgue's theorem) that

$$\lim_{s,t\to 0}\omega_{u,k,w,h}(s,t) = 0\,.$$

(iv) It follows from $F_i'(u) = S_i(u)$ in (2.15) and from the assumed symmetry of the Jacobians $\partial_\xi f(x,\xi)$ that for any $u, h, v \in H^1_Q(\Omega)^r$

$$\sum_{i=1}^r \langle F_i'(u)h, v_i\rangle_{H^1_Q(\Omega)} = \sum_{i=1}^r \langle h_i, F_i'(u)v\rangle_{H^1_Q(\Omega)}.$$

(v) For any $u, h \in H^1_Q(\Omega)^r$ we have

$$\sum_{i=1}^r \langle F_i'(u)h, h_i\rangle_{H^1_Q(\Omega)}$$
$$= \int_\Omega \Big(\sum_{i=1}^r a_i|\nabla h_i|^2 + \sum_{i,k=1}^r \partial_{\xi_k} f_i(x,u)h_k h_i\Big)\,dx + \int_{\Gamma_\beta} \frac{\alpha}{\beta}\sum_{i=1}^r a_i h_i^2\,d\sigma\,.$$

Hence from assumptions (C3)-(C4) we have

$$\sum_{i=1}^r \langle F_i'(u)h, h_i\rangle_{H^1_Q(\Omega)}$$
$$\geq m\int_\Omega \sum_{i=1}^r |\nabla h_i|^2\,dx + \kappa\int_\Omega \sum_{i=1}^r h_i^2\,dx + m\int_{\Gamma_\beta}\frac{\alpha}{\beta}\sum_{i=1}^r h_i^2\,d\sigma$$
$$= m\|h\|^2_{H^1_Q(\Omega)^r}$$

using $\kappa = cm$ (see Def.2.2). Further,

$$\begin{aligned}\sum_{i=1}^r \langle F_i'(u)h, h_i\rangle_{H^1_Q(\Omega)} &\leq m'\sum_{i=1}^r\int_\Omega |\nabla h_i|^2\,dx + \eta m'\sum_{i=1}^r\int_{\Gamma_\beta} h_i^2\,d\sigma \\ &\quad + \int_\Omega \Big(\kappa' + \gamma\sum_{j=1}^r |u_j|^{p-2}\Big)\sum_{k=1}^r h_k^2\,dx\,.\end{aligned} \tag{2.16}$$

Here

$$\kappa'\sum_{k=1}^r\int_\Omega h_k^2\,dx \leq \frac{\kappa'}{\lambda_1}\|h\|^2_{H^1_Q(\Omega)^r}$$

from Corollary 2.9 (b). Further, from Hölder's inequality, using $\frac{p-2}{p} + \frac{2}{p} = 1$, we obtain

$$\begin{aligned}\sum_{j,k=1}^r \int_\Omega |u_j|^{p-2}h_k^2\,dx &\leq \sum_{j,k=1}^r \Big[\int_\Omega \big(|u_j|^{p-2}\big)^{\frac{p}{p-2}}\Big]^{\frac{p-2}{p}}\Big[\int_\Omega \big(h_k^2\big)^{p/2}\Big]^{2/p} \\ &= \sum_{j,k=1}^r \|u_j\|^{p-2}_{L^p(\Omega)}\|h_k\|^2_{L^p(\Omega)} \\ &= \Big(\sum_{j=1}^r \|u_j\|^{p-2}_{L^p(\Omega)}\Big)\Big(\sum_{k=1}^r \|h_k\|^2_{L^p(\Omega)}\Big)\,.\end{aligned}$$

An elementary extreme value calculation shows that for $x \in \mathbb{R}^r$, $\sum_{j=1}^r x_j^2 = R^2$ the values of $\Big(\sum_{j=1}^r |x_j|^{p-2}\Big)^{\frac{2}{p-2}}$ lie between R^2 and $r^{\frac{4-p}{p-2}}R^2$, i.e.

$$\sum_{j=1}^r |x_j|^{p-2} \leq \mu_p\Big(\sum_{j=1}^r |x_j|^2\Big)^{\frac{p-2}{2}}$$

where $\mu_p = \max\{1, r^{\frac{4-p}{p-2}}\}$. Hence

$$\begin{aligned}\sum_{j,k=1}^{r}\int_\Omega |u_j|^{p-2}h_k^2\,dx &\le \mu_p\Big(\sum_{j=1}^{r}\|u_j\|^2_{L^p(\Omega)}\Big)^{\frac{p-2}{2}}\Big(\sum_{k=1}^{r}\|h_k\|^2_{L^p(\Omega)}\Big)\\ &\le \mu_p K^p_{p,\Omega}\Big(\sum_{j=1}^{r}\|u_j\|^2_{H^1_Q}\Big)^{\frac{p-2}{2}}\Big(\sum_{k=1}^{r}\|h_k\|^2_{H^1_Q}\Big)\\ &= \mu_p K^p_{p,\Omega}\|u\|^{p-2}_{H^1_Q(\Omega)^r}\|h\|^2_{H^1_Q(\Omega)^r}\,.\end{aligned}$$

Summing up, (2.16) yields

$$\sum_{i=1}^{r}\langle F_i'(u)h, h_i\rangle_{H^1_Q(\Omega)} \le\ M(u)\|h\|^2_{H^1_Q(\Omega)^r}\quad (u,h\in H^1_Q(\Omega)^r)$$

with

$$M(u) = \max\{1,\eta\}m' + \kappa'\lambda_1^{-1} + \gamma K^p_p(\Omega)\mu_p\|u\|^{p-2}_{H^1_Q(\Omega)^r}\,.$$

Since Corollary 2.9 (b) implies $\|u\|_{H^1_Q} \le \lambda_1^{-1/2}\|Bu\|_{L^2(\Omega)}$ $(u\in D_Q)$, therefore the radius $r_0 = m^{-1}\|F(u^0)-b\|_{H^1_Q(\Omega)^r}$ (defined in Corollary 2.6) fulfils

$$r_0 \le m^{-1}\lambda_1^{-1/2}\big(\sum_{i=1}^{r}\|T_i(u^0)-g_i\|^2_{L^2(\Omega)}\big)^{1/2},$$

using $BF_{i|D} = T_i$. Hence for $u\in B(u^0,r_0) = \{u\in H^1_Q(\Omega)^r :\ \|u-u^0\|_{H^1_Q(\Omega)^r}\le r_0\}$ we have

$$\sum_{i=1}^{r}\langle F_i'(u)h, h_i\rangle_{H^1_Q(\Omega)} \le\ M\|h\|^2_{H^1_Q(\Omega)^r}\quad (u\in B(u^0,r_0), h\in H^1_Q(\Omega)^r)$$

with M defined in (2.9). □

Remark 2.15. Theorem 2.14 holds similarly in the following cases:
(a) with other smoothness assumption s on $\partial\Omega$ and the coefficients of T, when the inclusion $R(T_i)\subset R(B)$ is fulfilled with suitable domain $D(T_i)$ of T_i instead of D^r_Q (cf. (2.5)).
(b) with more general linear part $-\operatorname{div}(A_i(x)\nabla u_i)$ of T_i, where $A_i\in C^1(\overline{\Omega},\mathbb{R}^{N\times N})$, in the case of Dirichlet boundary condition.

The above theorem is the extension of the cited earlier results on the gradient method to system (2.4). We note that the *conjugate gradient method* might be similarly extended to (2.4), following its application in [18] for a single Dirichlet problem. As mentioned earlier, the CGM constructs two simultaneous sequences, and it improves the convergence ratio of the gradient method to $(\sqrt{M}-\sqrt{m})(\sqrt{M}+\sqrt{m})$ at the price of an extra work which comprises the calculation of required integrals and numerical root finding for the stepsize.

Compared to *Newton-like methods* (which can provide higher order convergence than linear), we emphasize that in the iteration of Theorem 2.3 the auxiliary problems are of fixed (Poisson) type, whereas Newton-like methods involve stepwise different linearized problems with variable coefficients. Hence in our iteration one can exploit the efficient solution methods that exist for the Poisson equation. (More discussion on this will be given in subsection 4.2.)

2.3. **Sobolev embedding properties.** The construction of the sequence (2.10) in Theorem 2.3 needs an explicit value of the constant M in (2.9). The parameters involved in M are defined in conditions (C3)-(C4) only with the exception of the embedding constants $K_{p,\Omega}$ in (2.8). (The eigenvalue λ_1 fulfils $\lambda_1 = K_{2,\Omega}^{-2}$ by virtue of Corollary 2.9 (b).) Consequently, in order to define the value of M, the exact value or at least a suitable estimate is required for the constants $K_{p,\Omega}$.

Although most of the exact constant problems have been solved in $\mathbb{R}^n$, even in the critical exponent case (see [32, 34]), for bounded domains there is not yet complete answer. For $n \geq 3$ and for small square in the case $n = 2$, the embedding constants of $H^1(\Omega)$ to $L^p(\Omega)$ are given in [7, 8]; an estimate is given for functions partly vanishing on the boundary for the critical exponent case in [24]. Consequently, a brief study of the embeddings is worth wile to obtain estimates of the embedding constants which are valid for $n = 2$. Our estimates, presented in two dimensions, take into account the boundary values of the functions.

Besides the constants $K_{p,\Omega}$ in (2.8), for any set $\Gamma \subset \partial\Omega$ we denote by $K_{p,\Gamma}$ the embedding constant in the estimate

$$\|u\big|_\Gamma\|_{L^p(\Gamma)} \leq K_{p,\Gamma}\|u\|_{H^1_Q} \quad (u \in H^1_Q(\Omega)).$$

Lemma 2.16. *Let $I = [a,b] \times [c,d] \subset \mathbb{R}^2$, $p_i \geq 1$ $(i = 1,2)$. The boundary ∂I is decomposed into $\Gamma_1 = \{a,b\} \times [c,d]$ and $\Gamma_2 = [a,b] \times \{c,d\}$. Then*

$$K_{p_1+p_2,I}^{p_1+p_2} \leq \frac{1}{2}\big(K_{p_1,\Gamma_1}^{p_1} + p_1 K_{2(p_1-1),I}^{p_1-1}\big)\big(K_{p_2,\Gamma_2}^{p_2} + p_2 K_{2(p_2-1),I}^{p_2-1}\big)\,.$$

Proof. Let $u \in H^1(I)$. We define the function $u_a(y) \equiv u(a,y)$, and similarly u_b, u_c and u_d. Then for any $x, y \in I$ we have

$$\begin{aligned} |u(x,y)|^{p_1} &= |u_a(y)|^{p_1} + p_1 \int_a^x |u(\xi,y)|^{p_1-2} u(\xi,y)\partial_1 u(\xi,y)\,d\xi \\ &\leq |u_a(y)|^{p_1} + p_1 \int_a^b |u|^{p_1-1}|\partial_1 u|\,dx\,. \end{aligned} \tag{2.17}$$

Similarly, we obtain

$$|u(x,y)|^{p_2} \leq |u_c(y)|^{p_2} + p_2 \int_c^d |u|^{p_2-1}|\partial_2 u|\,dy\,. \tag{2.18}$$

Multiplying (2.17) and (2.18) and then integrating over I, we obtain

$$\begin{aligned} \int_I |u|^{p_1+p_2} &\leq \Big(\int_c^d |u_a|^{p_1} + p_1 \int_I |u|^{p_1-1}|\partial_1 u|\Big)\Big(\int_a^b |u_c|^{p_2} + p_2 \int_I |u|^{p_2-1}|\partial_2 u|\Big) \\ &\leq \Big(\int_c^d |u_a|^{p_1} + p_1\|u\|^{p_1-1}_{L^{2(p_1-1)}(I)}\|\partial_1 u\|_{L^2(I)}\Big) \\ &\quad \times \Big(\int_a^b |u_c|^{p_2} + p_2\|u\|^{p_2-1}_{L^{2(p_2-1)}(I)}\|\partial_2 u\|_{L^2(I)}\Big)\,. \end{aligned}$$

The same holds with u_b and u_d instead of u_a and u_b. Using the elementary inequality

$$\begin{aligned} &(\alpha_1 + r_1\gamma_1)(\alpha_2 + r_2\gamma_2) + (\beta_1 + r_1\gamma_1)(\beta_2 + r_2\gamma_2) \\ &\leq (\alpha_1 + \beta_1 + r_1\sqrt{\gamma_1^2 + \gamma_2^2})(\alpha_2 + \beta_2 + r_2\sqrt{\gamma_1^2 + \gamma_2^2}) \end{aligned}$$

for $\alpha_i, \beta_i, r_i, \gamma_i \geq 0$ $(i = 1, 2)$, we obtain

$$\begin{aligned} 2\int_I |u|^{p_1+p_2} &\leq \Big(\int_c^d (|u_a|^{p_1} + |u_b|^{p_1}) + p_1\|u\|^{p_1-1}_{L^{2(p_1-1)}(I)}\|\nabla u\|_{L^2(I)}\Big) \\ &\quad \times \Big(\int_a^b (|u_c|^{p_2} + |u_d|^{p_2}) + p_2\|u\|^{p_2-1}_{L^{2(p_2-1)}(I)}\|\nabla u\|_{L^2(I)}\Big) \\ &\leq \Big(K^{p_1}_{p_1,\Gamma_1} + p_1 K^{p_1-1}_{2(p_1-1),I}\Big)\Big(K^{p_2}_{p_2,\Gamma_2} + p_2 K^{p_2-1}_{2(p_2-1),I}\Big)\|u\|^{p_1+p_2}_{H^1_Q}\,. \end{aligned}$$

□

Corollary 2.17. *Let $\Omega \subset \mathbb{R}^2$ with $\partial\Omega \in C^1$ and let us consider Dirichlet boundary condition s in* (2.4)*, i.e. $Qu \equiv u$ and $H^1_Q(\Omega) = H^1_0(\Omega)$. Then*

$$K^{p_1+p_2}_{p_1+p_2,\Omega} \leq \frac{p_1 p_2}{2} K^{p_1-1}_{2(p_1-1),\Omega} K^{p_2-1}_{2(p_2-1),\Omega}\,. \tag{2.19}$$

Proof. Ω is included in some $I = [a, b] \times [c, d]$, and for any $u \in H^1_0(\Omega)$ its extension $\tilde{u} \in H^1_0(I)$ is defined as zero on $I \setminus \Omega$. Then for any $p \geq 1$ we have $K_{p,\Gamma_i} = 0$ and $K_{p,\Omega} = K_{p,I}$. □

The case when p is an even integer is of particular importance since we have this situation in the case of (2.6) owing to (C4). Then the functional inequality (2.19) leads directly to an estimate:

Corollary 2.18. *Let $\lambda_1 > 0$ be the smallest eigenvalue of $-\Delta$ on $H^1_0(\Omega)$. Then*

(a) $K_{2n,\Omega} \leq 2^{-1/2}\left(\frac{2}{\lambda_1}\right)^{1/2n}(n!)^{1/n}$ $(n \in \mathbb{N}^+)$*;*

(b) $K_{2n,\Omega} \leq 0.63 b_n n$ $(n \in \mathbb{N}^+, n \geq 2)$ *where $b_n = (2/\lambda_1)^{1/2n}$ (and thus $\lim b_n = 1$).*

Proof. (a) Let $h(p) = K^p_{p,\Omega}$ $(p \geq 1)$. Then for $n \in \mathbb{N}^+$ (2.19) implies the recurrence

$$h(2n) \leq h(2n-2)\frac{n^2}{2},$$

hence

$$h(2n) \leq h(2)\frac{(n!)^2}{2^{n-1}} = \frac{2}{\lambda_1}\frac{(n!)^2}{2^n}\,,$$

since Corollary 2.9 gives $K_{2,\Omega} = \lambda_1^{-1/2}$.
(b) The estimate $(n!)^{1/n} \leq 0.891\, n$ $(n \geq 2)$ is used. □

The boundary embedding constants K_{p,Γ_i} can be estimated in terms of suitable $K_{p'}(I)$ as follows.

Lemma 2.19. *Let I and Γ_i $(i = 1, 2)$ be as in Lemma 2.16, $p \geq 1$. Then*

$$K^p_{p,\Gamma_i} \leq \frac{2}{b-a}K^p_{p,I} + p\sqrt{2}K^{p-1}_{2(p-1),I}\,.$$

Proof. We prove the lemma for Γ_1. Similarly to Lemma 2.16 we have

$$|u_a(y)|^p \leq |u(x,y)|^p + p\int_a^b |u|^{p-1}|\partial_1 u|\,dx \quad (x, y \in I).$$

Integrating over I, we obtain

$$(b-a)\int_c^d |u_a|^p \le \int_I |u|^p + p(b-a)\int_I |u|^{p-1}|\partial_1 u| \\ \le \|u\|^p_{L^p(\Omega)} + p(b-a)\|u\|^{p-1}_{L^{2(p-1)}(\Omega)}\|\partial_1 u\|_{L^2(\Omega)}\,,$$

and similarly for u_b. Hence, summing up and using $\|\partial_1 u\|_{L^2(\Omega)} + \|\partial_2 u\|_{L^2(\Omega)} \le \sqrt{2}\|\nabla u\|_{L^2(\Omega)} \le \sqrt{2}\|u\|_{H^1_Q}$, we have

$$\|u\|^p_{L^p(\Gamma_1)} \le \big(\frac{2}{b-a}K^p_{p,I} + p\sqrt{2}K^{p-1}_{2(p-1),I}\big)\|u\|^p_{H^1_Q}\,.$$

□

Corollary 2.20. *For any $n \in \mathbb{N}^+$ we have*

$$K_{2n,\Gamma_i} \le 1.26 c_n n\,,$$

where $c_n = \frac{1}{2}\left(\frac{4}{\lambda_1(b-a)} + \frac{4^{n+1}}{\sqrt{2\lambda_1}}\right)^{1/2n}$ *(and thus* $\lim c_n = 1$*).*

The proof of this corollary follow using Corollary 2.18 (a) and again $n! < (0.891n)^n$.

Remark 2.21. Lemmas 2.1 and 2.2 may be extended from the interval case to other domains, depending on the actual shape of Ω, if portions Γ of $\partial\Omega$ are parametrized as e.g. $t \mapsto (t, \varphi(t))$ and inequalities of the type $\int_\alpha^\beta |u|^p\,dx \le \int_\alpha^\beta |u(t,\varphi(t))|^p(1+\varphi'(t))^{1/2}\,dt = \int_\Gamma |u|^p$ are used. Then estimates can be derived depending on the portions Γ_i of $\partial\Omega$ where $u\big|_{\Gamma_i} = 0$ (i.e. $K_{p,\Gamma_i} = 0$). The detailed investigation is out of the scope of this paper. (For a model problem a calculation of the corresponding estimate for mixed boundary condition s will be given in section 6.)

3. Implementation of the method

In Section 2, the Sobolev space gradient method was developed for systems of the form (2.4). Thereby a theoretical iteration is executed directly for the original boundary-value problem in the Sobolev space, and it is shown that the iteration converges in the corresponding energy norm.

One of the main features of this approach is the reduction of computational questions to those arising for the auxiliary linear Poisson problems. Namely, in the application to a given system of boundary-value problems one's task is to choose a numerical method for the Poisson problems and solve the latter to some suitable accuracy. This means that from now on two issues have to be handled: from the aspect of convergence, error control for the stepwise solution of the auxiliary problems, and from the aspect of cost, the efficient solution of the Poisson equations.

Section 4 is devoted to these topics. First a discussion of corresponding error estimates is given for the numerically constructed sequences. Then we will refer very briefly to some efficient Poisson solvers. Here we note that the efficiency of the whole iteration much relies on the fact that all the linear problems are of the same Poisson type, for which efficient solvers are available.

In Sections 5–6 we consider the simplest case of realization as an elementary illustration of the theoretical results. This suits the semilinear setting of this paper and involves the case of polynomial nonlinearity (2.6), connected to reaction-diffusion systems. Namely, on some special domains the GM is applied in effect directly to

the BVP itself since the Poisson equations are solved exactly. This is due to keeping the iteration in special function classes. We give a brief summary of some cases of such special domains. Then the paper is closed with an example that illustrates the convergence result.

4. Error control and efficiency

4.1. **Error estimates for the numerical iterations.** The theoretical iteration $(u^k) = (u_1^k, \dots, u_r^k)$ $(k \in \mathbb{N})$, defined in Theorem 2.1, can be written as

$$u^{k+1} = u^k - \frac{2}{M+m} z^k$$
$$\text{where } z^k = \mathcal{B}^{-1}(T(u^k) - g),$$

using the notation

$$\mathcal{B} : D_Q^r \to L^2(\Omega)^r, \quad \mathcal{B}(w_1, \dots, w_r) \equiv (Bw_1, \dots, Bw_r).$$

Recall that B is defined on D_Q (see (2.5)), containing the boundary conditions, and we have $Bu \equiv -\Delta u$ if $\alpha \not\equiv 0$ and $Bu \equiv -\Delta u + cu$ if $\alpha \equiv 0$ (i.e. for Neumann BC).

Any kind of numerical implementation of the GM in the Sobolev space $H_Q^1(\Omega)^r$ defines a sequence $(\overline{u}^k)$, constructed as follows:

$$\begin{gathered} \overline{u}^0 = u^0 \in D; \\ \text{for } k \in \mathbb{N}: \quad \overline{u}^{k+1} = \overline{u}^k - \frac{2}{M+m} \overline{z}^k, \\ \text{where } \overline{z}^k \approx z_*^k \equiv \mathcal{B}^{-1}(T(\overline{u}^k) - g) \\ \text{such that } \|\overline{z}^k - z_*^k\|_{H_Q^1(\Omega)^r} \le \delta_k \end{gathered} \tag{4.1}$$

where $(\delta_k) \subset \mathbb{R}^+$ is a real sequence. Then our task is to estimate $\|\overline{u}^k - u^*\|_{H_Q^1(\Omega)^r}$ in terms of the sequence (δ_k), where $u^* = (u_1^*, \dots, u_r^*) \in H_Q^1(\Omega)^r$ is the weak solution of the system (2.4).

We define

$$E_k \equiv \|u^k - \overline{u}^k\|_{H_Q^1(\Omega)^r}.$$

By Theorem 2.1 we have

$$\|\overline{u}^k - u^*\|_{H_Q^1(\Omega)^r} \le E_k + \frac{R_0}{m\sqrt{\lambda_1}} \left(\frac{M-m}{M+m}\right)^k \quad (k \in \mathbb{N}^+)$$

where $R_0 = \|T(u^0) - g\|_{L^2(\Omega)^r}$ denotes the initial residual. Hence the required error estimates depend on the behaviour of (E_k).

We have proved the following two results in [13] for a single Dirichlet problem. Since they are entirely based on the bounds m and M of the generalized differential operator (which is also the background for our Theorem 2.1), they can be immediately formulated in our setting, too.

Proposition 4.1 ([13]). *For all $k \in \mathbb{N}$*

$$E_{k+1} \le \frac{M-m}{M+m} E_k + \frac{2}{M+m} \delta_k.$$

Corollary 4.2 ([13]). *Let $0 < q < 1$ and $c_1 > 0$ be fixed, $\delta_k \le c_1 q^k$ $(k \in \mathbb{N})$. Then the following estimates hold for all $k \in \mathbb{N}^+$:*

(a) *if* $q > \frac{M-m}{M+m}$ *then* $\|\overline{u}^k - u^*\|_{H^1_Q(\Omega)^r} \le c_2 q^k$;

(b) *if* $q < \frac{M-m}{M+m}$ *then* $\|\overline{u}^k - u^*\|_{H^1_Q(\Omega)^r} \le c_3 \left(\frac{M-m}{M+m}\right)^k$

where $c_2 = \frac{\alpha c_1}{q-Q} + \frac{R_0}{m\sqrt{\lambda_1}}$, $c_3 = \frac{\alpha c_1}{Q-q} + \frac{R_0}{m\sqrt{\lambda_1}}$, $Q = \frac{M-m}{M+m}$.

Besides these convergence results, it is also of practical interest to ensure that we arrive only in a prescribed neighbourhood of the solution.

Proposition 4.3. *Let* $\varepsilon > 0$ *be fixed, let* $\delta_k \equiv m\varepsilon$. *Then we have for all* $k \in \mathbb{N}^+$

$$E_k < \varepsilon\,. \tag{4.2}$$

Proof. By definition $E_0 = \|u^0 - \overline{u}^0\|_{H^1_Q(\Omega)^r} = 0$. Assume that (4.2) holds for fixed $k \in \mathbb{N}^+$. Then Proposition 4.1 yields

$$E_{k+1} < \frac{M-m}{M+m}\varepsilon + \frac{2}{M+m}m\varepsilon = \varepsilon\,.$$

□

Corollary 4.4. *If* (δ_k) *is chosen as in Proposition 4.3, then for* $(k \in \mathbb{N}^+)$,

$$\|\overline{u}^k - u^*\|_{H^1_Q(\Omega)^r} \le \varepsilon + \frac{R_0}{m\sqrt{\lambda_1}}\Big(\frac{M-m}{M+m}\Big)^k.$$

4.2. **Discretization and efficient Poisson solvers.** The numerical solution of the Poisson equations is generally achieved by using some discretization, which is a finite difference method or a finite element method. In this respect we note two facts concerning the whole iteration. First, if we use one and the same fixed grid for each linear problem, then (4.1) is equivalent to a suitably precondition ed nonlinear FEM iteration such that the precondition er is the discrete Laplacian corresponding to the grid. On the other hand, the use of variable grids provide a multilevel type iteration.

We emphasize that for such iterations the convergence ratio in Theorem 2.1 yields an asymptotic value for those of the discretized problems, hence the numerical iterations using different grids (or grid sequences) exhibit mesh uniform convergence.

The numerical efficiency of the iteration (4.1) relies on the fact that all the linear problems are of the same Poisson type. Namely, various fast Poisson solvers are available that have been developed in the past decades. Many of these solvers were originally developed on rectangular domains and later extended to other domains via the fictitious domain approach. The fast direct solvers include the method of cyclic reduction, the fast Fourier transform and the FACR algorithm: comprehensive summaries on the solution of the Poisson equation with these methods are found in [10, 33]. Another family of fast solvers which has undergone recent development is formed by spectral methods [14]. The parallel implementation of these algorithms is also feasible. For the fictitious domain approach for general domains see [6].

5. Direct realization on special domains

We now consider the system (2.4) with polynomial nonlinearity (2.6). That is, we investigate the system

$$\begin{aligned} T_i(u_1,\dots,u_r) &\equiv -\operatorname{div}(a_i(x)\nabla u_i) + \sum_{|j|\le s_i} c_j^{(i)}(x) u_1^{j_1}\dots u_r^{j_r} = g_i(x) \\ Qu_i &\equiv (\alpha(x)u_i + \beta(x)\partial_\nu u_i)\big|_{\partial\Omega} = 0 \end{aligned} \tag{5.1}$$

that fulfils conditions (C1)-(C4) (given in Section 2), where the moving index is $j = (j_1,\dots,j_r) \in \mathbb{N}^r$ with $|j| = j_1 + \dots j_r$ and we have $c_j^{(i)} \in C(\overline{\Omega})$, $s_i \in \mathbb{N}^+$, $s_i \le p-1$. (This type of equations occurs e.g. in steady states or in time discretization of reaction-diffusion systems. Condition (C4) expresses that the reaction is of autocatalytic type.)

The main idea is the following: on special domains, first approximating the coefficients and right-sides of (5.1) by appropriate algebraic /trigonometric polynomials, the iteration (2.10) can also be kept in a suitable class of algebraic /trigonometric polynomials. Hence the solution of the auxiliary Poisson equations can be achieved directly by executing a linear combination (or, in the case of a ball, by solving a simply structured linear system of algebraic equations).

The solution of the approximate system (in which the coefficients and right-sides are approximated by polynomials) remains appropriately close to the solution of (5.1), depending on the accuracy of approximation. This can be formulated immediately in the case when the coefficients a_i and c_j^i are unchanged. (We typically have this situation with the operators T_i occurring in reaction-diffusion systems, wherein a_i and c_j^i are generally constant.) The evident but lengthy calculation for the variable coefficient case is left to the reader.

Proposition 5.1. *Let $\varepsilon > 0$. Let $\delta_i > 0$ $(i = 1,\dots,r)$ such that $(\sum_{i=1}^r \delta_i^2)^{1/2} < \lambda_1^{1/2} m\varepsilon$, and let $\|g_i - \tilde{g}_i\|_{L^2(\Omega)} < \delta_i$ $(i = 1,\dots,r)$. Denote by $u^* = (u_1^*,\dots,u_r^*)$ and $\tilde{u} = (\tilde{u}_1,\dots,\tilde{u}_r)$ the solution s of the systems*

$$T_i(u) = g_i(x) \quad \text{and} \quad T_i(u) = \tilde{g}_i(x)$$

(both with boundary condition $Qu_i\big|_{\partial\Omega} = 0$), respectively. Then $\|\tilde{u} - u^*\|_{H^1_Q(\Omega)^r} < \varepsilon$.

Proof. Note that

$$\begin{aligned} m\|u^* - \tilde{u}\|^2_{H^1_Q(\Omega)^r} &\le \sum_{i=1}^r \int_\Omega \big(T_i(u^*) - T_i(\tilde{u})\big)(u_i^* - \tilde{u}_i) \\ &\le \sum_{i=1}^r \|g_i - \tilde{g}_i\|_{L^2(\Omega)} \|u_i^* - \tilde{u}_i\|_{L^2(\Omega)} \\ &\le \lambda_1^{-1/2} \sum_{i=1}^r \delta_i \|u_i^* - \tilde{u}_i\|_{H^1_Q} \\ &\le \lambda_1^{-1/2} (\sum_{i=1}^r \delta_i^2)^{1/2} \|u_i^* - \tilde{u}_i\|_{H^1_Q(\Omega)^r} < m\varepsilon \|u_i^* - \tilde{u}_i\|_{H^1_Q(\Omega)^r}. \end{aligned}$$

□

In the sequel we consider (5.1) as having polynomial coefficients, i.e. being replaced by the approximate system. For simplicity of presentation, we only consider the two-dimension al case (the analogies being straightforward).

(a) Rectangle. We investigate the case when $\Omega \subset \mathbb{R}^2$ is a rectangle. It can be assumed that $\Omega = I \equiv [0,\pi]^2$ (if not, a linear transformation is done).

Denote by $\mathcal{P}_s$ and $\mathcal{P}_c$ the set of sine- and cosine-polynomials

$$\mathcal{P}_s = \{ \sum_{n,m=1}^{l} \sigma_{nm} \sin nx \sin my : \ l \in \mathbb{N}^+, \ \sigma_{nm} \in \mathbb{R} \, (n,m = 1,\dots,l)\},$$

$$\mathcal{P}_c = \{ \sum_{n,m=0}^{l} \sigma_{nm} \cos nx \cos my : \ l \in \mathbb{N}, \ \sigma_{nm} \in \mathbb{R} \, (n,m = 0,\dots,l)\},$$

respectively. The coefficients a_i and c_j^i of (5.1) can be approximated by cosine-polynomials to any prescribed accuracy, hence (as suggested above) we assume that (5.1) already fulfils $a_i, c_j^i \in \mathcal{P}_c$ $(i = 1,\dots,r; |j| \le s_i)$.

Dirichlet boundary conditions. We consider the case when $|j|$ is odd in each term in (5.1) and assume $g_i \in \mathcal{P}_s$. Then the operators T_i are invariant on $\mathcal{P}_s$. Further, for any $h \in \mathcal{P}_s$ the solution of

$$-\Delta z = h, \quad z\big|_{\partial I} = 0$$

fulfils $z \in \mathcal{P}_s$, namely, if $h(x,y) = \sum_{n,m=1}^{l} \sigma_{nm} \sin nx \sin my$, then

$$z(x,y) = \sum_{n,m=1}^{l} \frac{\sigma_{nm}}{n^2+m^2} \sin nx \sin my \,. \tag{5.2}$$

These imply that if (2.10) starts with $u_i^0 \in \mathcal{P}_s$ then we have $u_i^k \in \mathcal{P}_s$ throughout the iteration, and the auxiliary problems are solved in the trivial way (5.2).

Neumann boundary conditions. Similar to the Dirichlet case, now letting $g_i \in \mathcal{P}_c$. Here T_i are invariant on $\mathcal{P}_c$. Further, for $h \in \mathcal{P}_c$, the solution $h(x,y) = \sum_{n,m=0}^{l} \sigma_{nm} \cos nx \cos my$, of

$$-\Delta z + cz = h, \quad \partial_\nu z\big|_{\partial I} = 0$$

fulfils $z \in \mathcal{P}_c$, namely, $z(x,y) = \sum_{n,m=0}^{l} \frac{\sigma_{nm}}{n^2+m^2+c} \cos nx \cos my$.

Mixed boundary conditions. Denote by Γ_i $(i = 1,\dots,4)$ the boundary portions $[0,\pi) \times \{0\}$, $\{\pi\} \times [0,\pi)$, $(0,\pi] \times \{\pi\}$, $\{0\} \times (0,\pi]$, respectively. First, let $\alpha(x) \equiv \chi_{\{\Gamma_2 \cup \Gamma_4\}}$, $\beta(x) \equiv \chi_{\{\Gamma_1 \cup \Gamma_3\}}$, i.e. we have $u\big|_{\Gamma_2 \cup \Gamma_4} \equiv 0$, $\partial_\nu u\big|_{\Gamma_1 \cup \Gamma_3} \equiv 0$. Then the above method works on $\mathcal{P}_{sc} = \{\sum_{n=1}^{l} \sum_{m=0}^{l} \sigma_{nm} \sin nx \cos my \}$.

We can proceed similarly for other edgewise complementary characteristic functions α and β, using $\sin(n + \frac{1}{2})x$ type terms for mixed endpoint conditions.

(b) Disc. We investigate the case when $\Omega \subset \mathbb{R}^2$ is the unit disc $B = B_1(\mathbf{0})$. Now a_i, c_j^i and g_i are assumed to be algebraic polynomials. (Notation: $a_i, c_j^i, g_i \in \mathcal{P}_{alg}$). Then T_i are invariant on $\mathcal{P}_{alg}$.

Dirichlet boundary conditions. If $h \in \mathcal{P}_{alg}$, then the solution of

$$-\Delta z = h, \quad z\big|_{\partial B} = 0$$

can be found by looking for z in the form

$$z(x,y) = (x^2+y^2-1)q(x,y),$$

where $q \in \mathcal{P}_{alg}$ has the same degree as h (cf. [22]). Then the coefficients of q can be determined by solving a a simply structured linear system of algebraic equations (obtained from equating the coefficients of $-\Delta z$ and h). The matrix of this system has three diagonals and two other non-zero elements in each row.

3^{rd} *boundary conditions.* We examine the case $\alpha(x) \equiv \alpha > 0$, $\beta(x) \equiv \beta > 0$. If $h \in \mathcal{P}_{alg}$ then the solution of

$$-\Delta z = h, \quad (\alpha z + \beta \partial_\nu z)\big|_{\partial B} = 0$$

can be determined similarly to the Dirichlet case. Namely, let $q(x,y)$ be a polynomial with unknown coefficients $\{a_{nm}\}$ and of the same degree as h. Let

$$p(x,y) \equiv (x^2+y^2-1)q(x,y) = \sum_{s=0}^{l} \sum_{n+m=s} b_{nm} x^n y^m,$$

here the b_{nm}'s are linear combination s of the a_{nm}'s. We look for z as

$$z(x,y) = \sum_{s=0}^{l} \sum_{n+m=s} \frac{b_{nm}}{\alpha+\beta s} x^n y^m \,.$$

Equating the coefficients of $-\Delta z$ and h leads again to a linear system of algebraic equations for $\{a_{nm}\}$, from which we then determine $\{b_{nm}\}$. Then z fulfils the boundary condition, since on ∂B we have $\partial_\nu z = x\partial_x z + y\partial_y z$ and

$$\alpha z + \beta(x\partial_x z + y\partial_y z) = \sum_{s=0}^{l} \sum_{n+m=s} \frac{b_{nm}}{\alpha+\beta s}(\alpha+\beta(n+m)) x^n y^m = p(x,y) = 0\,.$$

(c) Annulus. Let $A = \{(x,y) \in \mathbb{R}^2 : R_1^2 < x^2+y^2 < R_2^2\}$ with given $R_2 > R_1 > 0$. Let $\Gamma_m = C(\mathbf{0}, R_m)$ $(m = 1,2)$ be the boundary circles. We investigate the system with radially symmetric coefficients, written in the form

$$\begin{gathered} T_i(u) \equiv -\overline{\nabla}\big(a_i(x^2+y^2)\overline{\nabla}u_i\big) + \sum_{|j|\le s_i} c_j^{(i)}(x^2+y^2) u_1^{j_1}\dots u_n^{j_n} = g_i(x^2+y^2) \\ Qu_i \equiv (\alpha u_i + \beta\partial_\nu u_i)\big|_{\Gamma_m} = 0 \quad (m=1,2) \end{gathered} \tag{5.3}$$

$(i = 1,\dots,n)$, with the notation $\overline{\nabla}u = \overline{\partial}_x u + \overline{\partial}_y u$, $\overline{\partial}_x u = x\partial_x u$, $\overline{\partial}_y u = y\partial_y u$. The functions $a_i \in C^1[R_1,R_2]$, $c_j^{(i)} \in C[R_1,R_2]$, $g_i \in L^2[R_1,R_2]$ and the numbers $\alpha,\beta \ge 0$ are such that the positivity and monotonicity conditions (C2)-(C4) are fulfilled.

Introducing the notations

$$r = (x^2+y^2)^{\frac{1}{2}}, \quad \hat{a}_i(r) = r a_i(r^2), \quad \hat{c}_j^{(i)}(r) = \frac{1}{r} c_j^{(i)}(r^2), \quad \hat{g}_i(r) = \frac{1}{r} g_i(r^2)$$

and using $\overline{\nabla}u = r\partial_r u$, the system (5.3) is written as

$$\begin{gathered} \hat{T}_i(u) \equiv -\frac{1}{2\pi r}\Big(\partial_r(\hat{a}_i(r)\partial_r u_i) + \sum_{|j|\le s_i} \hat{c}_j^{(i)}(r) u_1^{j_1}\dots u_n^{j_n}\Big) = \frac{1}{2\pi r}\hat{g}_i(r) \\ \hat{Q}_m u_i \equiv (\alpha u_i + (-1)^m \beta\partial_r u_i)\big|_{r=R_m} = 0 \quad (m=1,2). \end{gathered}$$

Let

$$Bu \equiv \frac{1}{2\pi r}(-\partial_r^2 u + cu) \quad \text{on } D(B) = D(T_i) = H^2_{rad}(A)$$

where $H^2_{rad}(A) = \{u \in H^2(A) : u \text{ is radially symmetric and } Qu\big|_{\partial A} = 0 \text{ in trace sense}\}$, and $c = \kappa/m$ as in Section 2. Then there holds

$$\int_A (Bu)v\,dx\,dy = \int_{R_1}^{R_2} \Big(\partial_r u\,\partial_r v + cuv\Big)\,dr + \beta_Q \quad (u, v \in H^2_{rad}(A)),$$

where $\beta_Q = 0$ if $\beta = 0$ and $\beta_Q = \frac{\alpha}{\beta}\sum_{i=1}^2 u(R_m)v(R_m)$ if $\beta > 0$. Hence $H_B = H^1_Q(A) \subset\{ u \in H^1_0(A)$: u is radially symmetric $\}$. Further, we have

$$\int_A \hat{T}_i(u)v\,dxdy = \int_{R_1}^{R_2} \Big(\hat{a}_i\partial_r u\partial_r v + \sum_{|j|\le s_i} \hat{c}_j^{(i)} u_1^{j_1}\dots u_n^{j_n}\,v\Big)\,dr + \tau_Q^{(i)},$$

where $\tau_Q^{(i)} = 0$ if $\beta = 0$ and $\tau_Q^{(i)} = \frac{\alpha}{\beta}\sum_{i=1}^2 \hat{a}_i(R_m)u(R_m)v(R_m)$ if $\beta > 0$.

Consequently, Theorem 2.1 applies in the setting $N = 1$ on (R_1, R_2). That is, (5.3) has a unique weak solution which is radial: $u^* = (u_1^*, \dots, u_n^*) \in H^1_Q(A)^n$. Further, for any $u_0 \in H^2_{rad}(A)^n$ the sequence (2.12) converges linearly to u^*. The auxiliary problem (2.12) is equivalent to

$$-\partial_r^2 z_i^k = \hat{g}_i^k\,, \quad (\hat{Q}_m u_i)\big|_{r=R_m} = 0$$

where $\hat{g}_i^k \equiv -\partial_r(\hat{a}_i\partial_r u_i^k) + \sum_{|j|\le s_i} \hat{c}_j^{(i)}(u_1^k)^{j_1}\dots(u_n^k)^{j_n} - \hat{g}_i$.

If $\alpha \not\equiv 0$ then $\hat{a}_i$, $\hat{c}_j^{(i)}$ and $\hat{g}_i$ are replaced by approximating algebraic polynomials. Then $\hat{T}_i$ are invariant on $\mathcal{P}_{rad} \equiv\{$ algebraic polynomials of $r\}$. For any $h \in \mathcal{P}_{rad}$ the solution of

$$-\partial_r^2 z = h, \quad (\hat{Q}_m z)\big|_{r=R_m} = 0$$

also fulfils $z \in \mathcal{P}_{rad}$ and is elementary to determine. Namely, if $h(r) = \sum_{j=0}^l a_j(r - R_1)^j$, then

$$z(r) = -\sum_{j=0}^{l} \frac{a_j}{(j+1)(j+2)}(r - R_1)^{j+2} + c(r - R_1) + d$$

where the constants c and d are determined from the boundary condition. In the case $\alpha \equiv 0$ the iteration is kept in the class of cosine-polynomials of r on (R_1, R_2) and the auxiliary problems are solved as in subsection (a) with Neumann conditions.

(d) Other domains. The methods given in paragraph (a) for a rectangle can be extended for other domains where the eigenfunctions of the Laplacian are known explicitly. Then the terms of the polynomials in $\mathcal{P}_s$, $\mathcal{P}_c$ and $\mathcal{P}_{sc}$ have to be replaced by the actual eigenfunctions. These are known e.g. for rectangular triangles and regular hexagons [25, 31].

(e) Domain transformation. If two domains are diffeomorphic, then (5.1) can be transformed from one to the other such that uniform monotonicity is preserved. The formulation of this property is restricted to the Dirichlet problem. Further, for simplicity of notation it is done for a single equation.

Proposition 5.2. *Let S, $\Omega \subset \mathbb{R}^N$ be bounded domains, $\Phi \in C^1(\overline{S}, \overline{\Omega})$, $\det \Phi'(x) \neq 0$ $(x \in \overline{S})$. Let $F : H_0^1(\Omega) \to H_0^1(\Omega)$ be given by*

$$\langle F(u), v\rangle_{H_0^1(\Omega)} = \int_\Omega \big(a(x)\,\nabla u \cdot \nabla v + \sum_{j=0}^{s} c_j(x) u^j v\big)\,dx \quad (u, v \in H_0^1(\Omega))$$

where a and c_j fulfil conditions (C1)-(C4), which means here that $a \in C^1(\overline{\Omega})$, $c_j \in C(\overline{\Omega})$, $a(x) \geq m > 0$, $0 \leq \sum_{j=0}^{s} j c_j(x) u^{j-1}$, $s \leq p-1$ with p in (C4). (That is, F is the generalized operator corresponding to some differential operator T of the studied kind.) Let $\tilde{u} \equiv u \circ \Phi$ $(u \in L^2(\Omega))$. Then the operator $\tilde{F} : H_0^1(S) \to H_0^1(S)$, defined by

$$\langle \tilde{F}(\tilde{u}), \tilde{v}\rangle_{H_0^1(S)} = \langle F(u), v\rangle_{H_0^1(\Omega)} \quad (u, v \in H_0^1(\Omega)),$$

fulfils

$$m_1 \|\tilde{h}\|^2_{H_0^1(S)} \leq \langle \tilde{F}'(\tilde{u})\tilde{h}, \tilde{h}\rangle_{H_0^1(S)} \leq M_1(\tilde{u}) \|\tilde{h}\|^2_{H_0^1(S)} \quad (\tilde{u}, \tilde{h} \in H_0^1(S))$$

with suitable constants m_1 and $M_1(\tilde{u})$, the latter depending on $\tilde{u}$.

Proof. Setting $\tilde{a} = (a \circ \Phi) \det \Phi'$ and $\tilde{c}_j = (c_j \circ \Phi) \det \Phi'$, we have

$$\langle \tilde{F}(\tilde{u}), \tilde{v}\rangle_{H_0^1(S)} = \int_S \big(\tilde{a}\,\widetilde{\nabla u} \cdot \widetilde{\nabla v} + \sum_{j=0}^{s} \tilde{c}_j \tilde{u}^j \tilde{v}\big)\,dx\,.$$

Then there holds

$$\langle \tilde{F}'(\tilde{u})\tilde{h}, \tilde{h}\rangle_{H_0^1(S)} = \int_S \big(\tilde{a}|\widetilde{\nabla h}|^2 + \sum_{j=0}^{s} j \tilde{c}_j \tilde{u}^{j-1} \tilde{h}^2\big)\,dx\,.$$

Using $0 < \min_{\overline{\Omega}} \det \Phi'$ and $\max_{\overline{\Omega}} \det \Phi' < +\infty$, it is easy to verify that $\tilde{F}'(\tilde{u})$ inherits uniform ellipticity from $F'(u)$. The details of calculations are left to the reader. □

If $\Phi \in C^2$, then, due to the smoothness conditions and the ellipticity result of Proposition 5.2, the differential operator $\tilde{T}$ that the generalized operator $\tilde{F}$ corresponds to inherits the properties of the original operator T. Consequently, if Ω is C^2-diffeomorphic to one of the above special domains (say S), then transforming the equation $T(u) = g$ from Ω to S via Φ leads to the described direct realization.

Finally, we note that the method described in this section also works for the general form of our system (2.4) if the nonlinearity $f(x, u)$ can be suitably approximated by polynomials, leading to the form (5.1).

6. Examples

We consider two model problems on the domain $I = [0, \pi]^2 \subset \mathbb{R}^2$: (1) a single equation with Dirichlet boundary conditions; (2) a system of two equations with mixed boundary conditions.

According to section 5 (a), the operations required for the numerical realization are elementary. Namely, the terms of the iteration are trigonometric polynomials. The iteration simply consists of executing multiplications and additions of the polynomials, further (for $-\Delta^{-1}$), linear combination of the form (5.2). Storing the polynomials as matrices of the coefficients, the algorithmization is straightforward.

General considerations and the algorithm. There is one more step to be inserted in the realization. Namely, the exact execution of the iteration yields exponentially growing degree of the polynomials. Trying to avoid this for obvious memory reasons, it turns out that the amount of terms which would cause the storing problem consists of essentially useless (almost 0) terms (the high-index coefficients of the stored polynomials), and most of them can be neglected with small prescribed change of accuracy. We formulate this below generally and conclude at the final algorithm for our example using a simple truncation procedure. (The idea may be put through to the other special domains.)

Definition 6.1. Let λ_{kl} and u_{kl} $(k,l \in \mathbb{N}^+)$ be the eigenvalues and eigenfunctions (normed in $L^2(\Omega)$) of the auxiliary problem

$$-\Delta u + cu = \lambda u, \quad Qu\big|_{\partial\Omega} = 0\,.$$

Let $u \in H^1_Q(\Omega)$ be fixed, $u = \sum_{k,l=1}^{\infty} a_{kl}u_{kl}$. Then

(a) for any $s \in \mathbb{N}^+$ we define $u_{[s]} \equiv \sum_{k+l\le s} a_{kl}u_{kl}$;

(b) for any $0 < \omega < \|u\|_{H^1_Q}$ the index $k_{u,\omega} \in \mathbb{N}^+$ is defined by the inequalities

$$\sum_{k+l>k_{u,\omega}} \lambda_{kl}a_{kl}^2 \le \omega^2, \qquad \sum_{k+l\ge k_{u,\omega}} \lambda_{kl}a_{kl}^2 > \omega^2. \tag{6.1}$$

Since $\|u\|^2_{H^1_Q} = \sum_{k+l\ge 1} \lambda_{kl}a_{kl}^2$, the index $k_{u,\omega} \in \mathbb{N}^+$ is the smallest one for which

$$\|u - u_{[k_{u,\omega}]}\|_{H^1_Q} \le \omega\,. \tag{6.2}$$

If the series of u consists of finitely many terms itself, then we can also define $k_{u,\omega}$ via (6.1) for $\omega = 0$.

Remark 6.2. In the considered examples we have $\lambda_{kl} = k^2 + l^2$.

Proposition 6.3. *Let $0 < \omega < \|u\|_{H^1_Q}$ be fixed, $u \in H^1_Q(\Omega)$, $(u^n)_{n\in\mathbb{N}} \subset H^1_Q(\Omega)$. If $u^n \to u$ in $H^1_Q(\Omega)$ then the sequence $(k_{u^n,\omega})$ is bounded.*

Proof. Let $n_0 \in \mathbb{N}^+$ such that $\|u^n - u\|_{H^1_Q(\Omega)} \le \frac{\omega}{2}$ $(n \ge n_0)$. Let $u = \sum_{k,l=1}^{\infty} a_{kl}u_{kl}$, $u^n = \sum_{k,l=1}^{\infty} a^n_{kl}u_{kl}$, and $k_2 = k_{u,\omega/2}$. Then for $n \ge n_0$ we have

$$\begin{aligned}\sum_{k+l>k_2} \lambda_{kl}(a^n_{kl})^2 &\le 2\Big(\sum_{k+l>k_2} \lambda_{kl}(a^n_{kl} - a_{kl})^2 + \sum_{k+l>k_2} \lambda_{kl}a_{kl}^2\Big)\\ &\le 2\left((\omega/2)^2 + (\omega/2)^2\right) = \omega^2.\end{aligned}$$

Hence $k_{u^n,\omega} \le k_2$ $(n \ge n_0)$. □

Note that the property holds similarly in product space. This proposition means that we can consider bounded number of terms in the series of u^n to attain u within a prescribed error. This motivates the following truncation procedure in the realization of the gradient method for our model problems on I.

The algorithm. Let the coefficients of T_i and the right sides g_i be trigonometric polynomials as in section 5 (a) (after suitable approximation). The sequence $u^n = (u^n_1, \dots, u^n_r) \in H^1_Q(I)^r$ is constructed as follows.

Let $u^0 = (0, \dots, 0)$. If, for $n \in \mathbb{N}$, u^n is obtained, then for $i = 1, \dots, r$ we define

$$\begin{aligned} &g_i^n = T_i(u^n) - g_i, \\ &z_i^n \quad \text{is the solution of} \quad -\Delta z_i^n = g_i^n, \quad Qz_i^n\big|_{\partial I} = 0; \\ &\text{we fix } \omega_n > 0, \\ &\tilde{z}_i^n = (z_i^n)_{[k_{z_i^n,\omega_n}]}, \\ &u_i^{n+1} = u_i^n - \frac{2}{M+m}\tilde{z}_i^n . \end{aligned} \tag{6.3}$$

Proposition 6.4. *Let $\varepsilon > 0$, $\omega = r^{-1/2} m\varepsilon$ (where m is the lower bound of T), $(\omega_n) \subset \mathbb{R}^+$, $\omega_n = \omega$ $(n \ge n_0)$. Then there exists $c > 0$ such that*

$$\|u^n - u^*\|_{H_Q^1(I)^r} \le \varepsilon + c\Big(\frac{M-m}{M+m}\Big)^n \quad (n \in \mathbb{N}). \tag{6.4}$$

Proof. Estimate (6.2) implies $\|z^n - \tilde{z}^n\|_{H_Q^1(I)^r} \le r^{1/2}\omega$ $(n \ge n_0)$, hence we can apply Corollary 4.4 with initial guess u^{n_0}. □

In each step we calculate the residual errors

$$r_i^n = \|T_i(u^n) - g_i\|_{L^2(I)} .$$

Since $T_i(u^n)$ and g_i are trigonometric polynomials, this only requires square summation of the coefficients. Letting

$$e_i^n = \frac{1}{m\sqrt{\lambda}} r_i^n ,$$

we then obtain

$$e^n \equiv \|u^n - u^*\|_{H_Q^1(I)^r} \le \Big(\sum_{i=1}^r (e_i^n)^2\Big)^{1/2}. \tag{6.5}$$

Remark 6.5. (a) Since z_i^n is a trigonometric polynomial, therefore (writing $z_i^n = \sum_{k+l\le s_n} \zeta_{kl} a_{kl}$ and $k_n = k_{z_i^n,\omega_n}$) (6.1) means a condition on finite sums:

$$\sum_{k_n < k+l \le s_n} (k^2+l^2)\zeta_{kl}^2 \le \omega^2, \qquad \sum_{k_n \le k+l \le s_n} (k^2+l^2)\zeta_{kl}^2 > \omega^2,$$

i.e. $\tilde{z}_i^n$ is determined by calculating appropriate square sums of cross-diagonals in the coefficient matrices.
(b) The choice of the constants ω_n is influenced by the following factors. The eventual values of ω_n (i.e. for large n) are determined by the required accuracy in virtue of Proposition 6.4. To keep the matrix sizes low, we choose ω_n large as long as possible. Based on (6.4), the necessity for decreasing ω_n arises when the error e^n (calculated by (6.5)) ceases to follow the theoretical linear convergence $(\frac{M-m}{M+m})^n$.

The truncation of the polynomials can be analoguously inserted in the algorithm in the case of other special domains.

Example 6.6. We consider the single Dirichlet problem

$$\begin{aligned} -\Delta u + u^3 &= \sin x \sin y \\ u\big|_{\partial I} &= 0. \end{aligned}$$

Since $g(x,y) = \sin x \sin y \in \mathcal{P}_s$, therefore, defining $u^0 = 0$, we have $(u^n) \subset \mathcal{P}_s$. According to the previous sections, the realization of the gradient method means

elementary matrix operations for the multiplication and addition of sine polynomials (stored as matrices of coefficients), further, linear combination of the form (5.2) for solving the auxiliary problems

$$-\Delta z^n = g^n$$
$$z^n\big|_{\partial I} = 0$$

(with $g^n = -\Delta u^n + (u^n)^3 - g$).

The calculations are made up to accuracy 10^{-4}. First the constants m and M for (6.3) are determined from (2.9) in Theorem 2.1. We now have $m = m' = 1$, $\eta = 0$, $p = 4$, $\kappa = \kappa' = 0$, $\gamma = 3$, $\lambda_1 = 1^2 + 1^2 = 2$, $K_{4,I} \le 1$ from Corollary 2.18 (a), $\mu_4 = 1$, and $\|g\|_{L^2(I)} = \frac{\pi}{2}$. Hence we have

$$M = 4.7011 \quad \text{and thus} \quad \frac{2}{M+m} = 0.3508.$$

The theoretical convergence quotient is

$$\frac{M-m}{M+m} = 0.6492\,.$$

The algorithm (6.3) has been performed in MATLAB, which is convenient for the matrix operations. The following table contains the error e^n (see (6.5)) versus the number of steps n. The truncation constant is $\omega_n \equiv \omega = 0.0005$ throughout the iteration.

step n	1	2	3	4	5	6	7	8
error e^n	1.1107	0.7186	0.4555	0.2821	0.1717	0.1034	0.0620	0.0375

step n	9	10	11	12	13	14	15	16
error e^n	0.0231	0.0149	0.0093	0.0058	0.0036	0.0022	0.0014	0.0009

step n	17	18	19	20	21	22
error e^n	0.0006	0.0004	0.0003	0.0002	0.0002	0.0001

Example 6.7. We consider the system with mixed boundary conditions

$$\begin{aligned} -\Delta u + u - v + u^3 &= g_1(x,y) \\ -\Delta v + v - u + v^3 &= \quad 0 \\ u\big|_{\Gamma_1} = v\big|_{\Gamma_1} = 0, \partial_\nu u\big|_{\Gamma_2} &= \partial_\nu v\big|_{\Gamma_2} = 0 \end{aligned} \tag{6.6}$$

where

$$g_1(x,y) = \frac{\sin x \cos y}{(2 - 0.249\cos 2x)(2 - 0.249\cos 2y)}$$

and $\Gamma_1 = \{0,\pi\}\times[0,\pi]$, $\Gamma_2 = [0,\pi]\times\{0,\pi\}$. Now the eigenfunctions are $\sin kx \cos ly$ which are in $\mathcal{P}_{sc}$; hence g_1 is approximated by $\tilde{g}_1 \in \mathcal{P}_{sc}$. Calculating up to accuracy 10^{-4}, we define the Fourier partial sum

$$\tilde{g}_1(x,y) = \sum_{\substack{k,l \, areodd \\ k+l\le 6}} a_{kl} \sin kx \cos ly\,, \quad a_{kl} = 4.0469 \cdot 4^{-(k+l)}$$

which fulfils $\|g_1 - \tilde{g}_1\|_{L^2(I)} \le 0.0001$. Replacing g_1 by $\tilde{g}_1$ in (6.6), Proposition 5.1 yields

$$\|\tilde{u} - u^*\|_{H^1_Q(I)^2} \le 0.0001,$$

where $u^* = (u_1^*, u_2^*)$ and $\tilde{u} = (\tilde{u}_1, \tilde{u}_2)$ are the solution s of the original and approximated systems, respectively.

Defining $u^0 = v^0 = 0$, we have $(u^n) \subset \mathcal{P}_{sc}$ and $(v^n) \subset \mathcal{P}_{sc}$. The iteration of the gradient method is realized through elementary matrix operations in an analogous way to Example 1. The calculations are made up to accuracy 10^{-4} again.

The constants for (2.9) are determined as follows. We have $m = m' = 1$, $\eta = 0$, $p = 4$, $\lambda_1 = 1^2 + 0^2 = 1$, $\mu_4 = 1$ from the given functions. Further, the eigenvalue s of the Jacobian of $f(u, v) = (u - v + u^3, v - u + v^3)$ are between 0 and $2 + 3(u^2 + v^2)$, hence $\kappa' = 2$ and $\gamma = 3$. Besides, we can use Lemmas 2.1–2.2: the boundary condition gives $K_{2,\Gamma_1} = 0$ and Lemma 2.19 yields $K_{2,\Gamma_2} \le 2.1535$, hence by Lemma 2.16 we have $K_{4,I} \le 1.6051$. Finally, $\|\tilde{g}_1\|_{L^2(I)} = 0.3988$. Hence we have

$$M = 6.1420 \quad \text{and thus} \quad \frac{2}{M+m} = 0.2801,$$

the theoretical convergence quotient now being

$$\frac{M-m}{M+m} = 0.7200\,.$$

The following table presents the errors (first given for u^n and v^n, resp.). The truncation constant is $\omega_n \equiv \omega = 0.0005$ up to step 10, then (observing that the error quotient rises over 0.8) $\omega_n \equiv \omega = 0.0001$ is chosen.

step n	1	2	3	4	5	6	7
e_1^n	0.2821	0.1219	0.0627	0.0358	0.0215	0.0132	0.0083
e_2^n	0	0.0532	0.0459	0.0316	0.0204	0.0129	0.0081
$[(e_1^n)^2 + (e_2^n)^2]^{1/2}$	0.2821	0.1330	0.0776	0.0477	0.0297	0.0184	0.0116

step n	8	9	10	11	12	13	14
e_1^n	0.0053	0.0036	0.0027	0.0022	0.0014	0.0009	0.0005
e_2^n	0.0053	0.0038	0.0030	0.0025	0.0015	0.0010	0.0006
$[(e_1^n)^2 + (e_2^n)^2]^{1/2}$	0.0075	0.0052	0.0041	0.0033	0.0020	0.0013	0.0008

step n	15	16	17	18
e_1^n	0.0003	0.0002	0.0001	0.0001
e_2^n	0.0003	0.0002	0.0001	0.0001
$[(e_1^n)^2 + (e_2^n)^2]^{1/2}$	0.0005	0.0003	0.0002	0.0001

References

[1] Adams, R.A., *Sobolev spaces*, Academic Press, New York-London, 1975.

[2] Axelsson, O., *Iterative solution methods,* Cambridge Univ. Press, 1994.

[3] Axelsson, O., Chronopoulos, A.T., *On nonlinear generalized conjugate gradient methods,* Numer. Math. 69 (1994), No. 1, pp. 1-15.

[4] Axelsson, O., Gustafsson, I., *An efficient finite element method for nonlinear diffusion problems,* Bull. GMS, 32 (1991), pp. 45-61.

[5] Axelsson, O., Layton, W., *Iteration methods as discretiz ation procedures,* Lecture Notes in Math. 1457, Springer, 1990.

[6] Börgers, C., Widlund, O. B., On finite element domain imbedding methods, *SIAM J. Numer. Anal.* 27 (1990), no. 4, 963–978.

[7] Burenkov, V.I., Gusakov, V.A., *On exact constants in Sobolev embeddings* (in Russian), Dokl. Akad. Nauk. 294 (1987), No. 6., pp. 1293-1297.

[8] Burenkov, V.I., Gusakov, V.A., *On exact constants in Sobolev embeddings III.*, Proc. Stekl. Inst. Math. 204 (1993), No. 3., pp. 57-67.

[9] Daniel, J.W., *The conjugate gradient method for linear and nonlinear operator equations*, SIAM J. Numer. Anal., 4 (1967), No.1., pp. 10-26.
[10] Dorr, F. W., The direct solution of the discrete Poisson equation on a rectangle, *SIAM Rev.* 12 (1970), 248–263.
[11] Egorov, Yu.V., Shubin, M.A., *Encyclopedia of Mathematical Sciences, Partial Differential Equations I.*, Springer, 1992.
[12] Faragó I., Karátson J., *Numerical solution of nonlinear elliptic problems via preconditioning operators: theory and application.* Advances in Computation, Volume 11, *NOVA Science Publishers*, New York, 2002.
[13] Faragó, I., Karátson, J., The gradient-finite element method for elliptic problems, *Comp. Math. Appl.* **42** (2001), 1043-1053.
[14] Funaro, D., *Polynomial approximation of differential equations,* Lecture Notes in Physics, New Series, Monographs 8, Springer, 1992.
[15] Gajewski, H., Gröger, K., Zacharias, K., *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen,* Akademie-Verlag, Berlin, 1974
[16] Hackbusch, W., *Theorie und Numerik elliptischer Differentialgleichungen*, Teubner, Stuttgart, 1986.
[17] Karátson, J., The gradient method for non-differentiable operators in product Hilbert spaces and applications to elliptic systems of quasilinear differential equations, *J. Appl. Anal.*, 3 (1997) No.2., pp. 205-217.
[18] Karátson, J., The conjugate gradient method for a class of non-differentiable operators, *Annales Univ. Sci. ELTE*, **40** (1997), 121-130.
[19] Karátson J., Gradient method in Sobolev space for nonlocal boundary value problems, *Electron. J. Diff. Eqns.*, Vol. 2000 (2000), No. 51, pp. 1-17.
[20] Karátson J., Faragó I., Variable preconditioning via quasi-Newton methods for nonlinear problems in Hilbert space, *SIAM J. Numer. Anal.* 41 (2003), No. 4, 1242-1262.
[21] Karátson J., Lóczi L., Sobolev gradient preconditioning for the electrostatic potential equation, to appear in *Comp. Math. Appl.*
[22] Kantorovich, L.V., Akilov, G.P., *Functional Analysis,* Pergamon Press, 1982.
[23] Kelley, C.T., *Iterative methods for linear and nonlinear equations,* Frontiers in Appl. Math., SIAM, Philadelphia, 1995.
[24] Lions. P.-L., Pacella, F., Tricario, M.,, *Best constants in Sobolev inequalities for functions vanishing on some part of the boundary*, Indiana Univ. Math. J., 37 (1988), No. 2., pp. 301-324.
[25] Makai, E., Complete systems of eigenfunctions of the wave equation in some special case, *Studia Sci. Math. Hung.,* 11 (1976), 139–144.
[26] Ortega, J.M., Rheinboldt, W.C., *Iterative solution of nonlinear equations in several variables,* Academic Press, 1970.
[27] Necas, J., *Introduction to the Theory of Elliptic Equations,* J.Wiley and Sons, 1986.
[28] Neuberger, J. W., Renka, R. J., *Sobolev gradients and the Ginzburg-Landau functional,* SIAM J. Sci. Comput. 20 (1999), No. 2, pp. 582–590.
[29] Neuberger, J. W., *Sobolev gradients and differential equations*, Lecture Notes in Math., No. 1670, Springer, 1997.
[30] Neuberger, J. W., *Steepest descent for general systems of linear differential equations in Hilbert space* , Lecture Notes in Math., No. 1032, Springer, 1983.
[31] Riesz F., Sz.-Nagy B., *Vorlesungen über Funktionalanalysis*, Verlag H. Deutsch, 1982.
[32] Rosen, G., *Minimum value for c in the Sobolev inequality* $\|\phi^3\| \le c\|\nabla\phi\|^3$, SIAM J. Appl. Math., 21 (1971), pp. 30-32.
[33] Swarztrauber, P. N., The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle, *SIAM Rev.* 19 (1977), no. 3, 490–501.
[34] Talenti, G., *Best constants in Sobolev inequality*, Ann. Mat. Pura Appl., Ser. 4 (1976), Vol. 110, pp. 353-372.

Department of Applied Analysis, ELTE University, H-1518 Budapest, Hungary
E-mail address: karatson@cs.elte.hu