Sobolev gradient preconditioning for elliptic reaction–diffusion problems with some nonsmooth nonlinearities

J. Karátson[1]

Abstract

The Sobolev gradient approach is an efficient way to construct preconditioned iterations for solving nonlinear problems. We extend this technique to be applicable for elliptic equations describing stationary states of reaction–diffusion problems if the nonlinearities have certain lack of differentiability. We derive convergence results of the Sobolev gradient method on an abstract level and then for our elliptic problem under different assumptions. Numerical tests show convergence as expected.

Keywords. Sobolev gradients, iteration, nonlinear elliptic equation

1 Introduction

Reaction–diffusion problems arise in various nonlinear models in applied mathematics, see, e.g., [7, 14] and the references there. It is often of interest to determine stationary states, which are described by elliptic problems. In order to solve numerically the arising nonlinear problems, an efficient approach is the Sobolev gradient method. The foundations and various contexts of the Sobolev gradient method have been described in [15, 16], for problems with potential see also the author's works [4, 9]. The main idea is that the gradient w.r.t. the Sobolev inner product provides a properly preconditioned iteration. Sobolev gradients have been succesfully used in many applications in the recent decade, such as image processing, Burgers' and Navier-Stokes equations, differential-algebraic equations, Gross-Pitaevskii equations and Ginzburg-Landau functionals, see [11, 12, 17, 18, 19, 20, 21]. Compared to Newton-like methods, which require less iterations, the Sobolev gradient approach often proves to be be still more favourable due to the simpler linearized problems (that, moreover, do not vary in course of the iteration process) and since it is directly fitted to minimization settings. A systematic comparison on a model problem has been executed in [13].

The standard convergence results for such iterations rely on suitable smoothness of the nonlinearity that describes the chemical reaction in the elliptic equation. In this paper we study problems where this smoothness condition is relaxed, and hence the above results cannot be applied. In general, we consider semilinear elliptic boundary value problems of the following form:

$$\left\{ \begin{aligned} -\mathrm{div}\big(k\,\nabla u\big) + q(x,u) &= g, \\ u_{|\partial\Omega} &= 0, \end{aligned} \right. \tag{1.1}$$

with a diffusion coefficient $k \in L^\infty(\Omega)$, $k \geq m > 0$, and with a continuous nonlinearity q such that $\xi \mapsto q(x,\xi)$ is increasing for any $x \in \Omega$. Later we will impose further technical restrictions on q to have well-posedness and then to achieve more concrete estimates on

[1]Department of Applied Analysis & MTA-ELTE Numerical Analysis and Large Networks Research Group, ELTE University; Department of Analysis, Technical University; Budapest, Hungary

the order of convergence of Sobolev gradients. However, these assumptions will still allow nonsmoothness of q. A typical example of such a nonlinearity is $q(x,u) \equiv q(u) = u^\gamma$ for $u > 0$ (or rewritten as $|u|^{\gamma-1}u$ for any u) with some exponent $0 < \gamma < 1$, see, e.g., [3]. Further details on this example will be mentioned later.

Note that for such nonlinear problems it is not possible to apply Newton's method. Moreover, semismooth Newton methods are not applicable either for such non-Lipschitz functions, since semismoothness requires at least local Lipschitz continuity [8] and is thus typically used for nonlinearities like e.g. $\max\{u, 0\}$.

As mentioned above, the Sobolev gradient method provides a properly preconditioned iteration by taking gradients w.r.t. the Sobolev inner product. An elegant property of this approach is that both the construction and the study of convergence can be carried out in the Sobolev space associated to the PDE, that is, on the continuous level. Then one can readily derive the analogous results for finite element discretizations using a proper projection into the FEM subspace, moreover, convergence is typically mesh-independent, see [4, 9]. We follow this vein in this paper, and we focus on the continuous level in the Sobolev space $H_0^1(\Omega)$. After a proper foundation of the problem and of underlying Hilbert space techniques, we derive convergence results of the Sobolev gradient method for our elliptic problem under different assumptions. Thus we obtain generalizations of the existing Sobolev gradient results.

2 The problem and its well-posedness

Let $\Omega \subset \mathbf{R}^d$ be a bounded domain, and let problem (1.1) satisfy the following

Assumptions 2.1.

(i) $k \in L^\infty(\Omega)$, $k \geq m > 0$;

(ii) $q : \overline{\Omega} \times \mathbf{R} \to \mathbf{R}$ is continuous;

(iii) $\xi \mapsto q(x,\xi)$ is increasing for any $x \in \Omega$, further, there exists a number $p \geq 1$ (if $d = 2$) or $1 \leq p \leq \frac{2d}{d-2}$ (if $d > 2$) such that

$$|q(x,\xi)| \leq c_1 + c_2|\xi|^{p-1} \qquad (\forall x \in \Omega,\ \xi \in \mathbf{R}); \tag{2.1}$$

(iv) $g \in L^2(\Omega)$.

The bound for the exponent p in (2.1) ensures the following Sobolev embedding, which will be required for the well-posedness:

$$H_0^1(\Omega) \subset L^p(\Omega), \qquad \|v\|_{L^p} \leq C_p \|v\|_{H_0^1} \quad (\forall v \in H_0^1(\Omega)) \tag{2.2}$$

for some constant $C_p > 0$ independent of v, see [1, Theorem 5.4]. The weak form of the problem reads in a usual way as follows: find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \big(k\,\nabla u \cdot \nabla v + q(x,u)v\big) = \int_\Omega gv \qquad (\forall v \in H_0^1(\Omega)). \tag{2.3}$$

Proposition 2.1 *Problem* (1.1) *has a unique weak solution.*

PROOF. Let us consider the Hilbert space $H_0^1(\Omega)$ with inner product

$$\langle u, v\rangle_{H_0^1} := \int_\Omega \nabla u \cdot \nabla v, \tag{2.4}$$

and the functional $\phi : H_0^1(\Omega) \to \mathbf{R}$,

$$\phi(u) := \int_\Omega \big(\tfrac{k}{2}\,|\nabla u|^2 + Q(x,u) - gu\big),$$

where Q is a potential of q w.r.t. ξ, i.e. $Q : \overline{\Omega} \times \mathbf{R} \to \mathbf{R}$ satisfies $\partial_\xi Q(x,\xi) = q(x,\xi)$ for all x, ξ. The estimate (2.1) implies a bound $|Q(x,u)| \le \tilde{c_1} + \tilde{c_2}|u|^p$, hence the embedding $H_0^1(\Omega) \subset L^p(\Omega)$ ensures that the integral in $\phi(u)$ is finite. Using standard techniques, one can check the Gateaux differentiability of ϕ: namely, for any $u, v \in H_0^1(\Omega)$, there exists

$$\langle \phi'(u), v\rangle_{H_0^1} = \lim_{t\to 0} \tfrac{1}{t}\big(\phi(u+tv)-\phi(u)\big) = \int_\Omega \big(k\,\nabla u\cdot\nabla v - gv\,\big) + \lim_{t\to 0}\int_\Omega \big(\,t\,\tfrac{k}{2}\,|\nabla v|^2 + \partial_\xi Q(x, u+t\theta v)v\,\big)$$

$$= \int_\Omega \big(k\,\nabla u \cdot \nabla v + q(x,u)v\ - gv\big), \tag{2.5}$$

where the last step follows because the integrand converges a.e. as $t \to 0$ and is majorized owing to (2.1). The functional ϕ has the following properties. It is uniformly convex, since Q is convex by the monotonicity of q. Further, the convexity of Q also implies $Q(x,\xi) \ge Q(x,0) + \partial_\xi Q(x,0)\xi = q(x,0)\xi$ if we choose Q such that $Q(x,0) \equiv 0$. Hence

$$\phi(u) \ge \int_\Omega \big(\tfrac{k}{2}\,|\nabla u|^2 + q(x,0)u - gu\big) \ge \tfrac{m}{2}\|u\|_{H_0^1}^2 - (\|q(x,0)\|_{L^2} + \|g\|_{L^2})\|u\|_{L^2}$$

$$\ge \tfrac{m}{2}\|u\|_{H_0^1}^2 - c\|u\|_{H_0^1} \ \to\ +\infty \quad \text{as} \quad \|u\|_{H_0^1} \to +\infty\,.$$

As is well-known (see, e.g., [25, Theorem 40.1]), these properties imply that ϕ has a unique minimizer, which is also its unique critical point, that is, where $\langle \phi'(u), v\rangle_{H_0^1} = 0$ for all $v \in H_0^1(\Omega)$. In virtue of (2.3) and (2.5), this minimizer is the unique weak solution of our problem. ∎

Example. A typical example for a reaction-diffusion problem with nonsmooth nonlinearity is

$$\left\{ \begin{array}{rcl} -\Delta u + u^\gamma & = & g, \\ u_{|\partial\Omega} & = & 0, \end{array} \right. \tag{2.6}$$

where $0 < \gamma < 1$ is the order of the reaction, see, e.g., [3]. The increasing function $u \mapsto u^\gamma$ describes an autocatalytic (endothermic) reaction, defined for $u \ge 0$. One can extend this function in an increasing manner as $u \mapsto |u|^{\gamma-1}u$ for any $u \in \mathbf{R}$, and for source functions $g \ge 0$ the maximum–minimum principle ensures that the solution of the problem

$$\left\{ \begin{array}{rcl} -\Delta u + |u|^{\gamma-1}u & = & g, \\ u_{|\partial\Omega} & = & 0 \end{array} \right. \tag{2.7}$$

is nonnegative, hence the solutions of (2.6) and (2.7) coincide. In this example the non-smoothness of q comes from non-differentiability in one point only ($u = 0$), but even this fact prohibits the applicability of some standard solution techniques. Note that the above nonlinearity is Hölder continuous, which will be a main assumption later to achieve more concrete estimates on convergence. Although this will be a strengthened continuity assumption, it is still general. It allows non-differentiability of q at even infinitely many points, as is the case with Cantor's function (or devil's staircase).

The weak form of our problem can be rewritten as an operator equation as follows. The Riesz repesentation theorem provides an operator $F : H_0^1(\Omega) \to H_0^1(\Omega)$ for which

$$\langle F(u), v\rangle_{H_0^1} = \int_\Omega \big(k\,\nabla u\cdot\nabla v + q(x,u)v\; - gv\big) \qquad (\forall v \in H_0^1(\Omega)). \tag{2.8}$$

Thus (2.3) becomes

$$\langle F(u), v\rangle_{H_0^1} = 0 \qquad (\forall v \in H_0^1(\Omega)),$$

or simply

$$F(u) = 0. \tag{2.9}$$

We wish to solve (2.9) with a steepest descent (or gradient) type method, based on the underlying potential ϕ. For this purpose, we first formulate general results in a Hilbert space setting.

3 Some results on steepest descent iterations in Hilbert space

Let H be a real Hilbert space and $F : H \to H$ be a potential operator, i.e. there exists a Gateaux differentiable functional $\phi : H \to \mathbf{R}$ such that

$$\phi' = F.$$

Further, assume that F is uniformly monotone, i.e. there exists $m > 0$ such that

$$\langle F(v) - F(u),\, v - u\rangle \geq m\,\|v - u\|^2 \qquad (\forall u, v \in H). \tag{3.1}$$

We wish to solve the equation

$$F(u) = 0.$$

The uniformly monotonicity assumption implies that there is a unique solution $u^* \in H$, which coincides with the minimizer of ϕ, see, e.g., [25]. Further, it is not restrictive to have a zero r.h.s., since any equation $F(u) = b$ can be rewritten as $F(u) - b = 0$ with an operator $u \mapsto F(u) - b$ satisfying the same conditions as F.

The steepest descent iteration is a sequence of fixed-point iteration form:

$$u_{n+1} := u_n - \alpha_n\,F(u_n) \tag{3.2}$$

with suitable constants $\alpha_n > 0$. In this case the search direction is $-F(u_n) = -\phi'(u_n)$. We present some convergence results under various continuity conditions on F.

We start with a general theorem, where we assume that F is uniformly continuous in the following sense: there exists a real function $r : \mathbf{R}^+ \to \mathbf{R}^+$ such that

$$r \text{ is continuous and increasing;} \tag{3.3}$$

$$r(0) = 0, \qquad \lim_{t\to+\infty} r(t) = +\infty; \tag{3.4}$$

$$\|F(u) - F(v)\| \le r\left(\|u - v\|\right) \qquad (\forall u, v \in H). \tag{3.5}$$

On the other hand, typically the function r has some simpler special form, such as $r(t) = Lt$ for Lipschitz continuity or $r(t) = Lt^\alpha$ for Hölder continuity. Hence we will later formulate the rate of convergence for these situations.

The following discussion will use the integral and average functions of r, respectively:

$$R(t) := \int_0^t r(s)\,ds\,, \qquad \rho(t) := \frac{1}{t}\int_0^t r(s)\,ds = \frac{R(t)}{t}\,. \tag{3.6}$$

Theorem 3.1 *Let F satisfy conditions* (3.1) *and* (3.3)–(3.5). *Then, for stepsizes*

$$\alpha_n := \tfrac{1}{\|F(u_n)\|}\ \rho^{-1}\bigl(\tfrac{\|F(u_n)\|}{2}\bigr)\,, \tag{3.7}$$

the error of iteration (3.2) *satisfies*

$$e_n := \|u_n - u^*\| \to 0.$$

Proof. The Newton-Leibniz formula yields

$$\phi(u_{n+1}) - \phi(u_n) = \int_0^1 \langle F\bigl(u_n + t(u_{n+1} - u_n)\bigr), u_{n+1} - u_n\rangle dt\,.$$

Subtracting $\langle F(u_n), u_{n+1} - u_n\rangle$, using (3.5) and (3.6), we obtain

$$\phi(u_{n+1}) - \phi(u_n) - \langle F(u_n), u_{n+1} - u_n\rangle \le \int_0^1 r\bigl(t\,\|u_{n+1} - u_n\|\bigr)\,\|u_{n+1} - u_n\|\,dt$$

$$= \int_0^{\|u_{n+1}-u_n\|} r(s)\,ds \;=\; R\bigl(\|u_{n+1} - u_n\|\bigr).$$

Using (3.2), we have

$$\phi(u_{n+1}) - \phi(u_n) \le -\alpha_n\,\|F(u_n)\|^2 + R\bigl(\alpha_n\|F(u_n)\|\bigr)\,.$$

Now let us define α_n as in (3.7). This makes sense since (owing to the assumptions) ρ is increasing and $\lim_{+\infty}\rho = +\infty$. Then

$$R\bigl(\alpha_n\|F(u_n)\|\bigr) = \alpha_n\|F(u_n)\|\,\rho\bigl(\alpha_n\|F(u_n)\|\bigr) = \alpha_n\frac{\|F(u_n)\|^2}{2}\,,$$

hence

$$\phi(u_{n+1}) - \phi(u_n) \le -\alpha_n \frac{\|F(u_n)\|^2}{2} = -\frac{\|F(u_n)\|}{2}\, \rho^{-1}\big(\tfrac{\|F(u_n)\|}{2}\big) =: -\sigma(\|F(u_n)\|),$$

where

$$\sigma(t) := \tfrac{t}{2}\, \rho^{-1}\big(\tfrac{t}{2}\big) \ \text{ is increasing and } \ \sigma(0) = 0.$$

Thus

$$\sum_{n=0}^{\infty} \sigma(\|F(u_n)\|) \le \sum_{n=0}^{\infty} \big(\phi(u_n) - \phi(u_{n+1})\big) = \phi(u_0) - \inf \phi =: D < \infty, \tag{3.8}$$

hence $\sigma(\|F(u_n)\|) \to 0$ and thus $\|F(u_n)\| \to 0$ as $n \to \infty$. Finally, from (3.1),

$$m\, \|u_n - u^*\| \le \|F(u_n) - F(u^*)\| = \|F(u_n)\| \to 0 \tag{3.9}$$

i.e. the theorem is proved. ■

Estimates on the rate of convergence can be given under special choices on r. In each case it suffices to estimate $\|F(u_n)\|$, since one can use (3.9). First we consider Hölder continuity.

Theorem 3.2 *Let F be Hölder continuous, i.e. satisfy condition* (3.5) *with $r(t) := Mt^\gamma$ for some constants $M > 0$ and $0 < \gamma \le 1$, and also satisfy* (3.1). *Then, with stepsizes α_n from* (3.7), *for some constant $c_1 > 0$,*

$$\min_{0 \le k \le n} e_k \le c_1\, n^{-\frac{\gamma}{\gamma+1}} \qquad (\forall n \in \mathbb{N}). \tag{3.10}$$

Proof. We follow [24], where the situation of Hölder continuity was considered for optimization on a finite dimensional space. For the function $r(t) := Mt^\gamma$, one can see that $\sigma(t) = c_0 t^{\frac{\gamma+1}{\gamma}}$ with some $c_0 > 0$. Here

$$c_0 \min_{0 \le k \le n} \|F(u_k)\|^{\frac{\gamma+1}{\gamma}} = \min_{0 \le k \le n} \sigma(\|F(u_k)\|) \le \frac{1}{n} \sum_{k=0}^{n} \sigma(\|F(u_k)\|) \le \frac{D}{n}$$

with D from (3.8), hence (3.9) yields

$$\min_{0 \le k \le n} e_k \le \frac{1}{m} \min_{0 \le k \le n} \|F(u_k)\| \le c_1\, n^{-\frac{\gamma}{\gamma+1}}. \qquad ■$$

Remark 3.1 (i) The stepsize for $r(t) := Mt^\gamma$ can be obtained with an elementary calculation from (3.7) :

$$\alpha_n = \big(\tfrac{\gamma+1}{2M}\big)^{\frac{1}{\gamma}} \|F(u_n)\|^{\frac{1}{\gamma}-1}. \tag{3.11}$$

(ii) For Theorem 3.2 to hold, it suffices to require Hölder continuity of r in a proper neighbourhood of 0, since the proof use this property for arguments that tend to 0. In other words, it suffices to require local Hölder continuity of F: for any $R > 0$ there exists a constant $M > 0$ such that

$$\|F(u) - F(v)\| \le M\|u - v\|^\gamma \qquad (\forall u, v \in H_0^1(\Omega) \ \text{ whenever } \ \|u\|, \|v\| \le R). \tag{3.12}$$

In the case of Lipschitz continuity the above proof (for $\gamma = 1$) would yield convergence of $O(n^{-\frac{1}{2}})$. However, with some further regularity properties, we then have even geometric (i.e., linear) convergence:

Theorem 3.3 *Let F be Lipschitz continuous, i.e. satisfy condition* (3.5) *with $r(t) := Mt$ for some constant $M > 0$, and also satisfy* (3.1). *Assume also that F itself is Gateaux differentiable, the operators $F'(u)$ are self-adjoint, and F' is hemicontinuous (i.e. weakly continuous on segments). Then, using constant stepsizes $\alpha := \frac{2}{M+m}$, denoting $c_1 := \frac{1}{m}\|F(u_0)\|$, we have*

$$e_n \le c_1 \left(\frac{M-m}{M+m}\right)^n \qquad (\forall n \in \mathbb{N}).$$

Proof. Under the assumptions on F', the uniform monotonicity and Lipschitz continuity conditions are equivalent to requiring

$$m\|h\|^2 \le \langle F'(u)h, h\rangle \le M\|h\|^2 \qquad (\forall u, h \in H),$$

in which situation the result is well-known, see, e.g., [4, Theorem 5.4]. ∎

4 Sobolev gradients on continuous and discrete level

Now we study the numerical solution of the boundary value problem (1.1), based on the results of the previous section. As described in the introduction, one may first formulate the iteration on the Sobolev space level. Then the iteration for a finite element discretization is obtained just by a projection into the FEM subspace used.

4.1 Sobolev gradient preconditioning

In our case, where $F = \phi'$, the gradients of ϕ are used to solve equation $F(u) = 0$. The Sobolev gradient iteration is defined by taking gradients w.r.t. the Sobolev inner product. This leads to linearized problems in each step, i.e. a properly preconditioned iteration. The linearized problems depend on the inner product used; we consider two situations. The standard $H_0^1(\Omega)$ inner product leads to auxiliary Poisson equations, which is favourable if there is an efficient Poisson solver to be used. If a weighted inner product is applied in $H_0^1(\Omega)$, then one can simplify the iteration.

Sobolev gradients with the standard $H_0^1(\Omega)$ inner product. Let $u_0 \in H_0^1(\Omega)$ be an arbitrary initial guess. For given u_n ($n \in \mathbb{N}$), the iteration step (3.2) can be written as

$$z_n := F(u_n), \qquad u_{n+1} := u_n - \alpha_n\, z_n\,,$$

where

$$\alpha_n := \left(\tfrac{\gamma+1}{2M}\right)^{\frac{1}{\gamma}} \|F(u_n)\|_{H_0^1}^{\frac{1}{\gamma}-1} \qquad \text{or} \qquad \alpha_n \equiv \alpha := \tfrac{2}{M+m} \tag{4.1}$$

when the conditions of Theorem 3.2 or Theorem 3.3 hold, respectively. Here $z_n := F(u_n)$ is equivalent to

$$\langle z_n, v\rangle_{H_0^1} = \langle F(u_n), v\rangle_{H_0^1} \qquad (\forall v \in H_0^1(\Omega)), \tag{4.2}$$

i.e. $z_n \in H_0^1(\Omega)$ is the function that satisfies

$$\int_\Omega \nabla z_n \cdot \nabla v = \int_\Omega \big(k\,\nabla u_n \cdot \nabla v + q(x,u_n)v \, - gv\big) \qquad (\forall v \in H_0^1(\Omega)). \tag{4.3}$$

That is, altogether, the iteration has the following form:

$$u_{n+1} := u_n - \alpha_n\, z_n\,, \tag{4.4}$$

where z_n is the weak solution of the linear elliptic problem

$$\left\{\begin{array}{rcl} -\Delta z_n &=& -\mathrm{div}\big(k\,\nabla u_n\big) + q(x,u_n) - g, \\ z_{n\,|\partial\Omega} &=& 0. \end{array}\right. \tag{4.5}$$

Weighted Sobolev gradients. Now let us use the weighted inner product

$$\langle u, v\rangle_{H_0^1} := \int_\Omega k\,\nabla u \cdot \nabla v \tag{4.6}$$

adapted to the problem. Then, instead of (4.5), z_n is the weak solution of the linear elliptic problem

$$\left\{\begin{array}{rcl} -\mathrm{div}\big(k\,\nabla z_n\big) &=& -\mathrm{div}\big(k\,\nabla u_n\big) + q(x,u_n) - g, \\ z_{n\,|\partial\Omega} &=& 0. \end{array}\right. \tag{4.7}$$

In this case the iteration can be rewritten in the following simpler form: letting

$$w_n := z_n - u_n\,,$$

if w_n is the weak solution of the linear elliptic problem

$$\left\{\begin{array}{rcl} -\mathrm{div}\big(k\,\nabla w_n\big) &=& q(x,u_n) - g, \\ z_{n\,|\partial\Omega} &=& 0. \end{array}\right. \tag{4.8}$$

then

$$u_{n+1} := (1-\alpha_n)u_n - \alpha_n\, w_n\,. \tag{4.9}$$

Here α_n is from (4.1), understanding the $H_0^1(\Omega)$-norm with weight function k added in (2.4).

4.2 Finite element discretization

The finite element method (FEM) looks for the numerical solution of problem (2.3) in a proper FEM subspace $V_h \subset H_0^1(\Omega)$: find $u \in V_h$ such that

$$\int_\Omega \big(k\,\nabla u \cdot \nabla v + q(x,u)v\,\big) = \int_\Omega gv \qquad (\forall v \in V_h). \tag{4.10}$$

The Sobolev gradient iteration for the FEM problem (4.10) is directly obtained with the projection of (4.3) into V_h. Namely, let $u_0 \in V_h$ be an arbitrary initial guess. For given u_n ($n \in \mathbb{N}$), we let $u_{n+1} := u_n - \alpha_n z_n$, where $z_n \in V_h$ is the function that satisfies

$$\int_\Omega \nabla z_n \cdot \nabla v = \int_\Omega \big(k \nabla u_n \cdot \nabla v + q(x, u_n)v - gv\big) \qquad (\forall v \in V_h), \tag{4.11}$$

i.e. $z_n \in V_h$ is the FEM solution of the linear elliptic problem (4.5), if the standard inner product is used. Finally, for weighted Sobolev gradients, we have the weight function k on the l.h.s. of (4.11), and the iteration can be simplified with w_n in the obvious way.

We note that the estimates are independent of the actual FEM subspace used. This property will also be reflected in the mesh independence obtained in the numerical tests in subsection (5.3).

5 Convergence estimates

5.1 Power order convergence

In this subsection we establish convergence rates under the assumption of Hölder continuity of q, i.e., we impose

Assumption 5.1. The function q is Hölder continuous w.r.t. ξ, i.e., there exist constants $c_q > 0$ and $0 < \gamma < 1$ such that

$$|q(x, \xi) - q(x, \tilde{\xi})| \le c_q |\xi - \tilde{\xi}|^\gamma \qquad (\forall x \in \Omega,\ \xi \in \mathbf{R}). \tag{5.1}$$

Proposition 5.1 *The operator F in* (2.8) *is locally Hölder continuous on $H_0^1(\Omega)$ in the sense of* (3.12).

PROOF. The operator F can be decomposed in a linear and a remaining part as

$$F = L + A,$$

where

$$\langle Lu, v\rangle_{H_0^1} = \int_\Omega k \nabla u \cdot \nabla v, \qquad \langle A(u), v\rangle_{H_0^1} = \int_\Omega \big(q(x, u)v - gv\big) \qquad (\forall v \in H_0^1(\Omega)).$$

First, L is Lipschitz continuous since it is a bounded linear operator:

$$\|Lu\|_{H_0^1} = \sup_{\|z\|_{H_0^1}=1} \langle Lu, z\rangle_{H_0^1} = \sup_{\|z\|_{H_0^1}=1} \int_\Omega k \nabla u \cdot \nabla z \le \|k\|_{L^\infty} \|\nabla u\|_{L^2} = \|k\|_{L^\infty} \|u\|_{H_0^1},$$

hence it is locally Hölder continuous. Second, we prove that A is Hölder continuous. Here

$$\|A(u) - A(v)\|_{H_0^1} = \sup_{\|z\|_{H_0^1}=1} \langle A(u) - A(v), z\rangle_{H_0^1} = \sup_{\|z\|_{H_0^1}=1} \int_\Omega \big(q(x, u) - q(x, v)\big) z$$

$$\leq c_q \sup_{\|z\|_{H_0^1}=1} \int_\Omega |u-v|^\gamma |z| \,.$$

We can first apply Hölder's inequality with the parameters $\frac{\gamma+1}{\gamma}$ and $\gamma+1$, since $\frac{\gamma}{\gamma+1} + \frac{1}{\gamma+1} = 1$, and then the Sobolev embedding

$$\|w\|_{L^{\gamma+1}} \leq C_{\gamma+1} \|w\|_{H_0^1} \tag{5.2}$$

from (2.2), since $\gamma+1 \leq 2 \leq \frac{2d}{d-2}$. Thus we obtain

$$\int_\Omega |u-v|^\gamma |z| \leq \left\| |u-v|^\gamma \right\|_{L^{\frac{\gamma+1}{\gamma}}} \|z\|_{L^{\gamma+1}} = \|u-v\|_{L^{\gamma+1}}^\gamma \|z\|_{L^{\gamma+1}} \leq C_{\gamma+1}^{\gamma+1} \|u-v\|_{H_0^1}^\gamma \|z\|_{H_0^1} \,.$$

The above inequalities yield

$$\|A(u)-A(v)\|_{H_0^1} \leq c_q C_{\gamma+1}^{\gamma+1} \|u-v\|_{H_0^1}^\gamma \tag{5.3}$$

that is, A is Hölder continuous (globally, and hence locally). Altogether, $F = L + A$ is also locally Hölder continuous. ■

Theorem 5.1 *Let Assumptions 2.1 and 5.1 hold. Let us construct the Sobolev gradient iteration, starting from arbitrary initial guess $u_0 \in H_0^1(\Omega)$, according to either* (4.4)–(4.5) *or* (4.8)–(4.9)*, and choose the stepsizes α_n from* (3.11)*. Then, for some constant $c_1 > 0$, the errors $e_k := \|u_k - u^*\|_{H_0^1}$ satisfy*

$$\min_{0\leq k\leq n} e_k \leq c_1 \, n^{-\frac{\gamma}{\gamma+1}} \,.$$

Proof. Proposition 5.1 and Remark 3.1 yield that Theorem 3.2 can be applied for F in $H_0^1(\Omega)$ with inner product (2.4), hence the corresponding Sobolev gradient iteration (4.4)–(4.5) satisfies (3.10). Further, the weighted inner product (4.6) induces a norm equivalent to the original one, hence Proposition 5.1 also holds w.r.t. the weighted norm, and hence the weighted Sobolev gradient iteration (4.8)–(4.9) also satisfies (3.10). ■

Remark 5.1 In order to calculate the stepsize (3.11), we need the values γ, $\|F(u_n)\|_{H_0^1}$ and M.

- Here γ is known since it coincides with the Hölder constant of q from (5.1).
- By construction, $F(u_n) =: z_n$ is computed as the solution of the auxiliary linear problem, hence we have to compute $\|z_n\|_{H_0^1}$ after solving the auxiliary problem and use this value in the stepsize.
- For M we can use any bound with which the Hölder continuity estimate holds. As seen in the proof of Proposition 5.1, such a bound requires an estimate for the

embedding constant $C_{\gamma+1}$ that appears in (5.3). For a simple estimate for this, we can use the intermediate L^2 space, since, as is well-known, for any $\beta \le 2$

$$\|w\|_{L^\beta} \le |\Omega|^{\frac{1}{\beta}-\frac{1}{2}} \|w\|_{L^2}, \qquad \|w\|_{L^2} \le \frac{diam(\Omega)}{d\sqrt{\pi}} \|w\|_{H_0^1},$$

see [4, 22]. Hence, with $\beta = \gamma + 1$, we obtain that (5.2) holds with

$$C_{\gamma+1} \le \tfrac{1-\gamma}{2(1+\gamma)} \frac{diam(\Omega)}{d\sqrt{\pi}}.$$

5.2 Local linear convergence for interior regular problems

In this subsection we study a special subclass of our boundary value problem (1.1). The main assumption is that the nonsmoothness of the nonlinearity (i.e. the singularity of its derivative) is restricted to the argument $\xi = 0$, as formulated below (5.5). As it will turn out, positive source functions imply $u > 0$ inside Ω, hence the singularity of such problems is restricted to the boundary where $u = 0$. This situation still covers the example (2.6).

Let us consider the problem

$$\begin{cases} -\Delta u + q(x,u) &= g, \\ u_{|\partial\Omega} &= 0 \end{cases} \tag{5.4}$$

under the following

Assumptions 5.2.

(i) $\Omega \subset \mathbf{R}^2$ is a bounded domain, and $\partial\Omega \in C^1$;

(ii) $q : \overline{\Omega} \times \mathbf{R} \to \mathbf{R}$ is continuous, further, $q(x,0) = 0$ $(\forall x \in \Omega)$;

(iii) for any $x \in \Omega$, $\xi \mapsto q(x,\xi)$ is increasing and it is C^1 on $\mathbf{R} \setminus \{0\}$, further, there exist constants $0 < \gamma < 1$ and $c_\gamma > 0$ such that

$$|\partial_\xi q(x,\xi)| \le c_\gamma |\xi|^{\gamma-1} \qquad (\forall x \in \Omega,\ \xi \in \mathbf{R} \setminus \{0\}); \tag{5.5}$$

(iv) $g \in C(\Omega)$ and $g > 0$ in Ω.

Remark 5.2 Assumptions 5.2 clearly imply that Assumptions 2.1 and 5.1 hold, i.e. we have a subclass of the previous boundary value problem. Namely, the diffusion coefficient is now $k \equiv 1$, further, using the Newton–Leibniz-formula, (5.5) implies both the Hölder continuity and the growth condition for q.

5.2.1 Properties of the exact solution

The positivity features of the solution will play an important role. For this we fomulate the following property, where $\partial_\nu u$ denotes the outer normal derivative.

Definition 5.1 A function $u : \Omega \to \mathbf{R}$ is called *strongly positive on* Ω if $u \in C^1(\overline{\Omega})$ and it satisfies $u > 0$ on Ω and $\partial_\nu u < 0$ on $\partial\Omega$.

Lemma 5.1 *If u is strongly positive, then* $|u|^{\gamma-1} \in L^r(\Omega)$ *for any* $1 < r < \frac{1}{1-\gamma}$.

PROOF. Since $\partial_\nu u < 0$ is continuous on $\partial\Omega$, there exists $m > 0$ such that $\partial_\nu u \le -m$ on $\partial\Omega$. For given $x \in \Omega$, let $d(x)$ denote the distance of x from $\partial\Omega$, and let $x_0 \in \partial\Omega$ such that $d(x) = |x - x_0|$. For given $\delta > 0$, let $\Omega_\delta := \{x \in \Omega : \ d(x) < \delta\}$. Then there exists $\delta > 0$ such that $u(x) \ge \frac{m}{2}|x - x_0|$ for $x \in \Omega_\delta$. Hence

$$|u(x)|^{r(\gamma-1)} \le c\,|x - x_0|^{r(\gamma-1)} \qquad (\forall x \in \Omega_\delta),$$

and since $r(\gamma-1) > -1$, we have $|u|^{r(\gamma-1)} \in L^1(\Omega_\delta)$, i.e. $|u|^{\gamma-1} \in L^r(\Omega_\delta)$. Further, $u \ge \varepsilon := \min_{\Omega\setminus\Omega_\delta} > 0$ in $\Omega \setminus \Omega_\delta$, thus $|u|^{\gamma-1} \in L^r(\Omega\setminus\Omega_\delta)$. Altogether, $|u|^{\gamma-1} \in L^r(\Omega)$. ∎

Proposition 5.2 *The solution of* (5.4) *is strongly positive.*

PROOF. It follows from [6, Theorem 12.4] that $u \in C^1(\overline{\Omega})$. Now, first we prove that $u > 0$. Under our monotonicity and positivity assumptions, we have $u \ge 0$ (see, e.g., [10]), hence we only have to exclude interior zeros. Assume for contrary that there exists $x_1 \in \Omega$ such that $u(x_1) = 0$. Then x_1 is a local minimizer, hence $\Delta u(x_1) \ge 0$. Further, from assumption (ii), $u(x_1) = 0$ implies $q(x_1,\, u(x_1)) = 0$, and from assumption (iv), $g(x_1) > 0$. These lead to the contradiction

$$0 \ge -\Delta u(x_1) + q(x_1,\, u(x_1)) = g(x_1) > 0.$$

Now we prove $\partial_\nu u < 0$ on $\partial\Omega$. By the strong maximum principle [6, Lemma 3.4], we have $\partial_\nu u(x_0) < 0$ for $x_0 \in \partial\Omega$ whenever $u(x_0) < u(x)$ for all $x \in \Omega$. In our case this holds on the whole $\partial\Omega$, since $u(x_0) = 0$ and $u > 0$ in Ω. ∎

Corollary 5.1 *For any* $1 < r < \frac{1}{1-\gamma}$, *the solution of* (5.4) *satisfies* $|u|^{\gamma-1} \in L^r(\Omega)$. *In other words,*

$$I_r(u) := \int_\Omega |u|^{r(\gamma-1)} \ < \ \infty. \tag{5.6}$$

5.2.2 The local linear convergence result

Based on the above, we are able to provide linear rate of convergence (hence an acceleration if compared with the bound in Theorem 5.1) provided that u_n is close enough to the exact solution to reproduce its positivity properties, formulated in Proposition 5.2 and Corollary 5.1. Namely, let us impose

Assumptions 5.3. There exists $n_0 \in \mathbb{N}$ such that

(i) u_n is strongly positive for any $n \ge n_0$;

(ii) $K_r := \sup\limits_{n\ge n_0} I_{r\,|\,[u_n,\, u_{n+1}]} \ < \ \infty.$

Here we used notation $\sup\limits_{n\ge n_0} I_{r\,|\,[u_n,\, u_{n+1}]} := \sup\limits_{n\ge n_0} \{I_r(u) : \ u = su_n + (1-s)u_{n+1},\ 0 \le s \le 1\}$.

Lemma 5.2 *If* (5.6) *holds, then the Gateaux derivative of* F *satisfies*

$$m\|h\|^2_{H^1_0} \leq \langle F'(u)h,h\rangle_{H^1_0} \leq \Big(\tilde{m} + c_\gamma\, C^2_{2s}\, I_r(u)^{\frac{1}{r}}\Big)\, \|h\|^2_{H^1_0} \qquad (\forall h \in H^1_0(\Omega)).$$

PROOF. It follows in a standard way [4] that the Gateaux derivative of F satisfies

$$\langle F'(u)h,v\rangle_{H^1_0} = \int_\Omega \big(k\,\nabla h\cdot\nabla v + \partial_\xi q(x,u)hv\,\big) \qquad (\forall h,v \in H^1_0(\Omega)).$$

To prove the upper bound, let us fix some number $1 < r < \frac{1}{1-\gamma}$. We can use Hölder's inequality with the parameters r and $s := \frac{r}{r-1}$, since $\frac{1}{r} + \frac{r-1}{r} = 1$. Then, with (2.2),

$$\int_\Omega |u|^{\gamma-1}\, h^2 \;\leq \big\||u|^{\gamma-1}\big\|_{L^r}\big\|h^2\big\|_{L^s} = I_r(u)^{\frac{1}{r}}\,\|h\|^2_{L^{2s}} \leq I_r(u)^{\frac{1}{r}}\, C^2_{2s}\|h\|^2_{H^1_0}\,,$$

hence, letting $\tilde{m} := \sup_\Omega k$,

$$\langle F'(u)h,h\rangle_{H^1_0} = \int_\Omega \big(k\,|\nabla h|^2 + \partial_\xi q(x,u)h^2\,\big) \leq \tilde{m}\int_\Omega |\nabla h|^2 + c_\gamma \int_\Omega |u|^{\gamma-1}\,h^2$$

$$\leq \Big(\tilde{m} + c_\gamma\, C^2_{2s}\, I_r(u)^{\frac{1}{r}}\Big)\,\|h\|^2_{H^1_0}\,.$$

The lower bound follows readily: since $\partial_\xi q \geq 0$, we have

$$m\|h\|^2_{H^1_0} = m\int_\Omega |\nabla h|^2 \leq \int_\Omega k\,|\nabla h|^2 \leq \langle F'(u)h,h\rangle_{H^1_0}\,. \qquad \blacksquare$$

Now we are in the position to prove linear convergence, for which we can adapt the techniques of [4, Theorem 5.4] in order to derive a local version of Theorem 3.3.

Theorem 5.2 *Let Assumptions 5.2–5.3 hold. Let us construct the Sobolev gradient iteration as in Theorem 5.1, but for* $n \geq n_0$ *redefine the stepsize as a constant* $\alpha_n \equiv \alpha := \frac{2}{M+m}$, *where* $M := \tilde{m} + c_\gamma\, C^2_{2s}\, K_r^{1/r}$. *Then, denoting* $c_1 := \frac{1}{m}\|F(u_{n_0})\|_{H^1_0}$, *the errors* $e_n := \|u_n - u^*\|_{H^1_0}$ *satisfy*

$$e_n \leq c_1\left(\frac{M-m}{M+m}\right)^{n-n_0} \qquad (\forall n \geq n_0).$$

PROOF. In virtue of the Newton-Leibniz formula, and since $u_{n+1} := u_n - \alpha\, F(u_n)$,

$$F(u_{n+1}) = F(u_n) + \int_0^1 F'\big(u_n + t(u_{n+1}-u_n)\big)\,(u_{n+1}-u_n)\,dt$$

$$= F(u_n) - \alpha\int_0^1 F'\big(u_n + t(u_{n+1}-u_n)\big)\,F(u_n)\,dt \;=: L_n\,F(u_n),$$

where

$$L_n := I - \alpha\int_0^1 F'\big(u_n + t(u_{n+1}-u_n)\big)\,dt$$

is a bounded linear operator on $H_0^1(\Omega)$, which is self-adjoint since the derivatives $F'(u)$ of the potential operator F are self-adjoint. Moreover, for any $u := u_n + t(u_{n+1} - u_n) \in [u_n,\, u_{n+1}]$ and $h \in H_0^1(\Omega)$, Lemma 5.2 and Assumption 5.3 (ii) yield

$$m\|h\|_{H_0^1}^2 \le \langle F'(u)h, h\rangle_{H_0^1} \le M\,\|h\|_{H_0^1}^2 , \qquad \text{where} \quad M := \tilde{m} + c_\gamma\, C_{2s}^2\, K_r^{\frac{1}{r}},$$

hence

$$(1-\alpha M)\|h\|_{H_0^1}^2 \le \langle L_n\, h,\, h\rangle_{H_0^1} \le (1-\alpha m)\|h\|_{H_0^1}^2 .$$

The stepsize $\alpha = \frac{2}{M+m}$ yields $1 - \alpha m = -(1-\alpha M) = \frac{M-m}{M+m}$, hence $\|L_n\| \le \frac{M-m}{M+m}$. Thus

$$\|F(u_{n+1})\|_{H_0^1} \le \tfrac{M-m}{M+m}\,\|F(u_n)\|_{H_0^1} \qquad (\forall n \ge n_0).$$

This and (3.9) altogether yield

$$e_n \le \tfrac{1}{m}\|F(u_n)\|_{H_0^1} \le \tfrac{1}{m}\|F(u_{n_0})\|_{H_0^1} \left(\tfrac{M-m}{M+m}\right)^{n-n_0} \qquad (\forall n \ge n_0). \qquad \blacksquare$$

5.3 Numerical experiments

We consider the stationary reaction-diffusion problem (2.6) with $\gamma = 1/2$ and $g \equiv 1$:

$$\begin{cases} -\Delta u + u^{1/2} &= 1, \\ u_{|\partial\Omega} &= 0. \end{cases} \tag{5.7}$$

For instance, such an equation with exponent 1/2 can describe the concentration u of chlorine in a gas phase reaction with chloroform if the latter has a larger magnitude considered as constant [23].

We used Courant finite elements on the unit square domain. For this problem, owing to the Laplacian principal part, the standard and weighted Sobolev gradients (4.5) and (4.7) coincide. We have run the iteration with three mesh parameters h, and calculated the residual errors and convergence quotients

$$\varepsilon_n := \|F(u_n)\|_{H_0^1} \qquad \text{and} \qquad Q_n := \frac{\varepsilon_n}{\varepsilon_{n-1}},$$

respectively. Owing to the positivity of the iterates, we have used the simple form $u^{1/2}$ of the nonlinearity in the code instead of $|u|^{-1/2}u$. The results are shown in Table 1.

$h = 0.1$		$h = 0.01$		$h = 0.001$	
ε_n	Q_n	ε_n	Q_n	ε_n	Q_n
0.1814322		0.1874409		0.1874726	
0.0301495	0.1661	0.0302984	0.1617	0.0302977	0.1616
0.0019173	0.0636	0.0018617	0.0614	0.0018612	0.0614
0.0001317	0.0687	0.0001235	0.0663	0.0001234	0.0663
0.0000089	0.0683	0.0000081	0.0660	0.0000081	0.0659
0.0000006	0.0684	0.0000005	0.0661	0.0000005	0.0660

Table 1. The residual errors and convergence quotients for the test problem.

One may observe that the convergence is linear in accordance with subsection 5.2, further, it shows a uniform behaviour independently of the mesh size h.

5.4 Concluding remarks

We have established the rate of convergence of the Sobolev gradient iteration in two situations in the Sobolev space $H_0^1(\Omega)$. For the finite element discretization, the iteration is obtained just by a projection into the FEM subspace. The convergence estimates for the FEM case coincide with the Sobolev space case, since the proofs that use functions in $H_0^1(\Omega)$ can be restricted to functions in V_h only. This approach is similar to the one used in [4, 15].

Besides obtaining convergence of Sobolev gradients, we mention that Newton's method or its semismooth versions might not be used here owing to the lack of Lipschitz continuity.

Finally we note that our results can be generalized in a straightforward way to certain more general situations: to problems with mixed boundary conditions and for gradient systems, i.e. PDE systems arising from the minimization of a joint energy functional. In these cases the Sobolev space $H_0^1(\Omega)$ has to be replaced by a subspace of $H^1(\Omega)$ associated to the given Dirichlet portion of the boundary, or by a product Sobolev space, respectively.

Acknowledgements. This research was carried out in the ELTE Institutional Excellence Program (1783-3/2018/FEKUTSRAT) supported by the Hungarian Ministry of Human Capacities, and supported by the Hungarian Scientific Research Fund OTKA, Grant No. K112157 and SNN125119.

References

[1] Adams, R.A., Fournier, J. F., *Sobolev Spaces*, Second edition, Pure and Applied Mathematics 140, Elsevier/Academic Press, Amsterdam, 2003.

[2] Céa, J. *Optimization - Theory and Algorithms*, Bombay, 1978.

[3] Díaz, J. I., Applications of symmetric rearrangement to certain nonlinear elliptic equations with a free boundary, *Nonlinear differential equations* (Granada, 1984), 155–181, Res. Notes in Math., 132, Pitman, 1985.

[4] 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.

[5] Faragó I., Karátson J., Sobolev gradient type preconditioning for the Saint-Venant model of elasto-plastic torsion, *Int. J. Numer. Anal. Model.* 5 (2008), no. 2, 206–221.

[6] Gilbarg, D., Trudinger, N. S., *Elliptic partial differential equations of second order* (2nd edition), Grundlehren der Mathematischen Wissenschaften 224, Springer, 1983.

[7] Grindrod, P., *The theory and applications of reaction-diffusion equations: patterns and waves*, Second edition, Oxford University Press, New York, 1996.

[8] Mifflin, R., Semismooth and Semiconvex Functions in Constrained Optimization, *SIAM J. Control Optim.* 1977, Vol. 15, No. 6, pp. 959–972.

[9] Karátson, J., Faragó, I., Preconditioning operators and Sobolev gradients for nonlinear elliptic problems, *Comput. Math. Appl.* 50 (2005), no. 7, 1077–1092.

[10] Karátson, J., Korotov, S., Discrete maximum principles for finite element solutions of nonlinear elliptic problems with mixed boundary conditions, *Numer. Math.* 99 (2005), 669–698.

[11] Kazemi, P., Danaila, I., Sobolev gradients and image interpolation, *SIAM J. Imaging Sci.* 5 (2012), no. 2, 601–624.

[12] Kazemi, P., Eckart, M,. Minimizing the Gross-Pitaevskii energy functional with the Sobolev gradient – analytical and numerical results, *Int. J. Comput. Methods* 7 (2010), no. 3, 453–475.

[13] Kovács B., A comparison of some efficient numerical methods for a nonlinear elliptic problem, *Central Eur. J. Math.*, 10 (2012), no. 1, 217–230.

[14] Marciniak-Czochra, A., *Reaction-diffusion models of pattern formation in developmental biology,* Mathematics and life sciences, 191?212, De Gruyter Ser. Math. Life Sci., 1, De Gruyter, Berlin, 2013.

[15] Neuberger, J. W., *Sobolev gradients and differential equations,* Second edition, Lecture Notes in Mathematics, 1670. Springer-Verlag, Berlin, 2010.

[16] Neuberger, J. W., Renka, R. J., Sobolev gradients: introduction, applications, problems. in: *Variational methods: open problems, recent progress, and numerical algorithms*, 85–99, Contemp. Math., 357, Amer. Math. Soc., Providence, RI, 2004.

[17] Raza, N., Sial, S., Neuberger, J. W., Numerical solution of Burgers' equation by the Sobolev gradient method, *Appl. Math. Comput.* 218 (2011), no. 8, 4017–4024.

[18] Nittka, R., Sauter, M., Sobolev gradients for differential algebraic equations, *Electron. J. Diff. Eqns.* 2008, No. 42, 31 pp.

[19] Raza, N., Sial, S., Butt, Asma R., Numerical approximation of time evolution related to Ginzburg-Landau functionals using weighted Sobolev gradients, *Comput. Math. Appl.* 67 (2014), no. 1, 210–216.

[20] Renka, R. J., Nonlinear least squares and Sobolev gradients, *Appl. Numer. Math.* 65 (2013), 91–104.

[21] Renka, R. J., A Sobolev gradient method for treating the steady-state incompressible Navier-Stokes equations, *Cent. Eur. J. Math.* 11 (2013), no. 4, 630–641.

[22] Rudin, W., *Functional analysis* (2nd edition), International Series in Pure and Applied Mathematics, McGraw-Hill, New York, 1991.

[23] Winning, I. H., The kinetics of the photo-chlorination of chloroform vapour, *Trans. Farad. Soc.*, 47 (1951), pp. 1084–1088.

[24] Yashtini, M., On the global convergence rate of the gradient descent method for functions with Hölder continuous gradients, *Optim. Lett.* 10 (2016), no. 6, 1361–1370.

[25] Zeidler, E., *Nonlinear functional analysis and its applications,* Vol. III., Springer, 1986