ANZIAM J. 46(2005), 449–470

A ROBUST LINEARISATION SCHEME FOR A NONLINEAR ELLIPTIC BOUNDARY VALUE PROBLEM: ERROR ESTIMATES

MARIAN SLODIČKA[1]

(Received 28 October, 2002; revised 27 May, 2004)

Abstract

We consider a nonlinear second-order elliptic boundary value problem in a bounded domain $\Omega \subset \mathbb{R}^N$ with mixed boundary conditions. The solution is found via linearisation. We design a robust and efficient approximation scheme. Error estimates for the linearisation algorithm are derived in $L_2(\Omega)$, $H^1(\Omega)$ and $L_\infty(\Omega)$ spaces under the minimal regularity assumptions of the exact solution.

1. Introduction

Linearisation methods have been used in the numerical analysis of nonlinear elliptic boundary value problems (BVPs) for quite a long time. Frequently, essential properties such as differentiability of the nonlinear operator, boundedness and invertibility properties of linearised operators are used. Linearisation methods are powerful tools when analysing the existence and convergence of approximations and many special techniques have been developed to solve these problems.

Many algorithms use a Newton-type linearisation. The classical Newton's method

$$f'(u_{k-1})(u_k - u_{k-1}) = -f(u_{k-1})$$

or its simplified version

$$f'(u_0)(u_k - u_{k-1}) = -f(u_{k-1})$$

converge for Lipschitz continuous f. A major drawback of both algorithms is that the initial guess needs to be near the exact solution, however it is well known that for initial data close enough to the exact solution, Newton's method converges quadratically.

[1]Department of Mathematical Analysis, Ghent University, Galglaan 2, B-9000 Gent, Belgium; e-mail: marian.slodicka@ugent.be.

Another group of approximations is based on the so-called relaxation schemes, one of which is the Jäger-Kačur scheme (see Jäger and Kačur [6, 7] or Kačur [8, 9]). In general, the proof of the convergence of such an algorithm is not an easy matter. The main disadvantage of the proposed relaxation scheme is the fact that the nonlinear function appearing in the equation must be strictly monotonically increasing.

An attractive group of linearisation schemes represents the method of upper and lower solutions (also known as barriers). Examples of such algorithms can be found, for example, in Amann [1], Deng *et al.* [4], Evans [5, page 507] and Pao [13, page 155]. The linearisation of a nonlinear problem relies on the ordering properties of solutions. One defines recursive sequences starting from a sub- and a super-solution, respectively and there exists a solution lying between them. The rates of such monotone convergence cannot be determined in general. This technique is often used in existence proofs, but it has a big disadvantage especially in the computation of evolution problems. Namely, one has to start far away from the real solution and the information from the previous time step cannot be used as the starting point for the approximation scheme. Otherwise it is not possible to prove the monotonicity of iterations. Nevertheless, these schemes create in some sense the basis of our approach. We will of course prove the convergence of iterations although they do not need to be monotone.

The need of a reliable, efficient and robust iteration scheme for the solution of nonlinear elliptic BVPs, which can start from arbitrary initial data, is evident. We propose such an algorithm in this paper. We consider a nonlinear second-order elliptic BVP, where the nonlinearity $\beta(u)$ (u stands for a solution and $\beta = g, g_R$) can appear as a source term in the equation or at the Robin-type boundary condition (BC). In both situations we assume that the function β is monotonically nondecreasing ($\beta' \geq 0$) and we distinguish the Lipschitz continuous ($0 \leq \beta' \leq L$) and the degenerate ($0 \leq \beta' \leq \infty$) case. We follow some ideas from Slodička [15, 17] and we extend these results to the case of unbounded β' taking into account the possible nonlinearity at the boundary. The Lipschitz continuous case has been considered in [15], where the function β could degenerate only in a single point in which it was β regularised by a suitably chosen β_k (k stands for the iteration parameter). Here we do not need such a regularisation and, moreover, we allow β to degenerate in the whole interval. In [17] this regularisation has been removed, but the problem setting there does not contain nonlinear BCs and the convection is independent from the solution, which is taken into account in our paper. The analysis of the mixed finite element discretisation for a Lipschitz continuous case can be found in Slodička [16].

In the Lipschitz continuous case, our algorithm is similar to the scheme proposed by Evans [5, page 507] (where there is a proof of convergence of a monotone approximation for the Dirichlet BVP), but the main difference is that we show the order of convergence of iterations (not necessarily monotone if we do not start from upper and lower solutions) for a more general setting (a nonlinear BC) without using the ordering

property of approximations. Hence we can start from arbitrary data and the iteration scheme will converge to the exact solution. In the degenerate case, we first apply a local regularisation to the nonlinear function β, and then we use a similar linearisation for the regular instance. The argument for convergence is more delicate since up to now there has existed no linearisation scheme for degenerate elliptic BVPs, which converges and which can start from arbitrary data.

The proposed algorithms (3.3) and (4.4) are in their spirit nothing more than an application of the well-known Banach fixed point theorem. We explain the main idea in the following example.

Given a Lipschitz continuous function g satisfying $0 < \gamma \le g' \le L$, we look for a solution x of the equation $g(x) = 0$. Define a function h by $h(s) = s - g(s)/L$, then

$$0 \le h'(s) = 1 - \frac{g'(s)}{L} \le 1 - \frac{\gamma}{L} = q < 1.$$

We try to approach the solution using the sequence $x_k = h(x_{k-1})$ of successive approximations. The Banach fixed point theorem implies the existence and uniqueness of a solution x to the equation $h(x) = x$, which immediately yields $g(x) = 0$. Moreover, the error bound

$$|x_k - x| \le \frac{q^k}{1-q} |x_1 - x_0|$$

can be established and $\lim_{k\to\infty} x_k = x$ is valid independently of the choice of the initial guess x_0.

This clever idea must be put into the context of PDEs and generalised to an appropriate form in order to handle the most interesting situations, namely $\gamma = 0$ or $L = \infty$, which cover degenerate nonlinear elliptic problems.

We recall that Pong and Yong [14] also applied the fixed-point argument to a Lipschitz continuous case for a simpler problem setting, but they were not able to establish the rate of convergence and also they did not discuss the degenerate case. Maitre [11] applied an iteration scheme for solving a nonlinear elliptic problem, but he was not able to handle nonlinearities of the type $g(x) = \operatorname{sign}(x)|x|^r$ for $0 < r < 1$.

The rate of convergence in the spaces $L_2(\Omega)$ and $H^1(\Omega)$ is shown in Theorems 3.2 and 4.3 and the main contribution of this paper is Theorem 4.4 (for strictly monotonically increasing nonlinearities), where convergence in $L_\infty(\Omega) \cap L_\infty(\Gamma_N)$ is shown. Here, the weak maximum principle proof-technique has been employed, which allows us to obtain the error estimates in the space $L_\infty(\Omega) \cap L_\infty(\Gamma_N)$ for a solution $u \in H^1(\Omega) \cap L_\infty(\Omega) \cap L_\infty(\Gamma_N)$.

The proposed technique can be also easily applied to a BVP with a nonlocal BC, see for example Slodička [15]. We have omitted this in our paper in order to focus on a new type of linearisation scheme and on the error estimates.

Throughout the paper C denotes a generic positive constant independent of the iteration parameter k.

2. Problem formulation and assumptions

Consider an open bounded set $\Omega \subset \mathbb{R}^N$, $N \geq 2$ with a Lipschitz continuous boundary Γ consisting of two complementary parts Γ_D and Γ_N. We assume that

$$|\Gamma_D| > 0. \tag{2.1}$$

We denote by $(w, z)_M$ the usual L_2-inner product of any real or vector-valued functions w, z on a set M.

We study the following nonlinear stationary BVP:

$$\begin{aligned} \nabla \cdot (-\boldsymbol{A}_{\rm dif}\nabla u - \boldsymbol{a}_{\rm con}u) + g(u) &= f && \text{in } \Omega, \\ u &= g_D && \text{on } \Gamma_D, \\ (-\boldsymbol{A}_{\rm dif}\nabla u - \boldsymbol{a}_{\rm con}u)\cdot \boldsymbol{\nu} - g_R(u) &= g_N && \text{on } \Gamma_N. \end{aligned} \tag{2.2}$$

The nonlinear functions g and g_R are supposed to be continuous and monotonically nondecreasing, that is, $0 \leq g', g_R'$, a.e. in $\mathbb{R}$. Later, we will also adopt some new assumptions on g and g_R depending on whether or not they are Lipschitz continuous.

The tensor $\boldsymbol{A}_{\rm dif}$ describing the diffusive properties of the material obeys the inequality

$$C_0\,|w|^2_{1,\Omega} \leq (\boldsymbol{A}_{\rm dif}\nabla w, \nabla w)_\Omega \leq C\,|w|^2_{1,\Omega}, \quad \forall w \in H^1(\Omega) \tag{2.3}$$

for given positive constants C_0 and C. The assumption (2.1) implies the fact that the seminorm $|\cdot|_{1,\Omega}$ represents an equivalent norm in $H^1(\Omega)$ to the usual norm $\|\cdot\|_{1,\Omega}$.

We consider such a type of convection $\boldsymbol{a}_{\rm con}$, which has been caused by an independent stationary process without spatially distributed sources. This can be mathematically described as

$$\begin{aligned} |\boldsymbol{a}_{\rm con}| &\leq C && \text{a.e. in } \Omega, \\ \boldsymbol{a}_{\rm con}\cdot\boldsymbol{\nu} &\geq 0 && \text{a.e. on } \Gamma_N, \\ \nabla\cdot\boldsymbol{a}_{\rm con} &= 0 && \text{a.e. in } \Omega. \end{aligned} \tag{2.4}$$

Further, we adopt standard assumptions on the source term and boundary data f, g_N and g_D:

$$f \in L_2(\Omega), \quad g_N \in L_2(\Gamma_N) \quad \text{and} \tag{2.5}$$

$$\exists \tilde{g} \in H^1(\Omega) \ \text{ such that } \ \tilde{g} = g_D \ \text{ on } \ \Gamma_D. \tag{2.6}$$

Let us introduce the standard subspace V of $H^1(\Omega)$,

$$V = \{\varphi \in H^1(\Omega);\ \varphi = 0 \text{ on } \Gamma_D, \},$$

as the space of all admissible test functions in a variational formulation. Define the bilinear form $a : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ as

$$a(u, \varphi) = (\boldsymbol{A}_{\text{dif}} \nabla u + \boldsymbol{a}_{\text{con}} u, \nabla \varphi)_\Omega, \quad \forall u, \varphi \in H^1(\Omega)$$

and the linear bounded functional $F : V \to \mathbb{R}$

$$\langle F, \varphi \rangle = (f, \varphi)_\Omega - (g_N, \varphi)_{\Gamma_N}, \quad \forall \varphi \in V.$$

The weak formulation of (2.2) has the following form: Find $u \in H^1(\Omega)$ such that $u - \tilde{g} \in V$ and

$$a(u, \varphi) + (g(u), \varphi)_\Omega + (g_R(u), \varphi)_{\Gamma_N} = \langle F, \varphi \rangle, \quad \forall \varphi \in V. \tag{2.7}$$

The well-posedness of this problem (existence and uniqueness) is guaranteed by the theory of monotone operators (see for example, Nečas [12]).

THEOREM 2.1. *Let the assumptions* (2.1) *and* (2.3)–(2.6) *be satisfied. Then there exists a unique weak solution* $u \in H^1(\Omega)$ *to the BVP* (2.7).

The main goal of this paper is to design a simple and efficient linear approximation scheme. We distinguish between two cases depending on the Lipschitz continuity of the nonlinear functions g and g_R. For simplicity we assume that both functions have (or have not) bounded derivatives. Of course, the case when one function is Lipschitz continuous and the other not is possible, and this can be obtained by suitable combination of the approximation schemes we will describe.

3. Lipschitz continuous functions g and g_R

We start with a simple case. Let both functions g and g_R be Lipschitz continuous with the Lipschitz constant L, that is,

$$\begin{aligned} |\beta(x) - \beta(y)| &\le L|x - y|, && \forall x, y \in \mathbb{R}, \\ 0 \le \beta' &\le L, && \text{a.e. in } \mathbb{R},\ \beta = g, g_R. \end{aligned} \tag{3.1}$$

Here, we follow the ideas from Slodička [17], where a BVP with Dirichlet BCs has been considered as a temporal problem by time discretisation. We design a recursive

sequence of linear elliptic BVPs, solutions of which will approach the weak solution u of (2.7). We start with any function u_0 satisfying

$$u_0 \in L_2(\Omega) \cap L_2(\Gamma_N)\,. \tag{3.2}$$

Further, u_k for $k = 1, 2, \dots$ is a weak solution to the following linear elliptic BVP: Find $u_k \in H^1(\Omega)$ such that $u_k - \tilde{g} \in V$ and

$$\begin{aligned} a(u_k, \varphi) + L(u_k, \varphi)_\Omega + L(u_k, \varphi)_{\Gamma_N} = \langle F, \varphi\rangle + L(u_{k-1}, \varphi)_\Omega - (g(u_{k-1}), \varphi)_\Omega \\ + L(u_{k-1}, \varphi)_{\Gamma_N} - (g_R(u_{k-1}), \varphi)_{\Gamma_N} \end{aligned} \tag{3.3}$$

holds for any $\varphi \in V$.

First, we show the well-posedness of the BVP (3.3).

LEMMA 3.1. *Let the assumptions* (2.1), (2.3)–(2.6), (3.1) *and* (3.2) *be satisfied. Then the sequence* $\{u_k\}_{k=1}^{\infty} \subset H^1(\Omega)$ *is well defined.*

PROOF. Let w be any function from V. Assumption (2.3) implies

$$C\,|w|^2_{1,\Omega} \geq a(w, w) \geq C_0\,|w|^2_{1,\Omega}\,. \tag{3.4}$$

The relation (2.4) together with the Friedrichs inequality and Green's theorem give the estimate

$$\begin{aligned} C\,|w|^2_{1,\Omega} \geq (\boldsymbol{a}_{\rm con} w, \nabla w)_\Omega &= \frac{1}{2}(\boldsymbol{a}_{\rm con}, \nabla w^2)_\Omega \\ &= -\frac{1}{2}(\nabla \cdot \boldsymbol{a}_{\rm con}, w^2)_\Omega + \frac{1}{2}(\boldsymbol{a}_{\rm con} \cdot \boldsymbol{\nu}, w^2)_\Gamma \\ &= \frac{1}{2}(\boldsymbol{a}_{\rm con} \cdot \boldsymbol{\nu}, w^2)_{\Gamma_N} \geq 0. \end{aligned} \tag{3.5}$$

Hence the left-hand side of (3.3) is a V-elliptic continuous bilinear form.

Take $k = 1$. The right-hand side of (3.3), according to (2.5), (3.1) and (3.2), is a bounded linear functional on V. Thus the existence and uniqueness of a weak solution $u_1 \in H^1(\Omega)$ to the BVP (3.3) follows from the Lax-Milgramm lemma.

If $u_{k-1} \in H^1(\Omega)$, the right-hand side of (3.3) is a bounded linear functional on V. Thus there exists a unique weak solution $u_k \in H^1(\Omega)$ satisfying (3.3).

We now define the following functions:

$$h(s) = g(s) - Ls, \quad h_R(s) = g_R(s) - Ls, \quad s \in \mathbb{R}. \tag{3.6}$$

Subtracting (2.7) from (3.3), we get the variational formulation for the error $u_k - u$

$$\begin{aligned} a(u_k - u, \varphi) + L(u_k - u, \varphi)_\Omega + L(u_k - u, \varphi)_{\Gamma_N} \\ = (h(u) - h(u_{k-1}), \varphi)_\Omega + (h_R(u) - h_R(u_{k-1}), \varphi)_{\Gamma_N}, \end{aligned} \tag{3.7}$$

which holds for any $\varphi \in V$.

Our next goal is to derive the error estimates in the $H^1(\Omega)$ space for the linearisation scheme (3.3).

THEOREM 3.2. *Let the assumptions of Lemma 3.1 be satisfied. Then there exist positive constants C and δ such that*

$$\begin{aligned}&\|u_k - u\|^2_{0,\Omega} + \|u_k - u\|^2_{0,\Gamma_N} + |u_k - u|^2_{1,\Omega}\\ &\qquad \le C\left(1 - \frac{\delta}{L+\delta}\right)^k \left[\|u_0 - u\|^2_{0,\Omega} + \|u_0 - u\|^2_{0,\Gamma_N}\right]\end{aligned}$$

holds for all $k = 1, 2, \ldots$.

PROOF. Choose $\varphi = u_k - u \in V$ in (3.7) and get

$$\begin{aligned}&a(u_k - u, u_k - u) + L\,\|u_k - u\|^2_{0,\Omega} + L\,\|u_k - u\|^2_{0,\Gamma_N}\\ &\qquad = (h(u) - h(u_{k-1}), u_k - u)_\Omega + (h_R(u) - h_R(u_{k-1}), u_k - u)_{\Gamma_N}. \qquad (3.8)\end{aligned}$$

The crucial point is to estimate the terms in the right-hand side containing the functions h and h_R. To do this, we use (3.1) and deduce

$$\begin{aligned}-L \le h'(s) = g'(s) - L \le 0 \qquad &\text{a.e. in } \mathbb{R}\\ -L \le h'_R(s) = g'_R(s) - L \le 0 \qquad &\text{a.e. in } \mathbb{R}.\end{aligned}$$

Hence the derivatives of both functions h and h_R are bounded by the constant L, that is, $|h'(s)| \le L$ and $|h'_R(s)| \le L$ a.e. in $\mathbb{R}$.

Therefore, using the Cauchy and Young inequalities we deduce

$$\begin{aligned}|(h(u) - h(u_{k-1}), u_k - u)_\Omega| &\le \|h(u) - h(u_{k-1})\|_{0,\Omega}\,\|u_k - u\|_{0,\Omega}\\ &\le L\,\|u - u_{k-1}\|_{0,\Omega}\,\|u_k - u\|_{0,\Omega}\\ &\le \frac{L}{2}\,\|u - u_{k-1}\|^2_{0,\Omega} + \frac{L}{2}\,\|u_k - u\|^2_{0,\Omega}.\end{aligned}$$

Analogously we have

$$|(h_R(u) - h_R(u_{k-1}), u_k - u)_{\Gamma_N}| \le \frac{L}{2}\,\|u - u_{k-1}\|^2_{0,\Gamma_N} + \frac{L}{2}\,\|u_k - u\|^2_{0,\Gamma_N}.$$

The left-hand side of (3.8), according to the V-ellipticity of the bilinear form a (see (3.4)), can be estimated from below by

$$L\,\|u_k - u\|^2_{0,\Omega} + L\,\|u_k - u\|^2_{0,\Gamma_N} + C_0\,|u_k - u|^2_{1,\Omega}.$$

The generalised Friedrichs inequality (see Křížek and Neittaanmäki [10, page 26]) and the continuous embedding $L_2(\partial\Omega) \hookrightarrow H^1(\Omega)$ imply the existence of a positive real number δ such that

$$\delta \|w\|^2_{0,\Omega} \le \frac{C_0}{2} |w|^2_{1,\Omega}, \quad \delta \|w\|^2_{0,\Gamma_N} \le \frac{C_0}{2} |w|^2_{1,\Omega}, \tag{3.9}$$

holds for all $w \in V$. Thus the lower bound of the left-hand side of (3.8) is

$$\left(L + \frac{\delta}{2}\right) \|u_k - u\|^2_{0,\Omega} + \left(L + \frac{\delta}{2}\right) \|u_k - u\|^2_{0,\Gamma_N} + \frac{C_0}{2} |u_k - u|^2_{1,\Omega}.$$

Summarising the foregoing results we arrive at

$$\begin{aligned}(L+\delta)\left[\|u_k - u\|^2_{0,\Omega} + \|u_k - u\|^2_{0,\Gamma_N}\right] + C_0 |u_k - u|^2_{1,\Omega} \\ \le L\left[\|u - u_{k-1}\|^2_{0,\Omega} + \|u - u_{k-1}\|^2_{0,\Gamma_N}\right],\end{aligned}$$

which after a simple calculation gives

$$\begin{aligned}\|u_k - u\|^2_{0,\Omega} + \|u_k - u\|^2_{0,\Gamma_N} + \frac{C_0}{L+\delta} |u_k - u|^2_{1,\Omega} \\ \le \left(1 - \frac{\delta}{L+\delta}\right)\left[\|u - u_{k-1}\|^2_{0,\Omega} + \|u - u_{k-1}\|^2_{0,\Gamma_N}\right].\end{aligned} \tag{3.10}$$

We omit the third term on the left for a moment and obtain the recursion formula

$$\|u_k - u\|^2_{0,\Omega} + \|u_k - u\|^2_{0,\Gamma_N} \le \left(1 - \frac{\delta}{L+\delta}\right)\left[\|u - u_{k-1}\|^2_{0,\Omega} + \|u - u_{k-1}\|^2_{0,\Gamma_N}\right].$$

This after k iterations implies

$$\|u_k - u\|^2_{0,\Omega} + \|u_k - u\|^2_{0,\Gamma_N} \le \left(1 - \frac{\delta}{L+\delta}\right)^k \left[\|u_0 - u\|^2_{0,\Omega} + \|u_0 - u\|^2_{0,\Gamma_N}\right].$$

The rest of the proof comes from the last inequality and (3.10).

4. Non Lipschitz continuous functions g and g_R

Throughout this section we assume that the derivatives of both functions g and g_R are unbounded. The most interesting types of nonlinearities are depicted in Figure 1. To cover all these cases, we introduce the following class $\mathscr{D}_b$ of all real-valued functions β associated with any point $b \in \mathbb{R}$ and satisfying the next relations

$$\begin{gathered}\beta \in C(\mathbb{R}), \\ 0 \le \beta'(s) \le \infty \quad \text{a.e. in } \mathbb{R}, \\ \beta'(s_+)\,\beta'(s_-) = \infty \quad \Longrightarrow \quad s = b, \\ \beta'' \le 0 \quad \text{a.e. in } (b, \infty), \\ \beta'' \ge 0 \quad \text{a.e. in } (-\infty, b).\end{gathered}$$

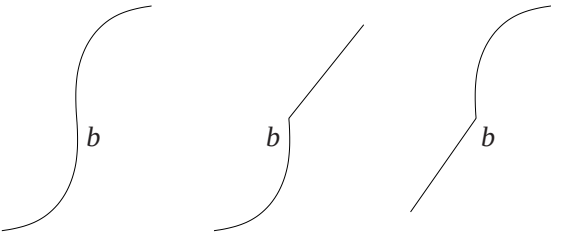


FIGURE 1. Examples of nonlinear functions from $\mathcal{Q}_b$

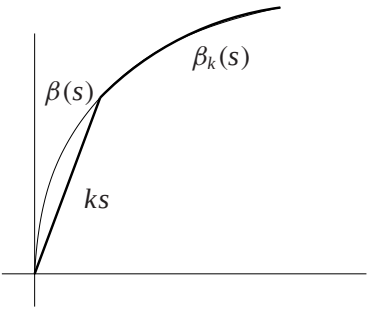


FIGURE 2. Local regularisation of β

Functions g and g_R can of course belong to different classes, but without loss of generality we assume that $g, g_R \in \mathcal{Q}_0$.

In light of the fact that the function β (stands for g or g_R) can degenerate, we regularise it first. Then we define a linearised approximation scheme, which is in some sense similar to the Lipschitz continuous case. We suppose that there exists a sequence of functions $\{\beta_k\}_{k=1}^{\infty}$ and positive real numbers ω and $\mathcal{L}$ satisfying

$$\begin{aligned} 0 \le \beta_k' &\le k\mathcal{L} && \text{a.e. in } \mathbb{R}, \\ |\beta(s) - \beta_k(s)| &\le Ck^{-\omega} && \forall k \ge k_0 \in \mathbb{N}, \\ \beta &= g, g_R. \end{aligned} \tag{4.1}$$

Without loss of generality one can assume that $k_0 = 1$. Similarly as in (3.6), we define $h_k(s) = g_k(s) - k\mathcal{L}s$ and $h_{R,k}(s) = g_{R,k}(s) - k\mathcal{L}s$, for $s \in \mathbb{R}$. In view of (4.1) we have

$$\begin{aligned} -k\mathcal{L} \le h_k'(s) = g_k'(s) - k\mathcal{L} \le 0 \quad \text{a.e. in } \mathbb{R}, \\ -k\mathcal{L} \le h_{R,k}'(s) = g_{R,k}'(s) - k\mathcal{L} \le 0 \quad \text{a.e. in } \mathbb{R}. \end{aligned}$$

Therefore the relations

$$|h_k'(s)| \le k\mathcal{L} \quad \text{and} \quad |h_{R,k}'(s)| \le k\mathcal{L} \tag{4.2}$$

are valid a.e. in $\mathbb{R}$.

We give a simple example of the regularisation to enhance readability.

EXAMPLE 1. Let the function β be defined as $\beta(s) = s|s|^{\alpha-1}$, where the real parameter α satisfies the condition $0 < \alpha < 1$. Clearly $\beta \in \mathscr{Q}_0$. We choose $\mathscr{L} = 1$. The regularisation β_k of β can be given as (see for example Figure 2)

$$\beta_k(s) = \begin{cases} \min\{\beta(s), ks\} & s > 0, \\ \max\{\beta(s), ks\} & s \le 0. \end{cases} \tag{4.3}$$

Clearly $0 \le \beta_k' \le k$ and one can easily check that

$$|\beta(s) - \beta_k(s)| \le C(\alpha) k^{-\alpha/(1-\alpha)}.$$

Now, we introduce a linearised scheme, the solution of which should approach the weak solution of (2.7). First, we replace the nonlinearity $\beta = g, g_R$ by its regularisation β_k, and then we apply a similar scheme to the Lipschitz case (3.3). The approximation scheme reads as: Find $u_k \in H^1(\Omega)$ such that $u_k - \tilde{g} \in V$ and

$$\begin{aligned} a(u_k, \varphi) + k\mathscr{L}(u_k, \varphi)_\Omega + k\mathscr{L}(u_k, \varphi)_{\Gamma_N} \\ = \langle F, \varphi\rangle + k\mathscr{L}(u_{k-1}, \varphi)_\Omega - (g_k(u_{k-1}), \varphi)_\Omega \\ + k\mathscr{L}(u_{k-1}, \varphi)_{\Gamma_N} - (g_{R,k}(u_{k-1}), \varphi)_{\Gamma_N} \end{aligned} \tag{4.4}$$

holds for any $\varphi \in V$.

The existence and uniqueness of a weak solution to the linear elliptic BVP (4.4) is guaranteed by the next lemma. The proof proceeds in the same way as in Lemma 3.1, therefore we omit it.

LEMMA 4.1. *Let $g, g_R \in \mathscr{Q}_0$. Assume (2.1), (2.3)–(2.6) and (3.2). Then the sequence $\{u_k\}_{k=1}^\infty \subset H^1(\Omega)$ is well defined.*

We subtract (2.7) from (4.4) and get the variational formulation for the error of the linearisation scheme

$$\begin{aligned} a(u_k - u, \varphi) + k\mathscr{L}(u_k - u, \varphi)_\Omega + k\mathscr{L}(u_k - u, \varphi)_{\Gamma_N} \\ = (g(u) - g_k(u), \varphi)_\Omega + (h_k(u) - h_k(u_{k-1}), \varphi)_\Omega \\ + (g_R(u) - g_{R,k}(u), \varphi)_{\Gamma_N} + (h_{R,k}(u) - h_{R,k}(u_{k-1}), \varphi)_{\Gamma_N}, \end{aligned} \tag{4.5}$$

which holds for any $\varphi \in V$.

The following lemma plays an important role in the derivation of the error estimates for the approximations u_k.

LEMMA 4.2. *Let a, b and ω be positive real numbers satisfying $b \neq \omega$. Assume that $\{y_k\}_{k=0}^{\infty}$ is a sequence of nonnegative real numbers obeying the following recursion formula*:

$$y_k \leq ak^{-1-\omega} + \left(1 - \frac{b}{k+b}\right) y_{k-1}, \quad k = 1, 2, \ldots.$$

Then there exists a positive constant $C = C(y_0, \omega, a, b)$ such that $y_k \leq Ck^{-\min\{b,\omega\}}$, $k = 1, 2, \ldots$.

PROOF. Suppose we have a recursion formula of the type $y_k \leq a_k + b_k y_{k-1}$, $k = 1, 2, \ldots$. One can prove by induction that

$$y_k \leq a_k + \sum_{j=1}^{k-1} a_j \prod_{i=j+1}^{k} b_i + y_0 \prod_{i=1}^{k} b_i \tag{4.6}$$

holds for all $k \in \mathbb{N}$. The details are left to the reader.

In our case we have $a_k = ak^{-1-\omega}$ and $b_k = 1 - b/(k+b)$. Now, we estimate all terms on the right-hand side of (4.6). We start with an obvious inequality for real numbers $1 + x \leq e^x$, for all $x \in \mathbb{R}$, which immediately gives $\prod_{i=1}^{m}(1+x_i) \leq e^{\sum_{i=1}^{m} x_i}$, for all $x_i \in \mathbb{R}$, $x_i \geq -1$. Therefore

$$\begin{aligned} y_0 \prod_{i=1}^{k} \left(1 - \frac{b}{i+b}\right) &\leq y_0 \exp\left(-b \sum_{i=1}^{k} \frac{1}{i+b}\right) \leq y_0 \exp\left(-b \int_1^{k+1} \frac{\mathrm{d}x}{x+b}\right) \\ &\leq y_0 \exp\left(b[\ln(1+b) - \ln(k+b)]\right) \\ &= y_0 (1+b)^b (k+1+b)^{-b} \leq Ck^{-b}. \end{aligned} \tag{4.7}$$

Similarly we estimate also the next term

$$\begin{aligned} &\sum_{j=1}^{k-1} \frac{C}{j^{1+\omega}} \prod_{i=j+1}^{k} \left(1 - \frac{b}{i+b}\right) \\ &\quad \leq C \sum_{j=1}^{k-1} \frac{1}{j^{1+\omega}} \exp\left(-b \sum_{i=j+1}^{k} \frac{1}{i+b}\right) \\ &\quad \leq C \sum_{j=1}^{k-1} \frac{1}{j^{1+\omega}} \exp\left(-b[\ln(k+1+b) - \ln(j+1+b)]\right) \\ &\quad = C \sum_{j=1}^{k-1} \frac{1}{j^{1+\omega}} \left(\frac{j+1+b}{k+1+b}\right)^b \leq C(k+1+b)^{-b} \sum_{j=1}^{k-1} (j+1+b)^{b-1-\omega} \\ &\quad \leq C(k+1+b)^{-b} \int_0^k (x+1+b)^{b-1-\omega}\, \mathrm{d}x \leq Ck^{-\min\{\omega,b\}}. \end{aligned} \tag{4.8}$$

Summarising the relations (4.6)–(4.8) we conclude the proof.

Now, let us turn our attention to the convergence proof of the approximations u_k.

THEOREM 4.3. *Let $g, g_R \in \mathscr{Q}_0$. Moreover, we assume* (2.1), (2.3)-(2.6), (3.2) *and* (4.1). *Then there exist positive constants C and δ such that*

$$\|u_k - u\|^2_{0,\Omega} + \|u_k - u\|^2_{0,\Gamma_N} \le Ck^{-\min\{2\omega,\delta/\mathscr{L}\}},$$
$$|u_k - u|^2_{1,\Omega} \le Ck^{1-\min\{2\omega,\delta/\mathscr{L}\}}$$

is valid for all $k \in \mathbb{N}$.

PROOF. Setting $\varphi = u_k - u \in V$ in (4.5) we have

$$\begin{aligned} a(u_k - u, u_k - u) + k\mathscr{L}(u_k - u, u_k - u)_\Omega + k\mathscr{L}(u_k - u, u_k - u)_{\Gamma_N} \\ = (g(u) - g_k(u), u_k - u)_\Omega + (h_k(u) - h_k(u_{k-1}), u_k - u)_\Omega \\ + (g_R(u) - g_{R,k}(u), u_k - u)_{\Gamma_N} \\ + (h_{R,k}(u) - h_{R,k}(u_{k-1}), u_k - u)_{\Gamma_N}. \end{aligned} \tag{4.9}$$

The term on the right-hand side containing the function g can be estimated using the Cauchy inequality, (4.1), Young's inequality, Sobolev's embedding theorem and at last the Friedrichs inequality. Successively we get for any $\eta \in \mathbb{R}_+$

$$\begin{aligned} |(g(u) - g_k(u), u_k - u)_\Omega| &\le \|g(u) - g_k(u)\|_{0,\Omega} \|u_k - u\|_{0,\Omega} \\ &\le Ck^{-\omega} \|u_k - u\|_{0,\Omega} \\ &\le C_\eta k^{-2\omega} + \eta \|u_k - u\|^2_{0,\Omega} \\ &\le C_\eta k^{-2\omega} + \eta \|u_k - u\|^2_{1,\Omega} \\ &\le C_\eta k^{-2\omega} + \eta |u_k - u|^2_{1,\Omega}. \end{aligned}$$

Analogously we deduce

$$\begin{aligned} |(g_R(u) - g_{R,k}(u), u_k - u)_\Omega| &\le C_\eta k^{-2\omega} + \eta \|u_k - u\|^2_{0,\Gamma_N} \\ &\le C_\eta k^{-2\omega} + \eta |u_k - u|^2_{1,\Omega}. \end{aligned}$$

Applying the Cauchy inequality, (4.2) and Young's inequality we obtain

$$\begin{aligned} |(h_k(u) - h_k(u_{k-1}), u_k - u)_\Omega| &\le \|h_k(u) - h_k(u_{k-1})\|_{0,\Omega} \|u_k - u\|_{0,\Omega} \\ &\le k\mathscr{L} \|u - u_{k-1}\|_{0,\Omega} \|u_k - u\|_{0,\Omega} \\ &\le \frac{k\mathscr{L}}{2} \|u - u_{k-1}\|^2_{0,\Omega} + \frac{k\mathscr{L}}{2} \|u_k - u\|^2_{0,\Omega} \end{aligned}$$

and in the same way we get

$$|(h_{R,k}(u) - h_{R,k}(u_{k-1}), u_k - u)_{\Gamma_N}| \le \frac{k\mathscr{L}}{2} \|u - u_{k-1}\|^2_{0,\Gamma_N} + \frac{k\mathscr{L}}{2} \|u_k - u\|^2_{0,\Gamma_N}.$$

According to the ellipticity of the bilinear form a, the left-hand side of (4.9) can be estimated from below by $k\mathscr{L}\,\|u_k-u\|^2_{0,\Omega}+k\mathscr{L}\,\|u_k-u\|^2_{0,\Gamma_N}+C_0\,|u_k-u|^2_{1,\Omega}$. Collecting the foregoing results together with (3.9), we arrive at

$$\frac{k\mathscr{L}+\delta}{2}\left[\|u_k-u\|^2_{0,\Omega}+\|u_k-u\|^2_{0,\Gamma_N}\right]+\left(\frac{C_0}{2}-\eta\right)|u_k-u|^2_{1,\Omega}$$
$$\le C_\eta k^{-2\omega}+\frac{k\mathscr{L}}{2}\left[\|u-u_{k-1}\|^2_{0,\Omega}+\|u-u_{k-1}\|^2_{0,\Gamma_N}\right].$$

Now, we choose $\eta=C_0/4$ and after a simple calculation we get

$$\|u_k-u\|^2_{0,\Omega}+\|u_k-u\|^2_{0,\Gamma_N}+\frac{C_0}{2(k\mathscr{L}+\delta)}\,|u_k-u|^2_{1,\Omega}$$
$$\le Ck^{-1-2\omega}+\left(1-\frac{\delta/\mathscr{L}}{k+\delta/\mathscr{L}}\right)\left[\|u-u_{k-1}\|^2_{0,\Omega}+\|u-u_{k-1}\|^2_{0,\Gamma_N}\right]. \tag{4.10}$$

Omit the third term on the left for a moment and get the following recursion formula:

$$\|u_k-u\|^2_{0,\Omega}+\|u_k-u\|^2_{0,\Gamma_N}$$
$$\le Ck^{-1-2\omega}+\left(1-\frac{\delta/\mathscr{L}}{k+\delta/\mathscr{L}}\right)\left[\|u-u_{k-1}\|^2_{0,\Omega}+\|u-u_{k-1}\|^2_{0,\Gamma_N}\right].$$

Lemma 4.2 yields

$$\|u_k-u\|^2_{0,\Omega}+\|u_k-u\|^2_{0,\Gamma_N}\le Ck^{-\min\{2\omega,\delta/\mathscr{L}\}}$$

and the rest of the proof follows from the last estimate and (4.10).

Theorem 4.3 proves the convergence of u_k to the exact solution u in the space $L_2(\Omega)\cap L_2(\Gamma_N)$. If $\min\{2\omega,\delta/\mathscr{L}\}>1$, then also the convergence in the norm of the Sobolev space $H^1(\Omega)$ is shown. This, of course, depends on the nonlinearity of g, g_R and also on the relation (3.9). The crucial point in the proof was the fact that the diffusion term has added a bit to the source term—see the relation (3.9). Let us note that if g', $g'_R>\gamma>0$, then the proof of Theorem 4.3 can be modified so that the relation (4.2) is replaced by

$$|h'_k(s)|\le k\mathscr{L}-\gamma\quad\text{and}\quad|h'_{R,k}(s)|\le k\mathscr{L}-\gamma, \tag{4.11}$$

which is valid a.e. in $\mathbb{R}$. Analogously one can prove

$$\begin{aligned}\|u_k-u\|^2_{0,\Omega}+\|u_k-u\|^2_{0,\Gamma_N}&\le Ck^{-\min\{2\omega,(\gamma+\delta)/\mathscr{L}\}}\\ |u_k-u|^2_{1,\Omega}&\le Ck^{1-\min\{2\omega,(\gamma+\delta)/\mathscr{L}\}}.\end{aligned} \tag{4.12}$$

EXAMPLE 2. Consider a real parameter α satisfying $0 < \alpha < 1$. Define the following function

$$\beta(s) = \begin{cases} s|s|^{\alpha-1} & \text{for } s \in [-1,1], \\ \alpha & \text{for } s \in \mathbb{R} \setminus [-1,1]. \end{cases}$$

Clearly $\beta' \geq \alpha$, that is, we can put $\gamma = \alpha$. Choose $\mathscr{L} = (1-\alpha)/2$ and define

$$\beta_k(s) = \begin{cases} \min\{\beta(s), k\mathscr{L}s\} & s > 0, \\ \max\{\beta(s), k\mathscr{L}s\} & s \leq 0. \end{cases}$$

Thus $0 \leq \beta_k' \leq k\mathscr{L}$ and a simple calculation gives for $k \geq 1/\mathscr{L}$

$$|\beta(s) - \beta_k(s)| \leq C(\alpha)\, k^{-\alpha/(1-\alpha)} = C\, k^{-\omega}.$$

Therefore

$$\min\left\{2\omega, \frac{\gamma+\delta}{\mathscr{L}}\right\} \geq \min\left\{2\omega, \frac{\gamma}{\mathscr{L}}\right\} = \min\left\{\frac{2\alpha}{1-\alpha}, \frac{2\alpha}{1-\alpha}\right\} = \frac{2\alpha}{1-\alpha} > 1$$

for $\alpha > 1/3$. According to the relation (4.12), we obtain

$$\lim_{k\to\infty} |u_k - u|_{1,\Omega} = 0 \quad \text{for } 1 > \alpha > 1/3.$$

The condition that $\beta' > \gamma > 0$, where β stands for g or g_R, is natural in some applications, see, for example, Barrett and Knabner [3]. Here, an equation of the type $\partial_t(u + [u]_+^p) - \Delta u = f$ with $0 < p < 1$ is considered. This, after time discretisation, leads to an elliptic equation of the form $v + [v]_+^p - \Delta v = F$.

Our next step is to prove convergence in the space $L_\infty(\Omega)$.

THEOREM 4.4. *Let the assumptions of Theorem 4.3 be satisfied. In addition we suppose $u \in L_\infty(\Omega) \cap L_\infty(\Gamma_N)$ and $0 < \gamma \leq \beta' \leq \infty$ for $\beta = g, g_R$. Then there exists a positive constant C such that*

$$\max\left\{\|u_k - u\|_{L_\infty(\Omega)}, \|u_k - u\|_{L_\infty(\Gamma_N)}\right\} \leq C\, k^{-\min\{\omega, \gamma/\mathscr{L}\}}$$

holds for all $k \in \mathbb{N}$.

PROOF. Fix the iteration parameter k and define the real constants A, B and M_{AB} in the following way:

$$A = (k\mathscr{L})^{-1} \|g(u) - g_k(u) + h_k(u) - h_k(u_{k-1})\|_{L_\infty(\Omega)},$$
$$B = (k\mathscr{L})^{-1} \left\|g_R(u) - g_{R,k}(u) + h_{R,k}(u) - h_{R,k}(u_{k-1})\right\|_{L_\infty(\Gamma_N)} \quad \text{and}$$
$$M_{AB} = \max\{A, B\}.$$

Denote by Ω^- and Γ_N^- the sets

$$\Omega^- = \{\boldsymbol{x} \in \Omega;\ u_k(\boldsymbol{x}) - u(\boldsymbol{x}) + M_{AB} < 0\} \quad \text{and}$$
$$\Gamma_N^- = \{\boldsymbol{x} \in \Gamma_N;\ u_k(\boldsymbol{x}) - u(\boldsymbol{x}) + M_{AB} < 0\}.$$

Let us suppose that at least one of these sets has a positive measure (the N- and $(N-1)$-dimensional measures are denoted by the same symbol), that is,

$$|\Omega^-| + |\Gamma_N^-| > 0.$$

We start again with the relation (3.7) and set $\varphi = (u_k - u + M_{AB})^- \in V$, where f^- stands for the usual cut-off function defined by $f^-(s) = \min\{f(s), 0\}$. We can write

$$\begin{aligned} &a(u_k - u, (u_k - u + M_{AB})^-) + k\mathscr{L}(u_k - u, (u_k - u + M_{AB})^-)_\Omega \\ &\quad + k\mathscr{L}(u_k - u, (u_k - u + M_{AB})^-)_{\Gamma_N} \\ &\quad = (h(u) - h(u_{k-1}), (u_k - u + M_{AB})^-)_\Omega \\ &\quad\quad + (h_R(u) - h_R(u_{k-1}), (u_k - u + M_{AB})^-)_{\Gamma_N}. \end{aligned}$$

This can be rewritten as

$$\begin{aligned} 0 &= \left(\boldsymbol{A}_{\mathrm{dif}}\nabla(u_k - u), \nabla(u_k - u + M_{AB})^-\right)_\Omega \\ &\quad + \left(\boldsymbol{a}_{\mathrm{con}}(u_k - u), \nabla(u_k - u + M_{AB})^-\right)_\Omega \\ &\quad + k\mathscr{L}\left(u_k - u - \frac{g(u) - g_k(u) + h_k(u) - h_k(u_{k-1})}{k\mathscr{L}}, (u_k - u + M_{AB})^-\right)_\Omega \\ &\quad + k\mathscr{L}\left(u_k - u - \frac{g_R(u) - g_{R,k}(u) + h_{R,k}(u) - h_{R,k}(u_{k-1})}{k\mathscr{L}}, (u_k - u + M_{AB})^-\right)_{\Gamma_N} \\ &= M_1 + M_2 + M_3 + M_4. \end{aligned} \tag{4.13}$$

The V-ellipticity of the matrix $\boldsymbol{A}_{\mathrm{dif}}$ (see (2.3)) implies the non-negativity of the term M_1, that is,

$$\begin{aligned} 0 &\le \left(\boldsymbol{A}_{\mathrm{dif}}\nabla(u_k - u + M_{AB})^-, \nabla(u_k - u + M_{AB})^-\right)_\Omega \\ &= \left(\boldsymbol{A}_{\mathrm{dif}}\nabla(u_k - u + M_{AB}), \nabla(u_k - u + M_{AB})^-\right)_\Omega \\ &= \left(\boldsymbol{A}_{\mathrm{dif}}\nabla(u_k - u), \nabla(u_k - u + M_{AB})^-\right)_\Omega \\ &= M_1. \end{aligned}$$

The convection term M_2 is also nonnegative. To show this, we apply Green's theorem,

(2.4) and (3.5). We successively obtain

$$
\begin{aligned}
M_2 &= \left(\boldsymbol{a}_{\mathrm{con}}(u_k-u), \nabla(u_k-u+M_{AB})^-\right)_\Omega \\
&= \left(\boldsymbol{a}_{\mathrm{con}}(u_k-u+M_{AB}), \nabla(u_k-u+M_{AB})^-\right)_\Omega \\
&\quad - M_{AB}\left(\boldsymbol{a}_{\mathrm{con}}, \nabla(u_k-u+M_{AB})^-\right)_\Omega && (\pm M_{AB}) \\
&= \underbrace{\left(\boldsymbol{a}_{\mathrm{con}}(u_k-u+M_{AB})^-, \nabla(u_k-u+M_{AB})^-\right)_\Omega}_{\geq 0} \\
&\quad + M_{AB}\left(\underbrace{\nabla\cdot\boldsymbol{a}_{\mathrm{con}}}_{=0}, (u_k-u+M_{AB})^-\right)_\Omega \\
&\quad - M_{AB}\left(\boldsymbol{a}_{\mathrm{con}}\cdot\boldsymbol{\nu}, (u_k-u+M_{AB})^-\right)_\Gamma && (\text{Green's thm.}) \\
&\geq -M_{AB}\left(\boldsymbol{a}_{\mathrm{con}}\cdot\boldsymbol{\nu}, (u_k-u+M_{AB})^-\right)_\Gamma && ((3.5),\,(2.4)) \\
&= -M_{AB}\left(\underbrace{\boldsymbol{a}_{\mathrm{con}}\cdot\boldsymbol{\nu}}_{\geq 0}, (u_k-u+M_{AB})^-\right)_{\Gamma_N} && ((2.4)) \\
&\geq 0.
\end{aligned}
$$

Using the obvious inequality

$$
u_k-u-\frac{h(u)-h(u_{k-1})}{k\mathscr{L}} \leq u_k-u+A \leq u_k-u+M_{AB} < 0,
$$

which is valid a.e. in Ω^-, we have

$$
M_3 = k\mathscr{L}\left(u_k-u-\frac{h(u)-h(u_{k-1})}{k\mathscr{L}}, (u_k-u+M_{AB})^-\right)_\Omega > 0.
$$

Analogously, applying the inequality (valid a.e. in Γ_N^-)

$$
u_k-u-\frac{h_R(u)-h_R(u_{k-1})}{k\mathscr{L}} \leq u_k-u+B \leq u_k-u+M_{AB} < 0,
$$

we get for M_4

$$
M_4 = k\mathscr{L}\left(u_k-u-\frac{h_R(u)-h_R(u_{k-1})}{k\mathscr{L}}, (u_k-u+M_{AB})^-\right)_{\Gamma_N} > 0.
$$

Collecting all the estimates for $M_1,\dots,M_4$ we arrive at

$$
M_1+M_2+M_3+M_4 > 0.
$$

This contradicts the relation (4.13) and the assumption $|\Omega^-| + |\Gamma_N^-| > 0$ fails to hold. In other words, we have just proved

$$\begin{aligned} u_k - u &\geq -M_{AB} \qquad \text{a.e. in } \Omega, \\ u_k - u &\geq -M_{AB} \qquad \text{a.e. in } \Gamma_N. \end{aligned} \tag{4.14}$$

The next step is to prove

$$\begin{aligned} u_k - u &\leq M_{AB} \qquad \text{a.e. in } \Omega, \\ u_k - u &\leq M_{AB} \qquad \text{a.e. in } \Gamma_N. \end{aligned} \tag{4.15}$$

Therefore, we introduce the sets Ω^+ and Γ_N^+ as

$$\begin{aligned} \Omega^+ &= \{\boldsymbol{x} \in \Omega;\ u_k(\boldsymbol{x}) - u(\boldsymbol{x}) - \max\{A, B\} > 0\} \quad \text{and} \\ \Gamma_N^+ &= \{\boldsymbol{x} \in \Gamma_N;\ u_k(\boldsymbol{x}) - u(\boldsymbol{x}) - M_{AB} > 0\}. \end{aligned}$$

We now put $\varphi = (u_k - u - M_{AB})^+ = \max\{u_k - u - M_{AB}, 0\} \in V$ into (3.7) and follow the same argument as before. So, we obtain (4.15).

In light of (4.14) and (4.15) we have

$$\max\left\{\|u_k - u\|_{L_\infty(\Omega)}, \|u_k - u\|_{L_\infty(\Gamma_N)}\right\} \leq M_{AB}. \tag{4.16}$$

The assumption $0 < \gamma \leq \beta'$ for $\beta = g, g_R$ implies the relation (4.11), which is valid a.e. in $\mathbb{R}$. Thus we successively get

$$\begin{aligned} A &= (k\mathscr{L})^{-1} \|g(u) - g_k(u) + h_k(u) - h_k(u_{k-1})\|_{L_\infty(\Omega)} \\ &\leq (k\mathscr{L})^{-1} \left(\|g(u) - g_k(u)\|_{L_\infty(\Omega)} + \|h_k(u) - h_k(u_{k-1})\|_{L_\infty(\Omega)}\right) \\ &\leq Ck^{-1-\omega} + \left(1 - \frac{\gamma}{k\mathscr{L}}\right) \|u - u_{k-1}\|_{L_\infty(\Omega)} \\ &\leq Ck^{-1-\omega} + \left(1 - \frac{\gamma/\mathscr{L}}{k + \gamma/\mathscr{L}}\right) \|u - u_{k-1}\|_{L_\infty(\Omega)} \\ &\leq Ck^{-1-\omega} + \left(1 - \frac{\gamma/\mathscr{L}}{k + \gamma/\mathscr{L}}\right) \max\left\{\|u_{k-1} - u\|_{L_\infty(\Omega)}, \|u_{k-1} - u\|_{L_\infty(\Gamma_N)}\right\} \end{aligned}$$

and

$$\begin{aligned} B &= (k\mathscr{L})^{-1} \left\|g_R(u) - g_{R,k}(u) + h_{R,k}(u) - h_{R,k}(u_{k-1})\right\|_{L_\infty(\Gamma_N)} \\ &\leq (k\mathscr{L})^{-1} \left(\left\|g_R(u) - g_{R,k}(u)\right\|_{L_\infty(\Gamma_N)} + \left\|h_{R,k}(u) - h_{R,k}(u_{k-1})\right\|_{L_\infty(\Gamma_N)}\right) \\ &\leq Ck^{-1-\omega} + \left(1 - \frac{\gamma}{k\mathscr{L}}\right) \|u - u_{k-1}\|_{L_\infty(\Gamma_N)} \\ &\leq Ck^{-1-\omega} + \left(1 - \frac{\gamma/\mathscr{L}}{k + \gamma/\mathscr{L}}\right) \|u - u_{k-1}\|_{L_\infty(\Gamma_N)} \\ &\leq Ck^{-1-\omega} + \left(1 - \frac{\gamma/\mathscr{L}}{k + \gamma\mathscr{L}}\right) \max\left\{\|u_{k-1} - u\|_{L_\infty(\Omega)}, \|u_{k-1} - u\|_{L_\infty(\Gamma_N)}\right\}. \end{aligned}$$

The last two estimates and (4.16) imply the following recursion formula for $k = 1, 2, \ldots$:

$$\max\left\{\|u_k - u\|_{L_\infty(\Omega)}, \|u_k - u\|_{L_\infty(\Gamma_N)}\right\} \leq Ck^{-1-\omega} + \left(1 - \frac{\gamma/\mathscr{L}}{k + \gamma/\mathscr{L}}\right) \max\left\{\|u_{k-1} - u\|_{L_\infty(\Omega)}, \|u_{k-1} - u\|_{L_\infty(\Gamma_N)}\right\}.$$

The rest of the proof can be obtained by a simple application of Lemma 4.2.

5. Numerical experiments

In this section we present two numerical examples to demonstrate the efficiency and robustness of the proposed linearisation schemes (3.3) and (4.4). For the numerical solution of a linear elliptic equation we have used the mixed non-conforming finite element formulation. This is equivalent to the mixed-hybrid method (see Arnold and Brezzi [2]). We explain very briefly the main idea of this approximation.

Let us consider a regular triangulation $\mathscr{T}_h$ (h denotes the mesh diameter) of the domain Ω. On each element $\mathscr{T} \in \mathscr{T}_h$ we define three linear basis functions associated with edges of $\mathscr{T}$, that is, a basis function has the value 1 at the midpoint of one edge and 0 at the midpoints of the different edges of one triangle. Further we define a bubble function on $\mathscr{T}$, which is a polynomial function of third order vanishing on the boundary $\partial\mathscr{T}$ and its integral average value on $\mathscr{T}$ is 1. In this way we have enriched the standard linear non-conforming space by bubbles, and we solve the linear elliptic problem in this space replacing the velocity field $\mathbf{q}$ by its projection on the Raviart-Thomas space RT_0. For more details see Arnold and Brezzi [2].

For the analysis of the mixed finite element discretisation for the Lipschitz continuous case (Dirichlet problem) we refer the reader to Slodička [16].

5.1. Lipschitz continuous case Let Ω be the unit square in $\mathbb{R}^2$, the boundary of which is split into two parts Γ_D and Γ_N, see Figure 3.

We consider the same nonlinear function in the domain and on Γ_N, that is, $g \equiv g_R$, which is defined as

$$g(s) = \begin{cases} \arctan s & \text{for } s \leq 1, \\ \pi/4 & \text{elsewhere.} \end{cases}$$

This is clearly continuous. For the derivative we have

$$g'(s) = \begin{cases} 1/(1+s^2) & \text{for } s < 1, \\ 0 & \text{elsewhere,} \end{cases}$$

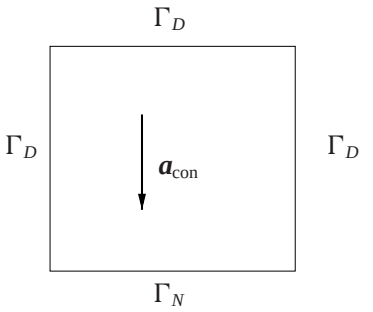


FIGURE 3. Domain Ω with convection $\boldsymbol{a}_{\text{con}}$

thus $0 \le g' < 1$ a.e. in $\mathbb{R}$.

The convection term $\boldsymbol{a}_{\text{con}} = (0, -1)$ clearly fulfills the assumption (2.4). We consider the following nonlinear elliptic BVP: Find $u \in H^1(\Omega)$ such that

$$\begin{aligned} \nabla \cdot (-\nabla u - \boldsymbol{a}_{\text{con}} u) + g(u) &= f \quad \text{in } \Omega, \\ u &= g_D \quad \text{on } \Gamma_D, \\ (-\nabla u - \boldsymbol{a}_{\text{con}} u) \cdot \boldsymbol{\nu} - g(u) &= g_N \quad \text{on } \Gamma_N, \end{aligned}$$

where the data functions f, g_D and g_N are defined in such a way that the exact solution of this BVP is

$$u(x, y) = x^3 - y^2 + x + \sin(\pi x) \sin(\pi y).$$

We have used the linearisation scheme (3.3) with $L = 1$ for computations.

Let us introduce a random function ran whose range is uniformly distributed over $(-1, 1)$. We present two computations. In the first case, we choose u_0 relatively close (up to 50% error) to the exact solution, that is,

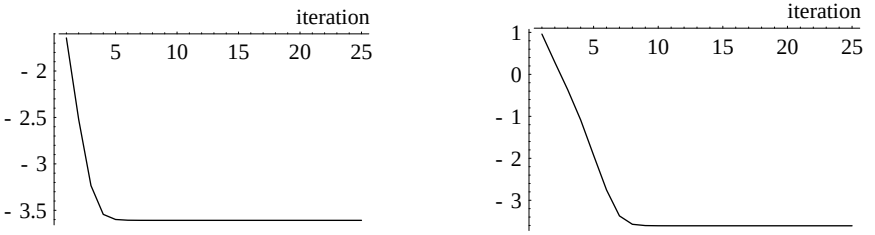
$$u_0(\boldsymbol{x}) = u(\boldsymbol{x})(1 + 0.5\,\text{ran}(\boldsymbol{x})).$$

In the second event we begin with u_0, which is far away from the solution u, that is,

$$u_0(\boldsymbol{x}) = 100\,\text{ran}(\boldsymbol{x}).$$

Let us note that the random function ran has been evaluated once per a given triangle or an edge.

We have used a fixed uniform mesh consisting of 5 000 triangles, which corresponds to $\Delta x = \Delta y = 0.02$, and we have computed 25 iterations. Then we have evaluated various errors of u_k and plotted them versus iterations $k = 1, \ldots, 25$. In order to get a better feeling for the rate of convergence, we have depicted *logarithms* of errors instead of errors on the y-axes—see Figure 4. Here, the left column represents the case for a good starting point u_0, while the right column corresponds to a very badly chosen u_0.



u_0 was close to u u_0 was far from u

FIGURE 4. Logarithms of $L_2(\Omega)$-errors for u_k versus iterations

5.2. Non Lipschitz continuous case Take $\Omega = [0,1]^2$. Consider the nonlinear function g given by

$$g(s) = \begin{cases} \sqrt{s} & \text{for } s > 0, \\ 0 & \text{elsewhere}, \end{cases}$$

which is clearly non Lipschitz continuous. We want to find a solution to the following nonlinear Dirichlet problem:

$$\begin{aligned} \nabla \cdot (-\nabla u) + g(u) &= f \quad \text{in } \Omega \\ u &= g_D \quad \text{on } \Gamma. \end{aligned}$$

The data functions f and g_D are defined in such a way that the exact solution of this BVP is

$$u(x,y) = x^3 - y^2 + x + \sin(\pi x)\sin(\pi y).$$

We have used the linearisation scheme (4.4) with $\mathscr{L} = 1$ for computations, where the approximation g_k is given by (4.3). We start from u_0, which is far away from the solution u, that is,

$$u_0(\boldsymbol{x}) = 100\,\mathrm{ran}(\boldsymbol{x}).$$

We have again used the same uniform mesh consisting of 5 000 triangles corresponding to $\Delta x = \Delta y = 0.02$, and we have computed 25 iterations. The results are depicted in Figure 5.

5.3. Conclusion Figures 4 and 5 show the behaviour of the $L_2(\Omega)$-error of the iteration process. One can also compute the $H^1(\Omega)$- and $L_\infty(\Omega)$-errors. The graphs will have the same character. The rapidly decreasing part at the beginning is followed by a more or less constant section. The reason for this is that the initiate dominant

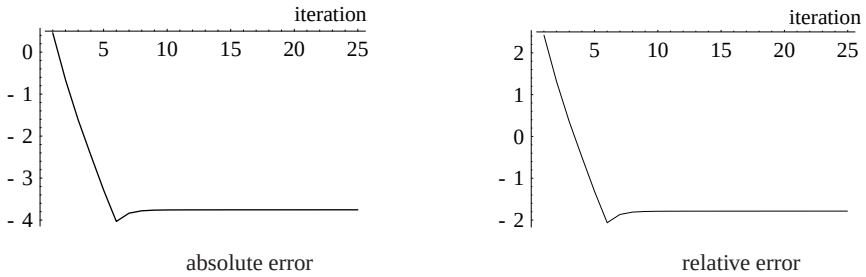


FIGURE 5. Logarithms of $L_2(\Omega)$-errors for u_k versus iterations

linearisation error becomes subjacent to the discretisation error as the number of iterations increases.

We can really observe that the linearisation schemes (3.3) and (4.4) are robust and that the approximations converge towards the exact solutions independently of where the iteration process has started. The robustness of the scheme allows the use of large time steps in the computation of evolution problems. The convergence at each time point of a suitable time partitioning is independent of the time step size. This is a big difference from other frequently used algorithms.

Moreover, both numerical schemes are efficient. In particular, we needed 7–8 iterations to get the best possible error for the given discretisation, although u_0 was really badly chosen. In the instance of a good starting point u_0, it is enough to do 3–4 iterations to achieve the discretisation error.

Acknowledgement

This work was supported by the BOF/GOA-project no. 12 052 499 of Ghent University.

References

[1] H. Amann, “Supersolution, monotone iteration and stability”, *J. Diff. Eq.* **21** (1976) 367–377.

[2] D. N. Arnold and F. Brezzi, “Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates”, *RAIRO Modél. Math. Anal. Numér.* **19** (1) (1985) 7–32.

[3] J. W. Barrett and P. Knabner, “Finite element approximation of the transport of reactive solutes in porous media. Part II: Error estimates for equilibrium adsorption processes”, *SIAM J. Numer. Anal.* **34** (1997) 455–479.

[4] Y. Deng, G. Chen, W.-M. Ni and J. Zhou, “Boundary element monotone iteration scheme for semilinear elliptic partial differential equations”, *Math. Comput.* **65** (1996) 943–982.

[5] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics 19 (American Mathematical Society, Providence, RI, 1998).

[6] W. Jäger and J. Kačur, “Solution of porous medium type systems by linear approximation schemes”, *Numer. Math.* **60** (1991) 407–427.
[7] W. Jäger and J. Kačur, “Solution of doubly nonlinear and degenerate parabolic problems by relaxation schemes”, *RAIRO Modél. Math. Anal. Numér.* **29** (5) (1995) 605–627.
[8] J. Kačur, “Solution of some free boundary problems by relaxation schemes”, *SIAM J. Numer. Anal.* **36** (1999) 290–316.
[9] J. Kačur, “Solution to strongly nonlinear parabolic problems by a linear approximation scheme”, *IMA J. Numer. Anal.* **19** (1) (1999) 119–145.
[10] M. Křížek and P. Neittaanmäki, *Mathematical and numerical modelling in electrical engineering. Theory and applications*, Volume 1 of *Mathematical modelling: Theory and applications*, (Kluwer, Dordrecht, 1996).
[11] E. Maitre, “Numerical analysis of nonlinear elliptic-parabolic equations”, *RAIRO Modél. Math. Anal. Numér.* **36** (2002) 143–153.
[12] J. Nečas, *Introduction to the theory of nonlinear elliptic equations* (John Wiley & Sons, Chichester, 1986).
[13] C. V. Pao, *Nonlinear parabolic and elliptic equations* (Plenum Press, New York, 1992).
[14] I. S. Pop and W.-A. Yong, “On the existence and uniqueness of a solution for an elliptic problem”, *Babes-Bolyayi* **45** (2000) 97–107.
[15] M. Slodička, “Error estimates of an efficient linearization scheme for a nonlinear elliptic problem with a nonlocal boundary condition”, *RAIRO Modél. Math. Anal. Numér.* **35** (2001) 691–711.
[16] M. Slodička, “Mixed finite element method for nonlinear second-order elliptic problems: Relaxation scheme”, in *Algoritmy 2002* (eds. A. Handlovičová, Z. Krivá, K. Mikula and D. Ševčovič), (Slovak University of Technology, Faculty of Civil Engineering, Department of Mathematics and Descriptive Geometry, Bratislava, 2002), 49–57.
[17] M. Slodička, “A robust and efficient linearization scheme for doubly nonlinear and degenerate parabolic problems arising in flow in porous media”, *SIAM J. Sci. Comput.* **23** (2002) 1593–1614.