Gradient-Fourier method for nonlinear elliptic partial differential equations in Sobolev space

Lajos Lóczi

Supervisor: János Karátson

Department of Applied Analysis
Eötvös Loránd University

June 2000

Contents

Chapter 1

Introduction

Our thesis deals with the iterative solution of a certain class of nonlinear elliptic partial differential equations. It incorporates two methods: the gradient method and the Fourier method.

The given partial differential equation can be written as $F(u) = 0$ with a suitable operator F and the unknown function u. The principle of variation will be made use of for the construction of the solution, that is, the solution of the equation is transformed into finding a minimum of an appropriately defined functional Φ using the concept of the potential operator. (The whole idea rests on the fact that if a function attains its minimum at a point, then its derivative vanishes at that point. Here the operator F plays the role of the derivative.) The technical background of this will be discussed in Chapter 2.

In the general case there is no guarantee that the functional Φ has a minimum. Or, if it has, it may have more than one: several local minima can occur. The theorems in Chapter 3 will give sufficient conditions which ensure the existence and uniqueness of a global minimum point, consequently the corresponding equation admits a unique solution.

Then the gradient method will be used to minimize the functional Φ. Another name of the method is the method of steepest descent, which better illustrates the gist of it. One can think of the values of Φ as defining the height of a surface. We pick an initial point on this surface and try to get progressively more accurate approximations to the minimum. A natural way to achieve this is to follow the path of steepest descent on the surface. In other words, we select the direction for the next iteration step to be the one in which the directional derivative—or gradient—of Φ is minimal. This procedure yields an iterative sequence which converges to the minimum point.

In Chapter 4 the gradient method will be applied to the given partial differential equation (or more precisely, to the boundary value problem). The iterative sequence obtained by the gradient method runs in a suitable Hilbert space (namely, in a Sobolev space). Our main reference here is [11].

The construction of the iterative sequence is effectively reduced to the solution of linear boundary value problems of Poisson type. Chapter 5 will show how to tackle these linear partial differential equations using the Fourier method, which is essentially based on Fourier series expansion. At the end, we give an error estimate of the combined Gradient-Fourier method.

Finally we list some further references. The gradient method in general is described in [10, Chapter XV]. The classical gradient method in finite dimensional spaces and its applications can be found in [15] and [17].

There are various other ways to solve nonlinear elliptic problems numerically. The finite element method and the method of finite differences transform the boundary value problem into an algebraic system of equations in $\mathbb{R}^N$ (see [3, 20]). Another approach is Newton's method (see for example in [10]), which is also of great practical importance.

The related Hilbert space theory of the finite element method is used to establish error estimates. Iterations using Hilbert space theory are for example in [2], [11], [14] and [18]. Our thesis also employs these techniques.

Acknowledgement

The author[1] is indebted to János Karátson for his clear and concise lectures, for the invaluable resources and explanations and for his helpful remarks and corrections about the thesis.

[1] lloczi@cs.elte.hu

Chapter 2

Solving equations and minimizing functionals

This chapter introduces some definitions and concepts used in the thesis, and gives the basic tools for later applications. Most of these propositions will not be proved here. (Many of the proofs can be carried out using their counterparts from real analysis.) For an account of the subject, see [10] and [20].

Definition 2.1 *Let X and Y be normed spaces, $u \in X$ and $A: X \to Y$. We say that the operator A is Gâteaux differentiable at u if*

(i) the directional derivative $\partial_v A(u) \in Y$ exists for all $v \in X$

(ii) $X \ni v \mapsto \partial_v A(u) \in Y$ is a continuous linear mapping.

The continuous linear mapping in (ii) will be denoted by $A'(u)$. If $Y = \mathbb{R}$ ($\Rightarrow A'(u) \in X^*$, the dual space), then instead of $A'(u)v$ we often write $\langle A'(u), v\rangle$.

The closed line segment between $u \in X$ and $v \in X$ will be denoted by $[u, v]$, that is $[u, v] := \{u + t(v - u) \mid t \in [0, 1]\} \subset X$. (Similarly, the open line segment is (u, v).)

Proposition 2.2 *Let X be a normed space, $u, v \in X$ and $\Phi: X \to \mathbb{R}$ such that $\Phi|_{[u,v]}$ is Gâteaux differentiable. Then there exists $\xi \in (u, v)$ such that*

$$\Phi(v) - \Phi(u) = \langle \Phi'(\xi), v - u\rangle .$$

Proposition 2.3 *Let X be a normed space, $u, v \in X$ and $\Phi: X \to \mathbb{R}$ such that $\Phi|_{[u,v]}$ is twice Gâteaux differentiable. Then there exists $\xi \in (u, v)$ such that*

$$\Phi(v) - \Phi(u) = \langle \Phi'(u), v - u\rangle + \frac{1}{2} \langle \Phi''(\xi)(v - u), v - u\rangle .$$

If X and Y are normed spaces, $B(X, Y)$ will denote the set of continuous linear mappings from X to Y.

Definition 2.4 *Let X, Y and Z be normed spaces.*
An operator $A: X \to B(Y, Z)$ is said to be

(i) hemicontinuous, if for all $u, v \in X$ and for all $w \in Y$ the mapping

$$\mathbb{R} \ni t \mapsto A(u + tv)w \in Z$$

is continuous.

(ii) bihemicontinuous, if for all $u, v_1, v_2 \in X$ and for all $w \in Y$ the mapping

$$\mathbb{R}^2 \ni (s, t) \mapsto A(u + sv_1 + tv_2)w \in Z$$

is continuous.

Proposition 2.5 *Let X be a normed space. Let $\Phi: X \to \mathbb{R}$ be Gâteaux differentiable such that Φ' is hemicontinuous. Then for all $u, v \in X$*

$$\Phi(v) - \Phi(u) = \int_0^1 \langle \Phi'(u + t(v - u)), v - u \rangle \mathrm{d}t.$$

Moreover, if Φ is twice Gâteaux differentiable, and Φ'' is hemicontinuous, then

$$\Phi'(v) - \Phi'(u) = \int_0^1 \Phi''(u + t(v - u))(v - u) \mathrm{d}t,$$

that is, for all $h \in X$

$$\langle \Phi'(v) - \Phi'(u), h \rangle = \int_0^1 \langle \Phi''(u + t(v - u))(v - u), h \rangle \mathrm{d}t.$$

For any functional $\Phi: X \to \mathbb{R}$ and vectors $u, v \in X$, let $\phi_{u,v}$ denote the following $\mathbb{R} \to \mathbb{R}$ mapping:

$$\phi_{u,v}(t) := \Phi(u + t(v - u)) \quad (t \in [0, 1]).$$

Definition 2.6 *Let X be a normed space. A functional $\Phi: X \to \mathbb{R}$ is said to be (strictly) convex, if for all $u, v \in X$ the function $\phi_{u,v}: \mathbb{R} \to \mathbb{R}$ is (strictly) convex.*

Definition 2.7 *Let X be a normed space. An operator $A: X \to X^*$ is said to be monotone, if for all $u, v \in X$*

$$\langle A(u) - A(v), u - v \rangle \geq 0.$$

Proposition 2.8 *Let X be a normed space. Let $\Phi: X \to \mathbb{R}$ be twice Gâteaux differentiable. Then the following statements are equivalent:*

(i) Φ is convex

(ii) $\Phi': X \to X^*$ is monotone

(iii) $\Phi'' \geq 0$, that is for all $u, h \in X$ $\langle \Phi''(u)h, h \rangle \geq 0$.

The following definitions and statements will establish connection between solving equations and minimizing functionals.

Proposition 2.9 *Let X be a normed space and let $\Phi: X \to \mathbb{R}$ be Gâteaux differentiable. If Φ attains its minimum at a point $u \in X$, then $\Phi'(u) = 0$.*

Definition 2.10 *Let X be a normed space, $A: X \to X^*$ and $b \in X^*$. Consider the equation $A(u) = b$, where $u \in X$ is the unknown vector. We say that a functional $\Phi: X \to \mathbb{R}$ is a minimizing functional (with respect to the given equation) if*

$$A(u^*) = b \iff \Phi(u^*) = \min_X \Phi.$$

Definition 2.11 *Let X be a real normed space. An operator $A: X \to X^*$ is said to be a potential operator, if there exists a functional $J: X \to \mathbb{R}$ such that J is Gâteaux differentiable and $J'(u) = A(u)$ $(u \in X)$. (When X is a Hilbert space, the dual of H will be identified with itself.)*

Proposition 2.12 *Let X be a real normed space and let $A: X \to X^*$ be a potential operator with a convex potential J. Let $b \in X^*$ arbitrary. Then the following functional*

$$\Phi(u) := J(u) - \langle b, u \rangle \in \mathbb{R} \quad (u \in X)$$

is a minimizing functional with respect to the equation $A(u) = b$.

Proposition 2.13 *Let X be a real normed space. Let $A: X \to X^*$ be Gâteaux differentiable and let A' be bihemicontinuous. Then the following are equivalent:*

(i) A is a potential operator

(ii) for all $w \in X$ $A'(w)$ is symmetric (that is $\langle A'(w)u, v \rangle = \langle A'(w)v, u \rangle$, for all $u, v \in X$).

A sufficient condition for the minimum to exist is

Proposition 2.14 *Let X be a reflexive normed space. Let $\Phi: X \to \mathbb{R}$ be a Gâteaux differentiable and convex functional, further suppose that*

$$\lim_{\|u\| \to \infty} \Phi(u) = \infty.$$

Then Φ attains its minimum in X. (Moreover, if Φ is strictly convex, then the minimum is unique.)

For the extension theorem we will need

Definition 2.15 *Let H be a Hilbert space with a dense subspace $D \subset H$. Let $B: D \to H$ be a linear operator with $B \geq \lambda \cdot I$, where $\lambda > 0$ is a constant and I is the identity operator. Further, define a scalar product*

$$\langle u, v \rangle_B := \langle Bu, v \rangle \quad (u, v \in D).$$

Then the completion of $(D, \langle \cdot, \cdot \rangle_B)$ is called the energy space (with respect to B), and denoted by H_B. The corresponding norm is

$$\|u\|_B := \sqrt{\langle u, u \rangle_B}.$$

Proposition 2.16 *Suppose the operator B has the same properties as above. Then the following inequality holds:*

$$\|u\|_B \le \frac{1}{\sqrt{\lambda}} \cdot \|Bu\| \quad (u \in D).$$

Finally we prove a lemma (see [6]).

Lemma 2.17 *Let H be a real Hilbert space. Let $F: H \to H$ be the operator defined by*

$$F(u) := u - \frac{2}{M+m} C(u) \quad (u \in H),$$

where m and M $(0 < m \le M)$ are constants, $C: H \to H$ is a Gâteaux differentiable operator, C' is hemicontinuous, further $C'(u)$ $(u \in H)$ is a bounded self-adjoint operator with

$$m \cdot \|u\|^2 \le \langle C'(u)h, h\rangle \le M \cdot \|u\|^2 \quad (u, h \in H). \tag{2.1}$$

Then for $u, v \in H$

$$\|F(v) - F(u)\| \le \frac{M-m}{M+m} \|v - u\|.$$

Proof. Let $u, v \in H$ be fixed. Using Proposition 2.5 we get

$$F(v) - F(u) = \int_0^1 F'(u + t(v-u))(v-u)\mathrm{d}t =$$

$$= v - u - \frac{2}{M+m} \cdot \int_0^1 C'(u + t(v-u))(v-u)\mathrm{d}t.$$

Define a linear operator $L: H \to H$ by

$$Lx := x - \frac{2}{M+m} \cdot \int_0^1 C'(u + t(v-u))x\mathrm{d}t \quad (x \in H).$$

Then L is a bounded self-adjoint operator, since $C'(w)$ has these properties for all $w \in H$. Further, using the estimate (2.1)

$$-\frac{M-m}{M+m}\|x\|^2 = \|x\|^2 - \frac{2}{M+m} M \cdot \|x\|^2 \le \langle Lx, x\rangle \le$$

$$\le \|x\|^2 - \frac{2}{M+m} m \cdot \|x\|^2 = \frac{M-m}{M+m}\|x\|^2 \quad (x \in H).$$

Hence $\|L\| \le \frac{M-m}{M+m}$. Since $F(v) - F(u) = L(v-u)$, the Lemma is proved. ∎

Chapter 3

The gradient method

The first theorem of this chapter introduces the gradient method. The method here will be used for the iterative solution of equations $A(u) = b$ in real Hilbert spaces. The operator A need not be linear, however, Gâteaux differentiability is required. The proof follows the one described in [12]. The application of the gradient method to continuous linear operators in Hilbert spaces is discussed in [9] and [10]. For the nonlinear case, see [7], [10] and [19]. The most important variant of the method, the conjugate gradient method, can be found in [1] and [5].

Theorem 3.1 *Let $(H, \langle\cdot,\cdot\rangle)$ be a real Hilbert space and $A\colon H \to H$ be a (nonlinear) operator with the following properties:*

(i) A is Gâteaux differentiable and A' is bihemicontinuous

(ii) $A'(u)$ is self-adjoint $(u \in H)$

(iii) there exist functions $m, M\colon [0,\infty) \to [0,\infty)$, m is monotone decreasing, M is monotone increasing and $\lim_{r\to\infty} r \cdot m(r) = \infty$ such that for all $u \in H$

$$m(\|u\|)\cdot I \le A'(u) \le M(\|u\|)\cdot I,$$

where I is the identity operator.

Then

(1) for any $b \in H$ there exists a unique $u^* \in H$ such that $A(u^*) = b$

(2) if $u_0 \in H$ is arbitrary, then there exist constants m_0, M_0 $(M_0 \ge m_0 > 0)$ and a constant $\mu_0 > 0$ such that if

$$u_{n+1} := u_n - \frac{2}{M_0 + m_0}(A(u_n) - b)$$

then

$$\|u_n - u^*\| \le \frac{1}{\mu_0}\,\|A(u_0) - b\| \cdot \left(\frac{M_0 - m_0}{M_0 + m_0}\right)^n \qquad (n = 0, 1, 2, \ldots)$$

PROOF. First we prove the existence and uniqueness of the solution. By Proposition 2.13, (i) and (ii) implies that A is a potential operator with potential $J\colon H \to \mathbb{R}$. Let $b \in H$ fixed and let $\Phi(u) := J(u) - \langle b, u\rangle \quad (u \in H)$. We will show that Φ is strictly convex and $\lim_{\|u\|\to\infty} \Phi(u) = \infty$.
In fact,

$$\Phi'(u) = A(u) - b \quad \text{and} \quad \Phi''(u) = A'(u) > 0,$$

thus, by Proposition 2.8, Φ is strictly convex. Using Proposition 2.3, there exists $\theta \in [0,1]$ such that

$$\Phi(u) = \Phi(0) + \Phi'(0)u + \frac{1}{2}\langle \Phi''(\theta u)u, u\rangle .$$

Due to the fact that $\|\theta u\| \in [0, \|u\|]$ and (iii)

$$\Phi(u) \geq \Phi(0) - \|\Phi'(0)\| \cdot \|u\| + \frac{1}{2}m(\|u\|)\cdot \|u\|^2 =$$

$$= \Phi(0) + \|u\| \left(-\Phi'(0) + \frac{1}{2}m(\|u\|)\cdot\|u\| \right).$$

Using the limit condition in (iii), the last term tends to infinity, as $\|u\| \to \infty$. Hence—by Proposition 2.14—Φ attains its unique minimum in X. Finally, Proposition 2.12 shows that this Φ is a minimizing functional with respect to the equation $A(u) = b$.

Now the convergence estimate will be proved. Let $u_0 \in H$ be arbitrary. Because of the limit condition in (iii) again, there exists $r_0 > 0$ such that

$$m(\|u_0\| + r_0)\cdot r_0 > \|\Phi'(u_0)\| .$$

Let B_0 denote the closed ball in H with centre u_0 and radius r_0. Further, let

$$m_0 := m(\|u_0\| + r_0),$$

$$M_0 := M(\|u_0\| + r_0),$$

$$t_0 := \frac{2}{M_0 + m_0},$$

$$Q_0 := \frac{M_0 - m_0}{M_0 + M_0}.$$

Then by the definition of u_{n+1}

$$u_{n+1} = u_n - t_0 \cdot \Phi'(u_n).$$

Now we show that if $u, v \in B_0$ and $v = u - t_0 \cdot \Phi'(u)$, then

$$\|\Phi'(v)\| \leq Q_0 \cdot \|\Phi'(u)\| \tag{3.1}$$

The proof goes the same as in Lemma 2.17 with $C := \Phi'$. It is easy to see that this operator C satisfies all requirements of the Lemma except (2.1). Instead, a weaker version of this condition holds. This time

$$Lx = x - t_0 \cdot \int_0^1 \Phi''(u + t(v-u))x\mathrm{d}t$$

but $u+t(v-u) \in [u,v] \subset B_0$, and B_0 is contained in the ball centered on the origin with radius $\|u_0\| + r_0$. Hence

$$m_0 = m(\|u_0\| + r_0) \leq m(\|u + t(v-u)\|) \tag{3.2}$$

and similarly

$$M(\|u + t(v-u)\|) \leq M(\|u_0\| + r_0) = M_0.$$

This implies

$$m_0 \cdot \|x\|^2 \leq \langle \Phi''(u+t(v-u))x, x\rangle \leq M_0 \cdot \|x\|^2 \quad (x \in H)$$

using (iii). Except for this small modification, the proof of Lemma 2.17 is unchanged. Finally we find that $\|L\| \leq \frac{M_0 - m_0}{M_0 + M_0}$. Hence $\Phi'(v) = L\Phi'(u)$ implies $\|\Phi'(v)\| \leq \|L\| \cdot \|\Phi'(u)\|$, which completes the proof of (3.1).

Inequality (3.1) will be used to verify that $u_n \in B_0 \quad (n \in \mathbb{N})$. We prove by induction. For $n = 0$ it is trivial from the definition of B_0. Suppose $u_0, \ldots, u_n \in B_0$. Then

$$\|u_{n+1} - u_0\| \leq \sum_{i=0}^{n} \|u_{i+1} - u_i\| = t_0 \cdot \sum_{i=0}^{n} \|\Phi'(u_i)\| \leq t_0 \cdot \sum_{i=0}^{n} Q_0^i \cdot \|\Phi'(u_0)\| \leq$$

$$\leq t_0 \cdot \|\Phi'(u_0)\| \cdot \frac{1}{1-Q_0} = \|\Phi'(u_0)\| \cdot \frac{2}{M_0 + m_0} \cdot \frac{1}{1 - \frac{M_0 - m_0}{M_0 + M_0}} = \frac{\|\Phi'(u_0)\|}{m_0} \leq r_0,$$

that is $u_{n+1} \in B_0$.

Finally, since $\Phi'(u^*) = 0$ holds for the unique solution u^*, for $n \in \mathbb{N}$ we have

$$\langle \Phi'(u_n) - \Phi'(u^*), u_n - u^*\rangle \leq \|\Phi'(u_n)\| \cdot \|u_n - u^*\| \leq Q_0^n \cdot \|\Phi'(u_0)\| \cdot \|u_n - u^*\|.$$

On the other hand, from Proposition 2.5 we get

$$\langle \Phi'(u_n) - \Phi'(u^*), u_n - u^*\rangle = \int_0^1 \langle \Phi''(u^* + t(u_n - u^*))(u_n - u^*), u_n - u^*\rangle \mathrm{d}t \geq$$

$$\geq \min\{m_0, m(\|u^*\|)\} \cdot \|u_n - u^*\|^2.$$

(The inequality depends on the same argument as (3.2), due to the fact that $u^* + t(u_n - u^*) \in [u_n, u^*]$, this time by comparing $\|u_0\| + r_0$ and $\|u^*\|$.) Let $\mu_0 := \min\{m_0, m(\|u^*\|)\}$. Hence we conclude

$$\mu_0 \cdot \|u_n - u^*\|^2 \leq Q_0^n \cdot \|\Phi'(u_0)\| \cdot \|u_n - u^*\| \quad (n \in \mathbb{N})$$

which was to be proved. ■

The next Remark immediately follows from the definition of μ_0.

Remark 3.2 *In the case of M and m being constant functions,*

$$M_0 := M, m_0 := m \quad \textit{and} \quad \mu_0 := m$$

is an appropriate choice (and independent of u_0).

The criterion of Gâteaux differentiability in this theorem is a rather strong one. Differential operators in general fail to be Gâteaux differentiable, so the above theorem can not be applied to them directly. To overcome the difficulties, we extend the previous result to (nonlinear) operators which themselves are not Gâteaux differentiable but suitably can be transformed into Gâteaux differentiable ones. The gradient method applied to discontinuous linear operators in Hilbert spaces is described in [4] and [10, Chapter XV]. For the nonlinear uniformly monotone case, see [11]; the nonlinear non-uniformly monotone case can be found in [12]. The following theorem is essentially presented there.

Theorem 3.3 *Let $(H, \langle\cdot,\cdot\rangle)$ be a real Hilbert space and $D \subset H$ be a dense subspace. Let $T\colon D \to H$ be a nonlinear (and not necessarily continuous) operator. Let $B\colon D \to H$ be a linear operator such that*

$$ran\, B \supset ran\, T \quad and \quad B \geq \lambda \cdot I,$$

where $\lambda > 0$ is a constant and I is the identity operator. Moreover suppose, that

(i) there exists a (nonlinear) operator $A\colon H_B \to H_B$ such that $A|_D = B^{-1}T$, where H_B denotes the corresponding energy space

(ii) A is Gâteaux differentiable and A' is bihemicontinuous

(iii) $A'(u)$ is self-adjoint $(u \in H_B)$

(iv) there exist functions $m, M\colon [0,\infty) \to [0,\infty)$, m is monotone decreasing, M is monotone increasing and $\lim_{r\to\infty} r \cdot m(r) = \infty$ such that for all $u, h \in H_B$

$$m(\|u\|) \cdot \|h\|^2 \leq \langle A'(u)h, h\rangle \leq M(\|u\|) \cdot \|h\|^2 .$$

Then

(1) for any $g \in ran\, B$ there exists a unique $u^* \in H_B$ such that

$$\langle A(u^*), v\rangle_B = \langle g, v\rangle \quad (\forall v \in H_B),$$

where $\langle\cdot,\cdot\rangle_B$ denotes the scalar product in the energy space. (Moreover, if $g \in ran\, T$, then $T(u^*) = g$.)

(2) if $u_0 \in D$ is arbitrary, then there exist constants m_0, M_0 $(M_0 \geq m_0 > 0)$ and a constant $\nu_0 > 0$ such that if

$$u_{n+1} := u_n - \frac{2}{M_0 + m_0} B^{-1}(T(u_n) - g)$$

then

$$\|u_n - u^*\|_B \leq \nu_0 \cdot \| T(u_0) - g\| \cdot \left(\frac{M_0 - m_0}{M_0 + m_0}\right)^n \qquad (n = 0, 1, 2, \ldots)$$

PROOF. Let $g \in ran\, B$ arbitrary. Define $b := B^{-1}g$. Then $b \in D \subset H_B$. Now apply the previous theorem to the operator $A: H_B \to H_B$ and the vector $b \in H_B$. (H_B plays the role of H in Theorem 3.1.) Hence we get a unique $u^* \in H_B$ such that $A(u^*) = b$. This implies

$$\langle A(u^*), v\rangle_B = \langle b, v\rangle_B = \langle g, v\rangle \quad (\forall v \in H_B).$$

One easily checks that if $g \in ran\, T$, then $u^* \in D$, thus $T(u^*) = g$.

Now let $u_0 \in D$ be arbitrary. Theorem 3.1 also yields

$$\|u_n - u^*\|_B \leq \frac{1}{\mu_0} \|A(u_0) - b\|_B \cdot \left(\frac{M_0 - m_0}{M_0 + m_0}\right)^n \qquad (n = 0, 1, 2, \ldots),$$

if

$$u_{n+1} := u_n - \frac{2}{M_0 + m_0}(A(u_n) - b).$$

It follows by induction that $u_n \in D \ \ (n \in \mathbb{N})$. In fact, since $ran\, B^{-1} \subset D$ and $A|_D = B^{-1}T$, then

$$A(u_n) - b = B^{-1}(T(u_n) - g) \in D,$$

hence $u_{n+1} \in D$.
For the convergence estimate use Proposition 2.16:

$$\|A(u_0) - b\|_B = \left\|B^{-1}(T(u_0) - g)\right\|_B \leq \frac{1}{\sqrt{\lambda}} \cdot \|T(u_0) - g\|.$$

This shows that $\nu_0 := \frac{1}{\sqrt{\lambda} \cdot \mu_0}$ is appropriate. ■

Finally, the following Remark is easy to be seen from the definition of ν_0 and from Remark 3.2.

Remark 3.4 *If the functions M and m are constant, then*

$$M_0 := M, m_0 := m \quad \textit{and} \quad \nu_0 := \frac{1}{\sqrt{\lambda} \cdot m}$$

can be chosen independently of u_0.

Chapter 4

Application to nonlinear elliptic partial differential equations

In this chapter the previous results will be applied to a certain class of nonlinear elliptic partial differential equations. (Other sources, where applications of the Hilbert space gradient method to nonlinear elliptic problems are described are [11] (theoretical iteration for the uniformly monotone case), [13] (for the non-uniformly monotone case) and for the numerical iteration see [6].) Now Theorem 3.3 will be used in the following setting.

Let $\Omega \subset \mathbb{R}^N$ be a bounded domain which is C^2–diffeomorphic to a convex one and let $H := \Big(L^2(\Omega), \langle \cdot, \cdot \rangle_{L^2(\Omega)}\Big)$, where $\langle f, g\rangle_{L^2(\Omega)} := \int_\Omega fg \quad (f, g \in L^2(\Omega))$. (Throughout the thesis $L^2(\Omega)$ is considered as a real Hilbert space.) We define a differential operator T with domain

$$dom\ T := D := H^2(\Omega) \cap H_0^1(\Omega).$$

(As usual, $H^2(\Omega)$ and $H_0^1(\Omega)$ denote the corresponding Sobolev spaces.)

Let

$$T(u) := -div\, f(x, \nabla u) \qquad (u \in D),$$

that is

$$(T(u))(x) := -\sum_{i=1}^{N} \partial_{x_i}[f_i(x, \partial_{x_1}u(x), \ldots, \partial_{x_N}u(x))]$$

with $x = (x_1, x_2, \ldots, x_N)$ and $f = (f_1, f_2, \ldots, f_N)$, where f_i's are all real-valued functions, satisfying $f_i \in C^1(\overline{\Omega} \times \mathbb{R}^N)$.

Moreover, suppose that

$$\partial_{p_j} f_i(x, p) = \partial_{p_i} f_j(x, p) \tag{4.1}$$

for all $x \in \overline{\Omega}$ and $p \in \mathbb{R}^N$ $(i, j = 1, 2, \ldots, N)$.

Assume that there exist constants m and M $(0 < m \le M)$ such that

$$m|\zeta|^2 \le \sum_{i,j=1}^{N} \partial_{p_j} f_i(x, p)\zeta_i\zeta_j \le M|\zeta|^2 \tag{4.2}$$

for any $(x,p) \in \overline{\Omega} \times \mathbb{R}^N$ and any $\zeta \in \mathbb{R}^N$.
(We remark that (4.2) is equivalent to demanding that all eigenvalues of $\partial_p f(x,p)$ lie in the interval $[m, M]$, where $\partial_p f(x,p)$ is the $N \times N$ (symmetric) matrix standing $\partial_{p_j} f_i(x,p)$ at the $(i,j)^{\text{th}}$ position.)

Finally, let

$$B := -\Delta, \, dom\ B := D.$$

(Throughout the thesis $(-\Delta)$ is considered as an $L^2(\Omega) \to L^2(\Omega)$ operator with domain $H^2(\Omega) \cap H_0^1(\Omega)$.)

The following boundary value problem will be considered:

$$\left.\begin{aligned} T(u) &= g \\ u|_{\partial\Omega} &= 0 \end{aligned}\right\} \tag{4.3}$$

where $g \in L^2(\Omega)$ is given and $u \in D$ is the unknown function.

First we prove an important lemma (see for example [16]).

Lemma 4.1 *The operator $(-\Delta)$ in $L^2(\Omega)$ is symmetric and for any $w \in H^2(\Omega) \cap H_0^1(\Omega)$ the following inequality holds:*

$$\langle -\Delta w, w \rangle_{L^2(\Omega)} \geq \lambda_1 \|w\|^2_{L^2(\Omega)},$$

where λ_1 is the smallest eigenvalue of $(-\Delta)$ on $H^2(\Omega) \cap H_0^1(\Omega)$.

Proof. Let $u, v \in H^2(\Omega) \cap H_0^1(\Omega)$. Using the Green–formula we have

$$\langle -\Delta u, v \rangle_{L^2(\Omega)} = \int_\Omega (-\Delta u) v = \int_\Omega \langle \nabla u, \nabla v \rangle_{\mathbb{R}^N} - \int_{\partial\Omega} (-\partial_\nu u) v.$$

The last integral vanishes since $v \in H_0^1(\Omega)$, thus the symmetry of the scalar product in $\mathbb{R}^N$ implies the symmetry of $(-\Delta)$.

Now let λ_i and e_i $(i = 1, 2, \ldots)$ denote the eigenvalues and eigenfunctions of $(-\Delta)$ on $H^2(\Omega) \cap H_0^1(\Omega)$, respectively. It is known that $0 < \lambda_1 \leq \lambda_2 \leq \ldots$, and the only accumulation point of the sequence (λ_i) is $(+\infty)$. Moreover, the eigenfunctions form an orthonormal basis and their linear combinations are dense in $L^2(\Omega)$ as well as in $H_0^1(\Omega)$. Let a_i $(i = 1, 2, \ldots)$ be the coefficients of $w \in H^2(\Omega) \cap H_0^1(\Omega)$ in its Fourier series expansion, that is $a_i := \int_\Omega w e_i$. Then $w = \sum_{i=1}^\infty a_i e_i$ in $L^2(\Omega)$–norm and $\|w\|^2_{L^2} = \sum_{i=1}^\infty a_i^2$. Using the continuity of the scalar product and the symmetry of $(-\Delta)$ we have

$$\langle -\Delta w, w \rangle_{L^2} = \left\langle -\Delta w, \sum_{i=1}^\infty a_i e_i \right\rangle_{L^2} = \sum_{i=1}^\infty a_i \langle -\Delta w, e_i \rangle_{L^2} =$$

$$= \sum_{i=1}^\infty a_i \langle w, -\Delta e_i \rangle_{L^2} = \sum_{i=1}^\infty a_i \langle w, \lambda_i e_i \rangle_{L^2} = \sum_{i=1}^\infty a_i \lambda_i \langle w, e_i \rangle_{L^2} = \sum_{i=1}^\infty \lambda_i a_i^2.$$

Since $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_i \leq \ldots$

$$\sum_{i=1}^{\infty} \lambda_i a_i^2 \geq \lambda_1 \sum_{i=1}^{\infty} a_i^2 = \lambda_1 \|w\|_{L^2}^2 ,$$

that is $\langle -\Delta w, w \rangle_{L^2} \geq \lambda_1 \|w\|_{L^2}^2$. ■

Proposition 4.2 *The above choices fulfill the assumptions of Theorem 3.3.*

Sketch of the proof. It is known that D is a dense subspace in the real Hilbert space $L^2(\Omega)$. From the previous Lemma we see that the operator B in $L^2(\Omega)$ and with domain D has a lower bound $\lambda_1 > 0$ (that is $B \geq \lambda_1 \cdot I$, where I is the identity operator). Further, $ran\, T \subset L^2(\Omega)$ and $ran\, B = L^2(\Omega)$ (see [8]), hence $ran\, T \subset ran\, B$.

Let $u \in D$. Using the definition of the scalar product in the energy space and the Green–formula again (as in the proof of the previous Lemma), there holds

$$\langle u, u \rangle_B = \langle Bu, u \rangle_{L^2(\Omega)} = \int_\Omega (-\Delta u)u = \int_\Omega \langle \nabla u, \nabla u \rangle_{\mathbb{R}^N} = \int_\Omega |\nabla u|^2,$$

that is

$$\|u\|_B^2 := \langle u, u \rangle_B = \int_\Omega |\nabla u|^2. \tag{4.4}$$

Here the last expression is a norm in $H_0^1(\Omega)$ too, and $D \subset L^2(\Omega)$ is dense, thus $H_B = H_0^1(\Omega)$.

For the construction of the operator A observe that if $u, v \in D$, then

$$\int_\Omega T(u)v = \int_\Omega -\sum_{i=1}^{N} \partial_{x_i}[f_i(x, \nabla u(x))] \cdot v(x) \mathrm{d}x =$$

$$= \int_\Omega \sum_{i=1}^{N} f_i(x, \nabla u(x)) \cdot \partial_i v(x) \mathrm{d}x = \int_\Omega \langle f(x, \nabla u(x)), \nabla v(x) \rangle_{\mathbb{R}^N} \mathrm{d}x,$$

using the divergence theorem. It is easy to see from Cauchy–Scwartz inequality and from (4.4) that for any fixed $u \in H_0^1(\Omega)$ the following assignment

$$v \mapsto \int_\Omega \langle f(x, \nabla u(x)), \nabla v(x) \rangle_{\mathbb{R}^N} \mathrm{d}x$$

defines a bounded linear functional on $H_0^1(\Omega)$. Hence Riesz's theorem gives a unique vector $A(u) \in H_0^1(\Omega)$ for any $u \in H_0^1(\Omega)$ such that

$$\langle A(u), v \rangle_{H_0^1(\Omega)} = \int_\Omega \langle f(x, \nabla u(x)), \nabla v(x) \rangle_{\mathbb{R}^N} \mathrm{d}x. \tag{4.5}$$

The induced nonlinear operator $A\colon H_0^1(\Omega) \to H_0^1(\Omega)$ has the required property $A|_D = B^{-1}T$, since

$$\langle A(u), v \rangle_{H_0^1(\Omega)} = \langle T(u), v \rangle_{L^2(\Omega)} = \left\langle B^{-1}T(u), v \right\rangle_B$$

for all $v \in D$, which is a dense subspace, and consequently

$$A(u) = B^{-1}T(u)$$

for all $u \in D$.

To prove that A is Gâteaux differentiable and A' is bihemicontinuous, let

$$\partial_p f(x, \nabla u(x))$$

denote the $N \times N$ matrix in which $\partial_{p_j} f_i(x, \nabla u(x))$ stands at the $(i,j)^{\text{th}}$ position. Let $u \in H_0^1(\Omega)$ be arbitrary, then using (4.2) and a similar argument as above we find that the assignment

$$H_0^1(\Omega) \times H_0^1(\Omega) \ni (h, v) \mapsto \int_\Omega \langle \partial_p f(x, \nabla u(x))\, \nabla h(x), \nabla v(x) \rangle_{\mathbb{R}^N} \, \mathrm{d}x \in \mathbb{R}$$

is a bounded bilinear functional. Again, Riesz's theorem proves the existence of an operator $S(u)\colon H_0^1(\Omega) \to H_0^1(\Omega)$ (for each $u \in H_0^1(\Omega)$) for which

$$\langle S(u)h, v \rangle_{H_0^1(\Omega)} = \int_\Omega \langle \partial_p f(x, \nabla u(x))\, \nabla h(x), \nabla v(x) \rangle_{\mathbb{R}^N} \, \mathrm{d}x. \tag{4.6}$$

Now it can be shown that $A'(u) = S(u)$, that is

$$\frac{1}{t} \|A(u+th) - A(u) - tS(u)h\|_{H_0^1(\Omega)} \to 0$$

as $t \to 0 \in \mathbb{R}$ for any fixed $u, h \in H_0^1(\Omega)$; or, since the dual space of $H_0^1(\Omega)$ can be identified with itself, it is equivalent to showing that

$$\frac{1}{t} \sup_{\|v\|_{H_0^1}=1} \langle A(u+th) - A(u) - tS(u)h, v \rangle_{H_0^1(\Omega)} \to 0$$

as $t \to 0$. In fact, writing back (4.5) and (4.6) and using Lagrange's equality, cancellation by t occurs. Then, due to the continuity assumptions imposed on f_i's, the integrands tend to 0 (almost everywhere) as $t \to 0$. Finally, since the integrands can be majorated by functions belonging to $L^1(\Omega)$, Lebesgue's theorem yields the desired result.
By the definition of bihemicontinuity we have to show that

$$\|A'(u + sw_1 + tw_2)h - A'(u)h\|_{H_0^1(\Omega)} \to 0$$

as $(s,t) \to (0,0) \in \mathbb{R}^2$ for any fixed $u, w_1, w_2, h \in H_0^1(\Omega)$. This can be proved using a similar reasoning as above. (For details, see for example [11] and [13].)

It follows from (4.1) that $\partial_p f(x, \nabla u(x))$ is a symmetric matrix, so (4.6) and $A'(u) = S(u)$ implies that $A'(u)$ is self-adjoint ($u \in H_0^1(\Omega)$).

Finally, the monotone functions can be chosen to be the constants m and M in (4.2), because

$$m\,\|h\|^2_{H_0^1(\Omega)} = \int_\Omega m|\nabla h|^2 \le \int_\Omega \langle \partial_p f(x, \nabla u(x))\, \nabla h(x), \nabla h(x) \rangle_{\mathbb{R}^N} \, \mathrm{d}x =$$

$$= \langle A'(u)h, h\rangle_{H_0^1(\Omega)} \leq \int_\Omega M|\nabla h|^2 = M \|h\|^2_{H_0^1(\Omega)}$$

for all $u, h \in H_0^1(\Omega)$. ■

Now applying Theorem 3.3 and taking into consideration Remark 3.4 we get the following

Corollary 4.3 *Consider the boundary value problem*

$$\left.\begin{array}{ll} T(u) & = g \\ u|_{\partial\Omega} & = 0 \end{array}\right\} \tag{4.7}$$

where T is defined as above. Let $g \in L^2(\Omega)$ be arbitrary. Then the problem (4.7) admits a unique weak solution $u^* \in H_0^1(\Omega)$:

$$\langle A(u^*), v\rangle_{H_0^1(\Omega)} = \langle g, v\rangle_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega),$$

that is for all $v \in H_0^1(\Omega)$

$$\int_\Omega \langle f(x, \nabla u^*(x)), \nabla v(x)\rangle_{\mathbb{R}^N} \, \mathrm{d}x = \int_\Omega g(x)v(x)\mathrm{d}x$$

holds. (If $u^* \in H^2(\Omega) \cap H_0^1(\Omega)$, then $T(u^*) = g$.)

Moreover, if $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ is arbitrary, then the following sequence

$$u_{n+1} := u_n - \frac{2}{M+m}(-\Delta)^{-1}(T(u_n) - g) \quad (n \in \mathbb{N})$$

converges to the solution u^*:

$$\|u_n - u^*\|_{H_0^1(\Omega)} \leq \frac{1}{m\sqrt{\lambda_1}} \| T(u_0) - g \|_{L^2(\Omega)} \cdot \left(\frac{M-m}{M+m}\right)^n,$$

where λ_1 is the smallest eigenvalue of $(-\Delta)$ on $H^2(\Omega) \cap H_0^1(\Omega)$. ■

The main advantage of the method so far is that the solution of the nonlinear equation (*i.e.* determining T^{-1}) is reduced to the solution of a sequence of linear equations: the inverse of $(-\Delta)$ is to be calculated in each iteration step, which is—in general—a much easier task.

Chapter 5

The Gradient-Fourier method

In the previous chapter we have constructed—using the gradient method—a sequence u_n which converges to the unique solution u^*. Obtaining the next term u_{n+1} of the sequence requires a linear Poisson equation to be solved at each step. (These equations will be referred to as *auxiliary equations*.) There are various ways to deal with Poisson equations. In this thesis we introduce the *Gradient-Fourier method*: solutions to the auxiliary equations will be determined using the Fourier method.

Formally, solving a Poisson equation is equivalent to applying the inverse of the Laplacian operator to the right hand side of the equation. In the vast majority of cases, however, symbolic computation is not possible, so in general the exact sequence u_n is only theoretically known. A numerical method (now the Fourier method) yields an approximation $\overline{u}_n$ of the sequence. Then our goal is to guarantee that the limit of the approximating sequence $\overline{u}_n$ still lies near the exact solution u^*.

Let us first introduce some notations. For $n = 0, 1, \ldots$, let

$$\begin{aligned} \overline{u}_0 &:= u_0, \\ \overline{u}_{n+1} &:= \overline{u}_n - \frac{2}{M+m}\overline{z}_n, \end{aligned}$$

where $\overline{z}_n$ is the solution to the auxiliary equation obtained by the Fourier method. (As in the previous chapter, homogeneous Dirichlet boundary conditions are imposed on the auxiliary equations.) We also define two other sequences

$$z_n := (-\Delta)^{-1}(T(u_n) - g),$$

$$z_n^* := (-\Delta)^{-1}(T(\overline{u}_n) - g),$$

i.e. z_n is the exact solution to the auxiliary equation with exact (but only theoretically known) right hand side, and z_n^* is the exact solution to the auxiliary equation with the approximate (*i.e.* numerically computed) right hand side. Because of the presence of the inverse of Laplacian, these functions are both unknown, only $\overline{z}_n$ is available in practice. Finally, let

$$E_n := \|\,\overline{u}_n - u_n\|\,.$$

(The unindexed norms are all understood to be $H_0^1(\Omega)$–norms throughout the chapter.)

5.1 Convergence estimate of the approximating sequence

In this section the distance of $\overline{u}_n$ and u^* will be estimated provided that $\overline{z}_n$ and z_n^* remain sufficiently close through the iteration process. The next section will discuss how this can be achieved in the actual construction of $\overline{z}_n$.

First we prove two lemmas. (For the first lemma, see [6].)

Lemma 5.1 *If $\|\overline{z}_n - z_n^*\| \le \delta_n$ for some δ_n $(n \in \mathbb{N})$, then*

$$E_{n+1} \le \frac{M-m}{M+m} E_n + \frac{2}{M+m}\delta_n,$$

where m and M are given in (4.2).

Proof. $\overline{u}_{n+1} - u_{n+1} = \overline{u}_n - \frac{2}{M+m}\overline{z}_n - (u_n - \frac{2}{M+m} z_n) =$

$= \overline{u}_n - u_n - \frac{2}{M+m}(\overline{z}_n - z_n + z_n^* - z_n^*) = \overline{u}_n - u_n - \frac{2}{M+m}(z_n^* - z_n) - \frac{2}{M+m}(\overline{z}_n - z_n^*) =$

$= \overline{u}_n - u_n - \frac{2}{M+m}[(-\Delta)^{-1}(T(\overline{u}_n)) - (-\Delta)^{-1}(T(u_n))] - \frac{2}{M+m}(\overline{z}_n - z_n^*) =$

$= [\overline{u}_n - \frac{2}{M+m}(-\Delta)^{-1}(T(\overline{u}_n))] - [u_n - \frac{2}{M+m}(-\Delta)^{-1}(T(u_n))] - \frac{2}{M+m}(\overline{z}_n - z_n^*).$

Now applying Lemma 2.17 with $C := (-\Delta)^{-1} \circ T$ yields

$E_{n+1} = \|\overline{u}_{n+1} - u_{n+1}\| = \left\|F(\overline{u}_n) - F(u_n) - \frac{2}{M+m}(\overline{z}_n - z_n^*)\right\| \le$

$\le \frac{M-m}{M+m}\|\overline{u}_n - u_n\| + \frac{2}{M+m}\|\overline{z}_n - z_n^*\| \le \frac{M-m}{M+m}E_n + \frac{2}{M+m}\delta_n.$ ■

Lemma 5.2 *Let $\varepsilon > 0$ be arbitrary and $\delta_n := m\varepsilon$ $(n \in \mathbb{N})$, then*

$$E_n \le \varepsilon.$$

Proof. If $n = 0$ then $0 = \|\overline{u}_0 - u_0\| = E_0 \le \varepsilon$. Suppose the statement is true for some $n \in \mathbb{N}$. Then by the previous Lemma we have

$E_{n+1} \le \frac{M-m}{M+m}E_n + \frac{2}{M+m}\delta_n \le \frac{M-m}{M+m}\varepsilon + \frac{2}{M+m}m\varepsilon = \frac{M+m}{M+m}\varepsilon = \varepsilon.$ ■

Corollary 5.3 *Let $\varepsilon > 0$. If $\|\overline{z}_n - z_n^*\| \le m\varepsilon$ $(n \in \mathbb{N})$, then*

$$\|\overline{u}_n - u^*\| \le \varepsilon + \frac{1}{m\sqrt{\lambda_1}}\|T(u_0) - g\|_{L^2(\Omega)} \cdot \left(\frac{M-m}{M+m}\right)^n,$$

where λ_1 is the smallest eigenvalue of $(-\Delta)$ on $H^2(\Omega) \cap H_0^1(\Omega)$.

Proof. $\|\overline{u}_n - u^*\| \le \|\overline{u}_n - u_n\| + \|u_n - u^*\| \le E_n + \frac{1}{m\sqrt{\lambda_1}}\|T(u_0) - g\|_{L^2(\Omega)} \cdot (\frac{M-m}{M+m})^n \le$
$\le \varepsilon + \frac{1}{m\sqrt{\lambda_1}}\|T(u_0) - g\|_{L^2(\Omega)} \cdot (\frac{M-m}{M+m})^n$ using Corollary 4.3 and the previous Lemma. ■

5.2 The Fourier method for the auxiliary equations

Now let us focus on a single iteration step (*i.e.* $n \in \mathbb{N}$ is fixed in the section). The Fourier method implemented here can be found for example in [16].

According to the definition of z_n^*, the auxiliary equation takes the form

$$\left.\begin{aligned} -\Delta z_n^* &= r_n^* \\ z_n^*|_{\partial\Omega} &= 0 \end{aligned}\right\} \tag{5.1}$$

with $r_n^* := T(\overline{u}_n) - g$. Instead of r_n^*, the right hand side will be replaced by its Fourier series approximation $\overline{r}_n$. The modified equation is

$$\left.\begin{aligned} -\Delta \overline{z}_n &= \overline{r}_n \\ \overline{z}_n|_{\partial\Omega} &= 0 \end{aligned}\right\} \tag{5.2}$$

which now can be solved by a formula.

As in the previous chapter, let λ_i and e_i $(i = 1, 2, \dots)$ denote the eigenvalues and eigenfunctions of $(-\Delta)$ on $H^2(\Omega) \cap H_0^1(\Omega)$, respectively.

Let c_i $(i = 1, 2, \dots)$ be the coefficients of r_n^* in its Fourier series expansion, that is

$$c_i := \int_\Omega r_n^* e_i.$$

(Of course $c_i = c_{i,n}$, but we simply use c_i since n is fixed.) Then

$$r_n^* = \sum_{i=1}^{\infty} c_i e_i.$$

Now we define $\overline{r}_n$ as a partial sum of the above infinite series. Let l $(= l(n))$ be a positive integer and

$$\overline{r}_n := \sum_{i=1}^{l} c_i e_i.$$

Define $\overline{z}_n$ as

$$\overline{z}_n := \sum_{i=1}^{l} \frac{c_i}{\lambda_i} e_i.$$

A simple calculation shows that these satisfy (5.2):

$$-\Delta \overline{z}_n = \sum_{i=1}^{l} \frac{c_i}{\lambda_i}(-\Delta e_i) = \sum_{i=1}^{l} \frac{c_i}{\lambda_i} \lambda_i e_i = \sum_{i=1}^{l} c_i e_i = \overline{r}_n.$$

We will now determine how the requirement of Corollary 5.3 can be fulfilled by the appropriate choice of l. The basis for this is that the weak solution to the equation $(-\Delta)(\overline{z}_n - z_n^*) = \overline{r}_n - r_n^*$ depends continuously on the right hand side, *i.e.* there exists a constant c_* such that

$$\| \overline{z}_n - z_n^* \| \leq c_* \| \overline{r}_n - r_n^* \|_{L^2(\Omega)} .$$

A suitable choice for c_* is $\frac{1}{\sqrt{\lambda_1}}$ due to the following

Lemma 5.4 *If*

$$\left. \begin{array}{ll} -\Delta w & = f \\ w|_{\partial\Omega} & = 0 \end{array} \right\}$$

then

$$\|w\| \leq \frac{1}{\sqrt{\lambda_1}} \|f\|_{L^2} .$$

PROOF. If $w \in H^2(\Omega) \cap H_0^1(\Omega)$, then according to the Green–formula

$$\int_\Omega (-\Delta w) w = \int_\Omega |\nabla w|^2,$$

that is $\|w\|^2 = \langle -\Delta w, w \rangle_{L^2}$. Using Lemma 4.1 we get

$$\|w\|^2 = \langle -\Delta w, w \rangle_{L^2} = \langle f, w \rangle_{L^2} \leq \|f\|_{L^2} \|w\|_{L^2} \leq \|f\|_{L^2} \frac{1}{\sqrt{\lambda_1}} \|w\| ,$$

thus

$$\|w\| \leq \frac{1}{\sqrt{\lambda_1}} \|f\|_{L^2} ,$$

which was to be proved. ■

Hence for Corollary 5.3 it is sufficient to guarantee that

$$\|\overline{r}_n - r_n^*\|_{L^2} \leq m\varepsilon\sqrt{\lambda_1}. \tag{5.3}$$

Taking into consideration that

$$\|r_n^* - \overline{r}_n\|_{L^2}^2 = \sum_{i=l+1}^{\infty} |c_i|^2 = \sum_{i=1}^{\infty} |c_i|^2 - \sum_{i=1}^{l} |c_i|^2 = \|r_n^*\|_{L^2}^2 - \|\overline{r}_n\|_{L^2}^2$$

we see that for (5.3) it is sufficient that

$$\|r_n^*\|_{L^2}^2 - \sum_{i=1}^{l} |c_i|^2 \leq \lambda_1 m^2 \varepsilon^2.$$

Recalling the definition of r_n^* we get the following

Corollary 5.5 *If l is a positive integer such that*

$$\sum_{i=1}^{l} |c_i|^2 \geq \| T(\overline{u}_n) - g\|_{L^2}^2 - \lambda_1 m^2 \varepsilon^2,$$

then

$$\| \overline{z}_n - z_n^* \| \leq m\varepsilon.$$

5.3 Final convergence estimate of the algorithm

This section summarizes the results of our thesis. The solution to the given boundary value problem is approximated by an iterative sequence $\overline{u}_n$, which is obtained by the gradient method. Computing each term of the sequence is effectively reduced to a series of numerical integrations via Fourier series expansion. In order to realize the algorithm, one needs to know the eigenvalues and eigenfunctions of the operator $(-\Delta)$ on $H^2(\Omega) \cap H_0^1(\Omega)$. With this information, exact error control can be given in $H_0^1(\Omega)$–norm. Formulating our results leads to the following

Theorem 5.6 *Let $\Omega \subset \mathbb{R}^N$ be the bounded domain and T be the nonlinear elliptic differential operator defined in Chapter 4. The lower and upper bounds of the coefficients of T are denoted by m and M, respectively, as in (4.2). Let $g \in L^2(\Omega)$ be arbitrary. Then the unique weak solution $u^* \in H_0^1(\Omega)$ to the boundary value problem*

$$\left.\begin{aligned} T(u) &= g \\ u|_{\partial\Omega} &= 0 \end{aligned}\right\}$$

can be approximated by the following iterative sequence $\overline{u}_n \in H^2(\Omega) \cap H_0^1(\Omega)$. Let $\overline{u}_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ be arbitrary, and for $n \in \mathbb{N}$

$$\overline{u}_{n+1} := \overline{u}_n - \frac{2}{M+m} \overline{z}_n,$$

where

$$\overline{z}_n := \sum_{i=1}^{l(n)} \frac{c_{i,n}}{\lambda_i} e_i,$$

and $l(n)$ is a suitable positive integer, further λ_i and e_i $(i = 1, 2, \ldots)$ denote the eigenvalues and eigenfunctions of $(-\Delta)$ on $H^2(\Omega) \cap H_0^1(\Omega)$, respectively, and

$$c_{i,n} := \int_\Omega (T(\overline{u}_n) - g)\, e_i.$$

Let $\varepsilon > 0$ be arbitrary. If $l(n)$ is chosen such that

$$\sum_{i=1}^{l(n)} |c_{i,n}|^2 \geq \| T(\overline{u}_n) - g\|^2_{L^2(\Omega)} - \lambda_1 m^2 \varepsilon^2 \quad (n \in \mathbb{N}),$$

then

$$\| \overline{u}_n - u^* \|_{H_0^1(\Omega)} \leq \varepsilon + \frac{1}{m\sqrt{\lambda_1}} \| T(\overline{u}_0) - g \|_{L^2(\Omega)} \cdot \left(\frac{M-m}{M+m} \right)^n .$$

PROOF. The proof simply follows from applying Corollary 5.5 and Corollary 5.3. ■

5.4 An example

Finally we briefly mention an example, which arises in plasticity theory (for details, see [6]).

Let Ω be the unit square in $\mathbb{R}^2$, that is $\Omega := [0,1] \times [0,1]$. Define the operator T as

$$T(u) := -div(\overline{g}(|\nabla u|^2)\nabla u),$$

$$u|_{\partial\Omega} = 0,$$

where $\overline{g}$ is a given scalar-valued function (the so-called strain-stress function) defined in a finite interval, and $u = u(x_1, x_2)$ is the unknown function. For example, $\overline{g}$ can have the form

$$\overline{g}(t) := \frac{constant}{1 + \sqrt{1 - \frac{t}{3}}},$$

if $0 \le t \le t_0$, with a suitable $t_0 < 3$ and $\overline{g}(t) := \overline{g}(t_0)$, if $t > t_0$.

Then—if the right hand side is also given—we can obtain approximate solutions to the boundary value problem using the previous Theorem and the sine-system as eigenfunctions on the unit square.

Bibliography

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

[2] Axelsson, O., Karátson, J., *Double Sobolev gradient preconditioning for elliptic problems*, Report 0016, Dept. Math., Univ. Nijmegen, April 2000.

[3] Ciarlet, Ph., *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[4] Czách, L., *The steepest descent method for elliptic differential equations* (in Russian), C.Sc. thesis, 1955.

[5] Daniel, J.W., *The conjugate gradient method for linear and nonlinear operator equations*, SIAM J. Numer. Anal., 4 (1967), No. 1., pp. 10–26.

[6] Faragó, I., Karátson, J., *The gradient-finite element method for elliptic problems*, to appear in Comp. Math. Appl.

[7] Gajewski, H. Gröger, K., Zacharias, K., *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974.

[8] Hackbusch, W., *Elliptic Differential Equations (Theory and Numerical Treatment)*, Springer, 1992.

[9] Kantorovich, L. V., *On an effective method of the solution of the extremal problem of the quadratic functional* (in Russian), Dokl. Akad. Nauk. SSSR, 48 (1945), pp. 483–487.

[10] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, 1982.

[11] 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.

[12] Karátson, J., *Gradient method for non-uniformly convex functionals in Hilbert space*, to appear in Pure Math. Appl.

[13] Karátson, J., *On the gradient method for some nonlinear elliptic boundary value problems*, to appear in Acta Math. Hung.

[14] Neuberger, J.W., *Sobolev gradients and differential equations*, Lecture Notes in Math., No. 1670, Springer, 1997.

[15] Ortega, J.M., Rheinboldt, W.C., *Iterative solution of nonlinear equations in several variables*, Academic Press, 1970.

[16] Simon, L. Baderko, E., *Linear Partial Differential Equations of Second Order* (in Hungarian), Tankönyvkiadó, Budapest, 1983.

[17] Stoyan, G., *Numerical methods* (in Hungarian), ELTE Typotex, 1993.

[18] Stoyan, G., *Optimal iterative Stokes solvers in the harmonic Velte subspace*, Report 99-07, SFB F013, University Linz, 1999.

[19] Vainberg, M., *Variational Method and the Method of Monotone Operators in the Theory of Nonlinear Equations*, J. Wiley, New York, 1973.

[20] Zeidler, E., *Nonlinear functional analysis and its applications*, Springer, 1986.