Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER Computers and Mathematics with Applications 50 (2005) 1093–1104
An International Journal
computers & mathematics with applications
www.elsevier.com/locate/camwa

Sobolev Gradient Preconditioning for the Electrostatic Potential Equation

J. KARÁTSON
Department of Applied Analysis, ELTE University
H-1518 Budapest, Hungary
karatson@cs.elte.hu

L. LÓCZI
Department of Numerical Analysis, ELTE University
H-1518 Budapest, Hungary
lloczi@math.bme.hu

Abstract—Sobolev gradient type preconditioning is proposed for the numerical solution of the electrostatic potential equation. A constructive representation of the gradients leads to efficient Laplacian preconditioners in the iteration thanks to the available fast Poisson solvers. Convergence is then verified for the corresponding sequence in Sobolev space, implying mesh independent convergence results for the discretized problems. A particular study is devoted to the case of a ball: due to the radial symmetry of this domain, a direct realization without discretization is feasible. The simplicity of the algorithm and the fast linear convergence are finally illustrated in a numerical test example.

1. INTRODUCTION

We consider the numerical solution of the problem

$$\begin{aligned} T(u) \equiv -\Delta u + e^u &= 0, \\ u|_{\partial\Omega} &= 0, \end{aligned} \tag{1}$$

on a bounded domain $\Omega \subset \mathbb{R}^3$ being C^2-diffeomorphic to a convex one. The function u describes the electrostatic potential in Ω (see e.g. [1]). In particular, we develop a direct realization when the domain Ω is a ball. We remark however that analogues of our method can easily be formulated for other nonlinearities as well, satisfying similar growth conditions to those of e^u.

We propose Sobolev gradient type preconditioning for problem (1) with the auxiliary problems in the iteration containing the Laplacian as a preconditioner. For discretizations of (1), this method defines the corresponding discrete Laplacian as preconditioning matrix, which has proved its efficiency in many applications. For linear problems, discrete Laplacian preconditioners were first used in [2,3] with finite-difference discretization on a rectangle, and hence, the corresponding steepest descent (or Richardson) iteration was later termed as D'yakonov-Gunn iteration. The

This research was supported by the Hungarian Research Fund OTKA under Grant No. F034840, further, the first author by the Hungarian post-doc scholarship Magyary Zoltán and the second author by the Hungarian Research Fund OTKA under Grant No. T037491.

0898-1221/05/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2005.08.011

efficiency of this iteration is based on the fast Poisson solvers developed in the same period. The extension of fast solvers to more general problems led to iterations where the discrete Laplacian is modified by scaling or adding another term (see e.g. [4,5]). A general study of the related conditioning properties has been given in [6].

For nonlinear problems, the Sobolev gradient approach has been developed in a series of publications of Neuberger and summarized in [7]. In this approach the iteration is constructed as a steepest descent method for a suitable minimizing—generally least-square—functional. The main principle is that convergence can be improved by using the Sobolev inner product instead of the original L^2 one, which leads to auxiliary discrete Poisson or Helmholtz equations. Some illustrative applications are summarized in [8], also containing the van Roosbroeck type modification of problem (1) in one dimension with nonlinearity $2\sinh u$. The authors' related earlier results include finite element and Fourier series realization [9,10] and also extend to systems and higher-order problems [11,12]; general results on preconditioning are summarized in [13].

In our proposal, we introduce a Sobolev gradient type iteration for the convex potential that corresponds to problem (1). Suitable regularity leads to a constructive representation of the gradients that involves Laplacian preconditioners in the iteration. The theoretical sequence in Sobolev space not only defines iterations for the discretizations of (1) in an obvious way, but also provides mesh independent convergence results. We study in particular the case when the domain Ω is a ball. For this special radially symmetric problem, we develop a direct realization which uses no discretization but actual Sobolev space preconditioning. The main advantage of our method is the simplicity of the algorithm, whose straightforward coding as well as the obtained fast linear convergence in a numerical test example are enclosed.

2. THE ITERATION WITH LAPLACIAN PRECONDITIONING

We first sketch the background of the iterative solution to be proposed for problem (1).

The solution of (1) can be obtained by minimizing the convex functional $\Psi : H_0^1(\Omega) \to \mathbb{R}$

$$\Psi(u) \equiv \int_\Omega \left(\frac{1}{2}|\nabla u|^2 + e^u\right) \tag{2}$$

using the Sobolev gradient idea on the continuous level, that is by defining a steepest descent iteration for Ψ in the Sobolev space $H_0^1(\Omega)$ equipped with the H_0^1-inner product

$$\langle u, v\rangle_{H_0^1} := \int_\Omega \nabla u \cdot \nabla v. \tag{3}$$

Vanishing of the derivative of Ψ

$$\langle \Psi'(u), v\rangle_{H_0^1} = \int_\Omega (\nabla u \cdot \nabla v + e^u v)$$

corresponds to the weak solution of our problem. According to Green's formula, for regular functions $u \in H^2(\Omega) \cap H_0^1(\Omega)$

$$\langle \Psi'(u), v\rangle_{H_0^1} = \int_\Omega T(u)v = \langle -\Delta^{-1}T(u), v\rangle_{H_0^1},$$

i.e., we have the constructive Sobolev gradient

$$\Psi'|_{H^2\cap H_0^1} = -\Delta^{-1}T. \tag{4}$$

Using this decomposition, the steepest descent iteration yields a sequence where the Laplacian acts as a preconditioner:

$$u_{n+1} = u_n - \alpha_n(-\Delta)^{-1}T(u_n)$$

with some steplengths α_n. We will consider optimal constant steplengths.

We underline that suitable numerical iterations can be derived directly from the above iteration by projecting it into the considered discretization subspace.

To commence the rigorous formulation of the construction and convergence of the proposed iteration, we rely on the authors' earlier results in [9,11]. Let us first cite a regularity result for the Laplacian, required for a constructive representation on the domain Ω and also in the proof of Theorem 2.

THEOREM 1. *(See [14].) Let Ω be C^2-diffeomorphic to a convex domain. Then for any $g \in L^2(\Omega)$ the unique weak solution $u^* \in H_0^1(\Omega)$ of the problem*

$$\begin{aligned} -\Delta u &= g, \\ u|_{\partial\Omega} &= 0, \end{aligned} \tag{5}$$

satisfies $u^* \in H^2(\Omega)$.

The well-posedness of problem (1) will be verified together with the construction of the iteration and the proof of the convergence. For this purpose, introduce the function

$$f(u) := \begin{cases} e^u, & (u \le 0), \\ 1+u, & (u > 0), \end{cases}$$

because if $u \in H^2(\Omega) \cap H_0^1(\Omega)$ solves (1), then $\Delta u \ge 0$ and the maximum principle [15] implies that $u \le 0$ a.e., and consequently, (1) is equivalent to

$$\begin{aligned} -\Delta u + f(u) &= 0, \\ u|_{\partial\Omega} &= 0, \end{aligned} \tag{6}$$

where the inequality $0 \le f'(u) \le 1$ gives a linear growth bound for the nonlinearity.

THEOREM 2.

(1) *Problem (6) has a unique weak solution $u^* \in H_0^1(\Omega)$, moreover, $u^* \in H^2(\Omega)$.*
(2) *For any $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, the sequence $(u_n) \subset H^2(\Omega) \cap H_0^1(\Omega)$ defined by*

$$u_{n+1} = u_n - \frac{2\varrho}{2\varrho+1} z_n, \qquad \text{where } -\Delta z_n = -\Delta u_n + f(u_n), \quad z_{n|\partial\Omega} = 0, \tag{7}$$

and $\varrho > 0$ is the smallest eigenvalue of $-\Delta$ on $H^2(\Omega) \cap H_0^1(\Omega)$, converges linearly to u^*, namely,

$$\|u_n - u^*\|_{H_0^1(\Omega)} \le \varrho^{-1/2} \, \|-\Delta u_0 + f(u_0)\|_{L^2(\Omega)} \left(\frac{1}{2\varrho+1}\right)^n, \qquad (n \in \mathbb{N}). \tag{8}$$

PROOF. Problem (6) is a special case of problem (1.1) in [9] with $g(x,\eta) = \eta$, $q(x,u) = e^u$ and $f(x) = 0$, hence, both the well-posedness and the convergence result follow from that paper with appropriate modifications: namely, by its Section 2 the problem has a unique weak solution $u^* \in H_0^1(\Omega)$. The regularity $u^* \in H^2(\Omega)$ follows from Theorem 1 above by setting $g = -e^{u^*} \in L^2(\Omega)$. Further, the ellipticity bounds m and M in (2.2) of [9] now satisfy

$$m = 1 \qquad \text{and} \qquad M = 1 + \varrho^{-1}, \tag{9}$$

with $\varrho > 0$ being the smallest eigenvalue of $-\Delta$ on $H^2(\Omega) \cap H_0^1(\Omega)$. Then Theorem 2.1 of [9] yields the required convergence result, using that

$$\frac{2}{M+m} = \frac{2\varrho}{2\varrho+1} \qquad \text{and} \qquad \frac{M-m}{M+m} = \frac{1}{2\varrho+1}. \tag{10}$$

(In fact, in that theorem Ω is assumed to be convex or satisfy $\partial\Omega \in C^2$, but Theorem 1 also ensures the same result for this more general Ω.) ∎

We note that by setting $w_n := z_n - u_n$, the iteration (7) takes the simpler form

$$u_{n+1} = \frac{1}{2\varrho+1}(u_n - 2\varrho w_n), \tag{11}$$

$$\text{where } -\Delta w_n = f(u_n), \qquad w_{n|\partial\Omega} = 0. \tag{12}$$

Now return to problem (1) with the original nonlinearity e^u, equivalent to (6) as verified before. By the maximum principle again, we get $w_n \geq 0$ in (12), and letting $u_0 \leq 0$, we have by induction that $u_n \leq 0$, for all $n \in \mathbb{N}$. Hence, $f(u_n)$ in (12) can be replaced again by the original e^{u_n}.

COROLLARY 1.

(1) *Problem (1) has a unique weak solution* $u^* \in H_0^1(\Omega)$, *moreover,* $u^* \in H^2(\Omega)$.

(2) *Let* $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, $u_0 \leq 0$. *Then the sequence* $(u_n) \subset H^2(\Omega) \cap H_0^1(\Omega)$ *defined by*

$$\begin{aligned} u_{n+1} &= \frac{1}{2\varrho+1}(u_n - 2\varrho w_n), \\ \text{where } -\Delta w_n &= e^{u_n}, \qquad w_{n|\partial\Omega} = 0, \end{aligned} \tag{13}$$

and $\varrho > 0$ *is the smallest eigenvalue of* $-\Delta$ *on* $H^2(\Omega) \cap H_0^1(\Omega)$, *converges linearly to* u^*, *namely,*

$$\|u_n - u^*\|_{H_0^1(\Omega)} \leq \varrho^{-1/2} \|-\Delta u_0 + e^{u_0}\|_{L^2(\Omega)} \left(\frac{1}{2\varrho+1}\right)^n, \qquad (n \in \mathbb{N}). \tag{14}$$

Now let us turn to the discretizations of problem (1):

$$T_h(u) \equiv -\Delta_h u + e_h(u) = 0 \tag{15}$$

in some finite-dimensional subspace $V_h \subset H_0^1(\Omega)$, where Δ_h stands for the discrete Laplacian and the nonlinear function $e_h : V_h \to V_h$ corresponds to the discretization of e^u in (1). An iteration for (15) can be obtained by projecting the theoretical sequence (13) into V_h: the sequence $(u_n) \subset V_h$ is defined by

$$\begin{aligned} u_{n+1} &= \frac{1}{2\varrho+1}(u_n - 2\varrho w_n), \\ \text{where } -\Delta_h w_n &= e_h(u_n), \qquad w_{n|\partial\Omega} = 0. \end{aligned} \tag{16}$$

The convergence of the sequence (16) is asymptotically the same as that of (13) as $h \to 0$. In particular, for FEM discretizations the convergence factor $1/(2\varrho+1)$ is an upper asymptotic bound for the convergence since the proof can be repeated in any FEM subspace with the same parameters. Therefore, the proposed iteration provides mesh independent convergence.

Efficiency of the Laplacian preconditioners is justified by the various available fast Poisson solvers developed in the past decades. Many of these were originally introduced for rectangular domains, then extended to other domains via the 'fictitious domain approach'. Comprehensive summaries on the fast direct solution of the Poisson equation—including the method of cyclic reduction, the fast Fourier transform and the FACR algorithm—are found in [16,17]. Parallel implementation of these algorithms is also feasible [18,19]. Another family of fast solvers on rectangles consists of the spectral methods [20,21] developed further recently. For the fictitious domain approach, see [22].

REMARK 1. COMPARISON WITH OTHER METHODS.

(i) Iteration (13) can be considered as an improvement of the corresponding 'linearized' iteration

$$\Delta u_{n+1} = e^{u_n}, \qquad u_{n+1|\partial\Omega} = 0, \tag{17}$$

realizing another frequently used approach for semilinear problems. Namely, the limiting case $\varrho \to \infty$ in (13) would give $u_{n+1} = -w_n$ turning (13) into the linearized iteration (17). Nevertheless, ϱ as the smallest eigenvalue of $-\Delta$ is fixed, and u_{n+1} in (13) is the proper convex combination of u_n and $-w_n$ that—owing to (10)—provides the optimal linear convergence quotient $1/(2\varrho+1)$ for the steepest descent iteration.

(ii) Superlinear convergence could be achieved by using Newton's method for (1). The price of this is twofold. First, instead of a fixed Laplacian the auxiliary equations need to be redefined in each step:

$$-\Delta w_n + e^{u_n} w_n = \Delta u_n - e^{u_n}, \qquad w_{n|\partial\Omega} = 0, \tag{18}$$

hence, in the discretized case the matrix of the auxiliary system has to be updated stepwise. On the other hand, one cannot use the above-mentioned fast solvers directly for problems in (18), hence, the cost required to solve them is no more negligible. Still, in order to take advantage of the fast solvers, the Laplacian could be applied as a preconditioner in the inner iterations for solving the problems in (18). However, the obtained overall iteration would then consist of Poisson equations, as (13) does, and it is easily seen that the order of required iterations to achieve a prescribed error would be the same as in the case of the steepest descent iteration, namely, $n = \mathcal{O}(\log \varepsilon)$ as the prescribed error $\varepsilon \to 0$.

3. DIRECT PRECONDITIONING ON A BALL

Now we consider problem (1) on the ball $B = B(0,R) \subset \mathbb{R}^3$ with radius R:

$$\begin{aligned} -\Delta u + e^u &= 0, \\ u|_{\partial B} &= 0. \end{aligned} \tag{19}$$

By [15], the unique weak solution of problem (19) is classical, i.e., $u^* \in C^2(\bar{B})$. Moreover, u^* is radially symmetric [23].

3.1. The Proposed Method

Thanks to the special form of the problem, a direct approach becomes possible which avoids discretization, instead it realizes actual Sobolev space preconditioning with the iteration being applied directly in the Sobolev space $H_0^1(B)$ and kept in the class of radially symmetric polynomials

$$\mathcal{P} = \left\{ \sum_{m=0}^{l} a_m r^{2m} : l \in \mathbb{N},\ a_m \in \mathbb{R} \right\}, \qquad \text{with } r = |x|, \quad \text{for } x \in B,$$

where the Laplacian can be inverted exactly.

In each step of (13), we approximate e^{u_n} by a suitable Taylor polynomial

$$p(u_n) = \sum_{j=0}^{k_n} \frac{u_n^j}{j!}, \tag{20}$$

and define the subsequent iterate by

$$u_{n+1} = \frac{1}{2\varrho+1}\left(u_n - 2\varrho w_n\right), \tag{21}$$

$$\text{where } -\Delta w_n = p(u_n), \qquad w_{n|\partial B} = 0. \tag{22}$$

Note that $u_n \in \mathcal{P}$ implies $p(u_n) \in \mathcal{P}$. Further, if

$$p(u_n)(r) = \sum_{m=0}^{l_n} a_m r^{2m}, \qquad (r \in [-R, R]), \tag{23}$$

then (22) is equivalent to

$$-\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial w_n}{\partial r}\right) = \sum_{m=0}^{l_n} a_m r^{2m}, \qquad w_n(-R) = w_n(R) = 0,$$

thus, its solution $w_n \in \mathcal{P}$ is given explicitly by

$$w_n(r) = \sum_{m=0}^{l_n} \frac{a_m}{(2m+3)(2m+2)} \left(R^{2m+2} - r^{2m+2}\right). \tag{24}$$

Arguing inductively, if $u_0 \in \mathcal{P}$ and $u_0 \leq 0$, then $u_n \in \mathcal{P}$ for all $n \in \mathbb{N}$ and the Poisson equations (22) are solved by (24).

Recall that by Corollary 1 the theoretical iteration (11),(12) converges according to the estimate (14). Now since the first positive root of the spherical Bessel function $j_0(r) = \sin r/r$ is π (see [24]), we have that the smallest eigenvalue $\varrho > 0$ of $-\Delta$ on $H^2(B) \cap H_0^1(B)$ is explicitly

$$\varrho = \left(\frac{\pi}{R}\right)^2. \tag{25}$$

Further, for the sake of simplicity choose $u_0 \equiv 0$, which implies $\|-\Delta u_0 + e^{u_0}\|_{L^2(B)} = |B|^{1/2}$, where $|B| = 4R^3\pi/3$ is the volume of B.

Convergence of the iteration (21),(22) is achieved by suitably choosing $p(u_n)$. To this end, given u_n, let w_n^* and w_n denote the solution of (13) and (22), respectively. Then, we have the estimate

$$\|w_n^* - w_n\|_{H_0^1(B)} \leq \varrho^{-1/2} \|e^{u_n} - p(u_n)\|_{L^2(B)} \leq \left(\frac{|B|}{\varrho}\right)^{1/2} \|e^{u_n} - p(u_n)\|_\infty \leq \left(\frac{|B|}{\varrho}\right)^{1/2} \frac{\|u_n\|_\infty^{k_n+1}}{(k_n+1)!}.$$

Further, since $u_n = z_n - w_n$ is already a polynomial, it is easy to see that

$$\|z_n^* - z_n\|_{H_0^1(B)} = \|w_n^* - w_n\|_{H_0^1(B)},$$

where z_n^* and z_n are the solutions of (7) and its polynomial approximation, respectively. Hence,

$$\|z_n^* - z_n\|_{H_0^1(B)} \leq \left(\frac{|B|}{\varrho}\right)^{1/2} \frac{\|u_n\|_\infty^{k_n+1}}{(k_n+1)!}. \tag{26}$$

Then, we use Lemma 3.2 from [10] to ensure a prescribed accuracy $\varepsilon > 0$ (independent of n) throughout the iteration: if one defines the steepest descent iteration for an elliptic operator with ellipticity bounds m and M and with the optimal steplength $2/(M+m)$, then an accuracy $m\varepsilon$ of the correction terms z_n yields accuracy ε for the sequence (u_n). In our case, choose the indices k_n, such that the right-hand side of (26) is bounded by some fixed ε throughout the iteration. Since by (9) $m = 1$, the iteration (21),(22) satisfies (14) up to accuracy ε:

$$\|u_n - u^*\|_{H_0^1(B)} \leq \left(\frac{|B|}{\varrho}\right)^{1/2} \left(\frac{1}{2\varrho+1}\right)^n + \varepsilon, \qquad (n \in \mathbb{N}). \tag{27}$$

REMARK 2. We may experience a rapid growth in the degrees of the polynomials u_n in the above iteration together with high-index coefficients becoming very small in magnitude, so dropping

these small enough terms within some given accuracy $\delta > 0$ may spare a considerable amount of memory and computing time. The truncation would only result in the addition of δ to the right-hand side of (27). Indeed, supposing u_n has the form

$$u_n(r) = \sum_{m=0}^{s_n} a_m r^{2m}, \tag{28}$$

elementary integration yields for any index $t_n \le s_n$ that

$$\left\| \sum_{m=t_n+1}^{s_n} a_m r^{2m} \right\|^2_{H_0^1(B)} \le 4\pi s_n \sum_{m=t_n+1}^{s_n} 4m^2 a_m^2 \frac{R^{4m+1}}{4m+1}. \tag{29}$$

Hence, letting $t_n \le s_n$ be the smallest index for which

$$4\pi s_n \sum_{m=t_n+1}^{s_n} \left(2m a_m R^{2m}\right)^2 \frac{R}{4m+1} \le \delta^2 \tag{30}$$

and defining

$$\bar{u}_n(r) = \sum_{m=0}^{t_n} a_m r^{2m}, \tag{31}$$

we obtain the estimate

$$\|u_n - \bar{u}_n\|_{H_0^1(B)} \le \delta. \tag{32}$$

3.2. The Algorithmic Form of the Iteration

Now we summarize our method in an algorithmic form. First define

$$\varrho = \left(\frac{\pi}{R}\right)^2, \tag{33}$$

then fix a tolerance $\varepsilon > 0$, i.e., the accuracy of the algorithm in $H_0^1(B)$ norm, and let

$$\omega = \varepsilon \left(\frac{\varrho}{|B|}\right)^{1/2}, \tag{34}$$

with $|B| = 4R^3\pi/3$ being the volume of B.

The proposed method constructs a sequence of radial polynomials via (20),(22) with the indices k_n being chosen in such a way that the right-hand side of (26) is bounded by ε in each step: see (b1) in (35) and the definition of ω. The algorithm reads as follows:

(a) $u_0 \equiv 0$;

for any $n \in \mathbb{N}$: if $u_n \in \mathcal{P}$ has been obtained, then let

(b1) $\mu_n = \max_B |u_n|$;

$k_n \in \mathbb{N}$ be the smallest number, such that $\dfrac{\mu_n^{k_n+1}}{(k_n+1)!} \le \omega$;

(b2) $p(u_n)(r) = \displaystyle\sum_{j=0}^{k_n} \frac{u_n(r)^j}{j!} \qquad (r \in [-R, R])$; (35)

(b3) $w_n \in \mathcal{P}$ be the solution of the problem

$-\Delta w_n = p(u_n), \quad w_{n|\partial B} = 0$

according to formula (24);

(b4) $u_{n+1} = \dfrac{1}{2\varrho+1}\left(u_n - 2\varrho w_n\right).$

We emphasize that the auxiliary Poisson equations in Step (b3) are solved exactly by (24). The convergence of the algorithm has been given by (27).

As noted in Remark 2, the polynomials u_n may contain a large number of high-index terms with almost zero coefficients. These terms are automatically dropped when their coefficients are below the roundoff accuracy. However, we may spare memory as well as computing time by also dropping some of the small terms larger than the roundoff accuracy. Therefore, our algorithm (35) is completed by the steps below: using the polynomial form of u_{n+1} in Step (b4), (b5) finds the decreased truncation index t_{n+1} and (b6) redefines u_{n+1} using only the terms with indices up to t_{n+1}. (For simplicity, the notation u_{n+1} is kept for the truncated polynomial also.) The corresponding error estimate is obtained by adding the truncation accuracy δ to the right-hand side of (27).

(b4) $$u_{n+1}(r) = \sum_{m=0}^{s_{n+1}} a_m r^{2m},$$

(b5) $t_{n+1} \in \mathbb{N}$ be the smallest index, such that

$$4\pi s_{n+1} \sum_{m=t_{n+1}+1}^{s_{n+1}} \left(2ma_m R^{2m}\right)^2 \frac{R}{4m+1} \le \delta^2; \tag{36}$$

(b6) $$u_{n+1}(r) = \sum_{m=0}^{t_{n+1}} a_m r^{2m}.$$

4. AN EXAMPLE

We have used MATHEMATICA[1] as a working environment. The test results correspond to problem (19) on the ball $B = B(0,2) \subset \mathbb{R}^3$ with radius $R = 2$, using the truncated version of the algorithm (b1)–(b6) in (35) and (36) with the prescribed tolerance being equal to the truncation accuracy

$$\delta = \varepsilon = 10^{-6}.$$

Now (33) and (34) read $\rho \approx 2.4674$ and $\omega \approx 2.7135 \cdot 10^{-7}$, hence, the convergence factor in the estimate (27) is

$$\frac{1}{2\rho+1} \approx 0.168498.$$

The proposed choice $u_0 \equiv 0$ yields $\mu_0 = k_0 = 0$ in Step (b1) of (35), so a consistent definition of $0^0 := 1$ is needed before starting the iteration. The brevity and simplicity of the code shows that the implementation of the direct gradient method involves no difficulty.

(a) Using the predefined constants, further the initial data

$$u_0[r\_] = 0;\ \mu_0 = 0; k_0 = 0;$$

$$Pu_0[r\_] = 1;\ w_0[r\_] = \frac{1}{6}\left(R^2 - r^2\right);\ u_1[r\_] = \frac{1}{2\rho+1}(u_0[r] - 2\rho w_0[r]);$$

(b) the actual code reads

```
Do[
μiter=-FindMinimum[-Abs[uiter[r]], {r,R/100,-R/200,-R,R}][[1]];
 kiter=0; While[μiter^(kiter+1)/(kiter+1)! > ω, kiter++];
Puiter[r_] =Sum[(uiter[r])^j/j!, {j,0,kiter}];
coefflist=CoefficientList[Puiter[r],r^2]//N;
```

[1]Copyright 1988-2000 Wolfram Research, Inc.

```
length=Length[coefflist];
w_iter[r_]=-Sum[coefflist[[m+1]]/((2m+3)(2m+2))(r^(2m+2)-R^(2m+2)), {m,0,length-1}];
u_iter+1[r_]=1/(2ρ+1)(u_iter[r]-2ρw_iter[r])//Expand;
chopcoefflist=CoefficientList[u_iter+1[r],r^2]//N;
choplength=Length[chopcoefflist];
δ=ε;
counter=choplength-1; tailsum=0;
While[counter≥0 && 4πchoplengthtailsum ≤ δ^2,
    tailsum += (2counterchopcoefflist[[counter+1]]R^2counter)^2 R/(4counter+1);
    counter--];
counter += 2;
u_iter+1[r_]=u_iter+1[r][[Range[counter]]],
{iter,1,iter_max}]
```

Here $\texttt{iter}_{\max}$ is the maximal number of iterations allowed, which can be determined by the error estimates. For this purpose, we estimate the theoretical errors

$$E_n = \|u_n - u^*\|_{H_0^1(B)}$$

in two ways. First, by adding the truncation error $\delta = \varepsilon$ to (27), an *a priori* estimate yields

$$E_n \le e_n \equiv \left(\frac{|B|}{\varrho}\right)^{1/2} \left(\frac{1}{2\varrho+1}\right)^n + 2\varepsilon.$$

Further, using the lower ellipticity bound $m = 1$ of the operator $T(u) = -\Delta u + e^u$, the corresponding residual estimate (cf. [10, Lemma 3.4]) implies an *a posteriori* residual estimate for E_n:

$$\|u_n - u^*\|_{H_0^1(B)} \le r_n \equiv \varrho^{-1/2} \| - \Delta u_n + e^{u_n}\|_{L^2(B)}.$$

Using the data of our example, we obtain

$$E_n \le \min\{e_n, r_n\} \qquad (n \in \mathbb{N})$$

with

$$e_n = 3.6853 \cdot 0.168498^n + 2 \cdot 10^{-6} \tag{38}$$

and

$$r_n = 0.6366 \cdot \|-\Delta u_n + e^{u_n}\|_{L^2(B)}. \tag{39}$$

We also calculate, as usual, the norm of the difference of two consecutive terms

$$d_n \equiv \|u_{n+1} - u_n\|_{H_0^1(B)}. \tag{40}$$

The following table lists these error estimates during the iteration. Observe that the residual error is smaller than the *a priori* estimate e_n. It first decreases below the accuracy $\varepsilon = 10^{-6}$ in Step 9, then it is stabilized slightly below that value.

Table 1. The errors e_n, r_n and d_n, defined in (38), (39), (40), respectively.

n	0	1	2	3	4	5	6
e_n	3.6853	0.6209	0.1046	0.01763	0.002972	$5.025 \cdot 10^{-4}$	$8.634 \cdot 10^{-5}$
r_n	3.6853	0.4298	0.0505	0.00696	0.001082	$1.717 \cdot 10^{-4}$	$2.826 \cdot 10^{-5}$
d_n	2.4856	0.2740	0.0194	0.00168	0.000179	$2.323 \cdot 10^{-5}$	$3.319 \cdot 10^{-6}$

n	7	8	9	10	11	12
e_n	$1.621 \cdot 10^{-5}$	$4.394 \cdot 10^{-6}$	$2.403 \cdot 10^{-6}$	$2.067 \cdot 10^{-6}$	$2.011 \cdot 10^{-6}$	$2.001 \cdot 10^{-6}$
r_n	$5.183 \cdot 10^{-6}$	$1.457 \cdot 10^{-6}$	$8.551 \cdot 10^{-7}$	$7.573 \cdot 10^{-7}$	$7.413 \cdot 10^{-7}$	$7.387 \cdot 10^{-7}$
d_n	$4.953 \cdot 10^{-7}$	$7.549 \cdot 10^{-8}$	$1.165 \cdot 10^{-8}$	$1.813 \cdot 10^{-9}$	$2.839 \cdot 10^{-10}$	—

Table 2. The coefficients of the iterative sequence.

	u_1	u_2	.....	u_{11}	u_{12}
a_0	-0.554335	-0.472618		-0.475685	-0.475685
a_1	0.138584	0.102961		0.103577	0.103577
a_2		0.003309		0.003218	0.003218
a_3		0.000109		0.000127	0.000127
a_4		$2.944 \cdot 10^{-6}$		$5.572 \cdot 10^{-6}$	$5.572 \cdot 10^{-6}$
a_5		$6.632 \cdot 10^{-8}$		$2.596 \cdot 10^{-7}$	$2.596 \cdot 10^{-7}$
a_6		$1.360 \cdot 10^{-9}$		$1.258 \cdot 10^{-8}$	$1.258 \cdot 10^{-8}$
a_7		$1.736 \cdot 10^{-11}$		$6.264 \cdot 10^{-10}$	$6.264 \cdot 10^{-10}$
a_8		$5.954 \cdot 10^{-13}$		$3.181 \cdot 10^{-11}$	$3.181 \cdot 10^{-11}$
a_9				$1.641 \cdot 10^{-12}$	$1.641 \cdot 10^{-12}$
a_{10}				$8.576 \cdot 10^{-14}$	$8.576 \cdot 10^{-14}$
a_{11}				$4.524 \cdot 10^{-15}$	$4.524 \cdot 10^{-15}$

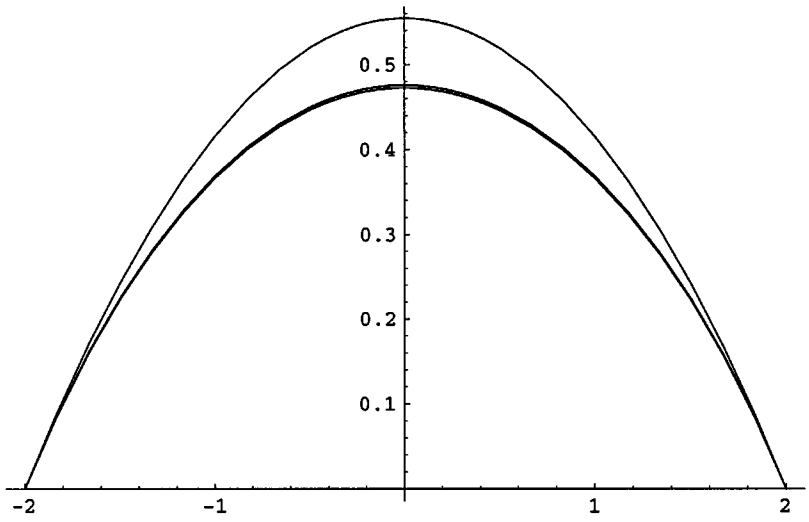

Figure 1. The first few terms of the sequence $-u_n(r)$.

Coefficients of the iterative sequence $u_n(r) = \sum_{m=0}^{s_n} a_m r^{2m}$ ($r \in [-R, R]$) consisting of radial polynomials of r are given in Table 2 for $n = 1, 2$ and $n = 11, 12$, respectively.

Figure 1 contains graphs of the first few terms of this sequence converging rapidly. (In fact, for the sake of positivity, the functions $-u_n(r)$ are plotted instead.)

Observe that u_{11} and u_{12} in Table 2 coincide up to the accuracy $\varepsilon = 10^{-6}$, moreover, since the corresponding residual error $r_{12} < 10^{-6}$, we accept

$$u_{12} \approx u^*$$

as the numerical solution. In order to better visualize the graph of u^* over B, one dimension is omitted in Figure 2 by plotting the surface of the 2D function which attains the values of u_{12} along the radii. (Again, for the sake of positivity, its modulus is plotted instead.)

Conclusions of the Experiment

In the example, we have realized direct Laplacian preconditioning for problem (19) on a ball. The proposed method uses no discretization but realizes actual Sobolev space preconditioning for the iteration directly in the space $H_0^1(B)$, due to keeping the iteration in the class of radially symmetric polynomials where the Laplacian is exactly invertible. The main advantage of this

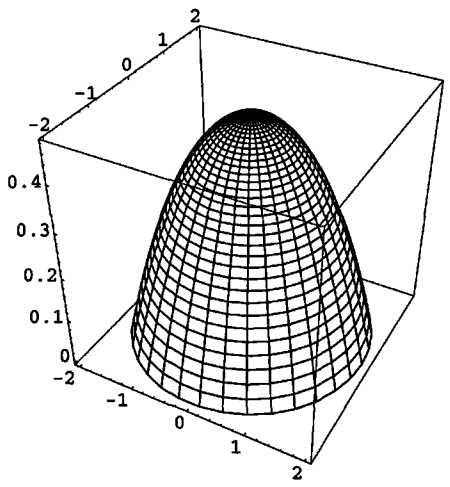

Figure 2. Graph of the modulus of the numerical solution u_{12}.

method is the simplicity of the algorithm (35) and (36), whose straightforward coding and the obtained fast linear convergence have been presented in this numerical test example.

REFERENCES

1. M. Křižek and P. Neittaanmäki, *Mathematical and Numerical Modelling in Electrical Engineering: Theory and Applications*, Kluwer Academic Publishers, (1996).
2. E.G. D'yakonov, On an iterative method for the solution of finite difference equations (in Russian), *Dokl. Akad. Nauk SSSR* **138**, 522–525, (1961).
3. J.E. Gunn, The numerical solution of $\nabla \cdot a\nabla u = f$ by a semi-explicit alternating direction iterative method, *Numer. Math.* **6**, 181–184, (1964).
4. P. Concus and G.H. Golub, Use of fast direct methods for the efficient numerical solution of nonseparable elliptic equations, *SIAM J. Numer. Anal.* **10**, 1103–1120, (1973).
5. A. Greenbaum, Diagonal scalings of the Laplacian as preconditioners for other elliptic differential operators, *SIAM J. Matrix Anal. Appl.*, **13**, 826–846, (1992).
6. V. Faber, T. Manteuffel and S.V. Parter, On the theory of equivalent operators and application to the numerical solution of uniformly ellipticpartial dofferential equations, *Adv. in Appl. Math.*, **11**, 109–163, (1990).
7. J.W. Neuberger, *Sobolev Gradients and Differential Equations, Lecture Notes in Math., Volume 1670*, Springer, (1997).
8. W.B. Richardson, Jr., Sobolev gradient preconditioning for PDE applications, In *Iterative Methods in Scientific Computation IV*, (Edited by D.R. Kincaid and A.C. Elster), pp. 223–234, IMACS, New Jersey, (1999).
9. I. Faragó and J. Karátson, The gradient finite-element method for elliptic problems, *Computers Math. Applic.* **42** (8/9), 1043–1053, (2001).
10. L. Lóczi, The Gradient-Fourier method for nonlinear elliptic partial differential equations in Sobolev space, *Annales Univ. Sci. ELTE* **43**, 139–149, (2000).
11. J. Karátson, The gradient method for non-differentiable operators in product Hilbert spaces and applications to elliptic systems of quasilinear differential equations,, *J. Appl. Anal.* **3** (2), 205–217, (1997).
12. J. Karátson, Sobolev space preconditioning of strongly nonlinear 4th order elliptic problems, In *Numerical Analysis and Its Applications, Lecture Notes Comp. Sci., Volume 1988*, Sec. Int. Conf. NAA 2000, Rousse, Bulgaria, (Edited by L. Vulkov, J. Wasniewski and P. Yalamov), pp. 459–46 6, Springer, (2001).
13. I. Faragó and J. Karátson, Numerical solution of nonlinear elliptic problems via preconditioning operators: Theory and applications, *Advances in Computation, Volume 11*, NOVA Science Publishers, New York, (2002).
14. J. Kadlec, On the regularity of the solution of the Poisson problem on a domain with boundary locally similar to the boundary of a convex open set, *Czechosl. Math. J.* **89** (14), 386–393, (1964).
15. M. Struwe, *Variational Methods. Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, Springer-Verlag, Berlin, (1990).
16. F.W. Dorr, The direct solution of the discrete Poisson equation on a rectangle, *SIAM Rev.* **12**, 248–263, (1970).
17. P.N. Swarztrauber, The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle, *SIAM Rev.* **19** (3), 490–501, (1977).
18. T. Rossi and J. Toivanen, A parallel fast direct solver for block tridiagonal systems with separable matrices of arbitrary dimension, *SIAM J. Sci. Comput.* (electronic) **20** (5), 1778–1796 electronic, (1999).

19. P.S. Vassilevski, R.D. Lazarov and S.D. Margenov, Vector and parallel algorithms in iteration methods for elliptic problems, In *Mathematics and Mathematical Education* (Albena, 1989), pp. 40–51, Bulgar. Akad. Nauk, Sofia, (1989).
20. J.P. Boyd, *Chebyshev and Fourier Spectral Methods*, Second Edition, Dover Publications, Mineola, (2001).
21. D. Funaro, *Polynomial Approximation of Differential Equations, Lecture Notes in Physics, New Series, Monographs 8*, Springer, (1992).
22. C. Börgers and O.B. Widlund, On finite element domain imbedding methods, *SIAM J. Numer. Anal.* **27** (4), 963–978, (1990).
23. B. Gidas, W.N. Ni and L. Nirenberg, Symmetry and related properties via the maximum principle, *Commun. Math. Phys.* **68**, 209–243, (1979).
24. M. Abramowitz and C.A. Stegun, Editors, Spherical Bessel functions, In *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Ninth Printing, pp. 437–442, Dover, New York, (1972).