JOURNAL OF COMPUTATIONAL PHYSICS 98, 26–32 (1992)

Enclosures of the Solution of the Thomas–Fermi Equation by Monotone Discretization

C. GROSSMANN

Dresden University of Technology, Institute of Numerical Mathematics, Mommsenstr. 13, D-O-8027 Dresden, Germany

Received January 5, 1990; revised July 13, 1990

In this paper a new discretization concept is proposed which generates uniform lower and upper bounds of the solution of the Thomas–Fermi equation. Their convexity is used to modify the occurring nonlinearity such that the right-hand side originally given is bounded from below and from above by piecewise tangents and secants, respectively.

1. INTRODUCTION

The Thomas–Fermi equation plays an important role in mathematical physics. Various authors (see [2–6, 10, 11], e.g.) have investigated the analysis of the problem and have proposed numerical methods approximating the solution. Especially monotone iteration techniques play an important role to derive upper and lower bounds, respectively, of the solution. In the fundamental paper [6] the original problem has been generalized and the existence of its solution has been shown by monotonically convergent algorithms which produce lower and upper bounds of the wanted solution. However, the implementation on computers additionally requires numerical integrations and spline interpolations to perform the iterations which base on some specific fixed point iteration where the related linear part has a known representation of its solution by means of modified Bessel functions.

The aim of the present paper consists in proposing a new algorithm for generating bounds for the solution of the Thomas–Fermi equation. This algorithm rests on an extension of the concept of monotone discretization recently proposed (see [8, 9]). This includes firstly a modification of the original monotone discretization method to the weakly singular case occurring in the Thomas–Fermi equation and, second, the use of convexity information w.r.t. the solution to simplify the construction of the bounding operators used in monotone discretization techniques. The method proposed here bases on piecewise simplifications of the nonlinearity. Thus, this technique can be implemented on computers without using special functions as Bessel functions. Furthermore, no numerical integration is needed to realize the method because the occurring integrals can be calculated explicitly. The problem can be reduced to a sequence of tridiagonal linear systems of finite dimension.

In this paper we deal with the original Thomas–Fermi equation

$$y'' = x^{-1/2} y^{3/2} \tag{1}$$

with the boundary condition

$$y(0) = 1, \qquad y(a) = 0. \tag{2}$$

Here a denotes some given positive constant. The boundary value problem (1), (2) occurs in the investigation of potentials and charge densities of ionized atoms. Existence, uniqueness and smoothness properties of the solution of (1), (2) have been shown in various publications (see [2, 6, 11], e.g.).

The idea of the method presented here rests on a simplification of the nonlinear part of the differential equation (1) in such a way that:

- the related auxiliary problems can be solved explicitly in some fixed finite dimensional space,
- the solution of the auxiliary problems forms a lower and an upper solution, respectively, of the original problem (1), (2).

2. MONOTONE DISCRETIZATION OF THE PROBLEM

In the sequel we base our approach on the weak formulation of the original problem which is an appropriate tool for handling discontinuous right-hand sides generated in the monotone discretization technique.

Let us abbreviate $\Omega := (0, a)$ the given interval of the boundary value problem (1), (2). The Sobolev space $H^1(\Omega)$

0021-9991/92 $3.00

denotes the set of all functions being quadratically integrable and possessing a quadratically integrable generalized first-order derivative, i.e.,

$$H^1(\Omega) := \left\{ u : \int_\Omega u^2(x)\,dx + \int_\Omega u'^2(x)\,dx < +\infty \right\}$$

equipped with the norm

$$\|u\| := \left[\int_\Omega u^2(x)\,dx + \int_\Omega u'^2(x)\,dx \right]^{1/2}.$$

Here the generalized derivative u' of the function u is defined via the identity

$$\int_\Omega u(x)\,v'(x)\,dx = -\int_\Omega u'(x)\,v(x)\,dx \qquad \text{for any} \quad v \in C^1(\bar{\Omega}) \text{ with } v(0) = v(a) = 0. \tag{3}$$

Especially $H_0^1(\Omega)$ denotes the subspace of $H^1(\Omega)$ formed by those functions having vanishing traces at the boundary points, i.e.,

$$H_0^1(\Omega) := \{ u \in H^1(\Omega) : u(0) = u(a) = 0 \}.$$

Throughout this paper we use in $H_0^1(\Omega)$ the same norm as in $H^1(\Omega)$. An equivalent norm in $H_0^1(\Omega)$ often applied in the literature is given by

$$|||u||| := \left[\int_\Omega u'^2(x)\,dx \right]^{1/2}.$$

A mapping l: $H_0^1(\Omega) \to R$ is called a linear functional on $H_0^1(\Omega)$ if

$$\langle l, \alpha u + \beta z \rangle = \alpha \langle l, u \rangle + \beta \langle l, z \rangle \qquad \text{for any} \quad u, z \in H_0^1(\Omega), \quad \alpha, \beta \in R$$

holds. Here $\langle l, u \rangle$ denotes the value of the functional l at the argument u. The linear functional l is called continuous if some constant $c > 0$ exists such that

$$|\langle l, u \rangle| \leqslant c\,\|u\| \qquad \text{for any} \quad u \in H_0^1(\Omega)$$

holds. The set of all continuous linear functionals on $H_0^1(\Omega)$ equipped with the norm

$$\|u\|_* := \sup_{u \in H_0^1(\Omega), u \neq 0} \frac{\langle l, u \rangle}{\|u\|} \tag{4}$$

is denoted by $H^{-1}(\Omega)$ the so-called dual space to $H_0^1(\Omega)$. In the sequel we abbreviate $V := H^{-1}(\Omega)$.

Let U denote the linear manifold defined by

$$U := \{ u \in H^1(\Omega) : u(0) = 1, u(a) = 0 \}.$$

We define operators L, G: $U \to V$ by

$$\langle Lu, v \rangle := \int_\Omega u'(x)\,v'(x)\,dx \qquad \text{for any} \quad u \in U, \quad v \in H_0^1(\Omega)$$

and

$$\langle Gu, v \rangle := \int_\Omega x^{-1/2} [u(x)]_+^{3/2}\,v(x)\,dx \qquad \text{for any} \quad u \in U, \quad v \in H_0^1(\Omega).$$

There $[\cdot]_+$ denotes the positive part, i.e.,

$$[t]_+ := \max\{0, t\} \qquad \text{for any} \quad t \in R$$

and $\langle \cdot, \cdot \rangle$ is the dual pairing. The operator $L + G$ has the following properties:

- $\langle (L+G)u - (L+G)v, u - v \rangle \geqslant \gamma\,\|u - v\|^2$ for any $u, v \in U$ with some $\gamma > 0$;
- $(L+G)u \leqslant (L+G)v$, $u, v \in U$ implies $u \leqslant v$.

Throughout this paper "$\leqslant$" denotes the natural semi-ordering in $H^1(\Omega)$ and the related dual semi-ordering in V, i.e.,

$$u, v \in H^1(\Omega): u \leqslant v \Leftrightarrow u(x) \leqslant v(x) \qquad \text{a.e. in } \Omega$$

and

$$l, f \in H^{-1}(\Omega): l \leqslant f \Leftrightarrow \langle l, v \rangle \leqslant \langle f, v \rangle \qquad \text{for any} \quad v \in H_0^1(\Omega) v \geqslant 0,$$

respectively.

The first property is the coerciveness of $L + G$ and the second one characterizes $L + G$ to be of a monotone type.

The operator L^{-1} has good smoothing properties in the sense that

$$Ly = f, f \in L_2(\Omega) \qquad \text{implies} \quad y \in H^2(\Omega). \tag{5}$$

Here $L_2(\Omega)$ and $H^2(\Omega)$ denote the Lebesgue space of quadratically integrable functions and the Sobolev space with quadratically integrable generalized derivatives up to the second order, respectively. The relationship (5) is a direct consequence of (3) and of the definition of the operator L.

Using the continuous embedding $H^2(\Omega) \subsetneq C^1(\bar{\Omega})$ can restrict us to the manifold $W := U \cap C^1(\bar{\Omega})$ and we obtain the operator equation

$$y \in W \qquad (L+G)\,y = 0 \tag{6}$$

as the weak formulation related to the originally given boundary value problem (1), (2).

The classical solution of (1), (2) which has been shown to exist in [6], e.g., forms a solution of the weak formulation also. Because the operator $L+G$ is coercive, this solution is unique. Thus, the given two-point boundary value problem (1), (2) is equivalent to the operator equation (6).

The main idea of monotone discretization consists in reducing problem (6) to appropriate auxiliary problems which are solvable explicitly with piecewise analytical representations. Let us select some grid $\{x_i\}_{i=0}^N$ on Ω, i.e.,

$$0=x_0<x_1<\cdots<x_{N-1}<x_N=a.$$

We denote $h_i:=x_i-x_{i-1}$, $\Omega_i:=(x_{i-1},x_i)$, $i=1(1)N$. The mesh size h of the grid is given by

$$h:=\max_{1\leqslant i\leqslant N} h_i.$$

We select a finite dimensional subspace $V_h\subset V$ by

$$V_h:=\operatorname{lin}\{\zeta_{ij}\}_{i=1,\ j=0}^{N\quad 1}$$

with

$$\zeta_{ij}(x):=\begin{cases} x^{-1/2}(x_i-x)^j, & x\in\Omega_i,\\ 0, & \text{otherwise,}\end{cases}$$

$i=1(1)N$, $j=0,1$. Thus, the functions $v\in V_h$ can be represented by

$$v(x)=\sum_{i=1}^{N}\sum_{j=0}^{1} w_{ij}\zeta_{ij}(x)\qquad \text{a.e. in } \Omega \tag{7}$$

with some $w_{ij}\in R$, $i=1(1)N$, $j=0,1$.

In monotone discretization (compare [8, 9]) we replace the nonlinear operator $G\colon U\to V$ by appropriate bounding operators $\underline{G}_h$, $\bar{G}_h\colon W\subset U\to V_h\subset V$, respectively. Then the auxiliary problems

$$\underline{y}_h,\bar{y}_h\in W:\qquad (L+\underline{G}_h)\,\underline{y}_h=0,\qquad (L+\bar{G}_h)\,\bar{y}_h=0, \tag{8}$$

are to be solved instead of (6). We define

$$[\underline{G}_h u](x):=x^{-1/2}\left([u_i]_+^{3/2}-\frac{[u_i]_+^{3/2}-[u_{i-1}]_+^{3/2}}{h_i}(x_i-x)\right) \tag{9}$$

and

$$[\bar{G}_h u](x):=x^{-1/2}([u_i]_+^{3/2}-\tfrac{3}{2}[u_i]_+^{1/2}[u_i']_-(x_i-x)), \tag{10}$$

for any $x\in\Omega_i$.

Here denote $[t]_-:=\min\{0,t\}$, $t\in R$, and $u_i:=u(x_i)$, $u_i':=u'(x_i)$, $i=1(1)N$.

A first proposal for a piecewise linearization has been made in [13]. In the present paper we concentrate our attention to the special kind of weakly singular boundary value problems and to algorithmic realizations. Furthermore, we investigate the convergence of the proposed monotone discretization method for $h\to+0$.

The existence and the local uniqueness of solution $\underline{y}_h$, $\bar{y}_h$ of the auxiliary problems (8) for sufficiently small $h>0$ can be shown by the technique used in [8].

THEOREM 1. *Let $\underline{y}_h,\bar{y}_h\in W$ denote solutions of the auxiliary problems* (8). *Then the estimations*

$$\underline{y}_h\leqslant y\leqslant \bar{y}_h$$

hold for the solution y of the original problem (6).

Proof. The definitions (9), (10) of the bounding operators $\underline{G}_h$, $\bar{G}_h$ result in

$$\underline{G}_h u\geqslant 0,\qquad \bar{G}_h u\geqslant 0\qquad \text{for any}\quad u\in W.$$

With (8) and with the definition of the operator L this leads to the convexity of the functions $\underline{y}_h$, $\bar{y}_h$. Because the functions $[\cdot]_+$ and $(\cdot)^{3/2}$ are convex and monotone nondecreasing the superpositions $[\underline{y}_h]_+^{3/2}$, $[\bar{y}_h]_+^{3/2}$ are convex functions, also. Now, using (9) we obtain

$$\underline{G}_h\underline{y}_h\geqslant G\underline{y}_h. \tag{11}$$

Because of $[\bar{y}_h]_+^{3/2}\geqslant 0$ and $[\bar{y}_h]_+^{3/2}(a)=0$ we have

$$([\bar{y}_h]_+^{3/2})'(a)\leqslant 0.$$

With the convexity of $[\bar{y}_h]_+^{3/2}$ this leads to

$$([\bar{y}_h]_+^{3/2})'\leqslant 0\qquad \text{on } \Omega.$$

Thus, the identity

$$[\bar{G}_h\bar{y}_h](x)=x^{-1/2}([\bar{y}_i]_+^{3/2}-\tfrac{3}{2}[\bar{y}_i]_+^{1/2}[\bar{y}_i']_-(x_i-x)),\qquad x\in\Omega_i$$

holds with $\bar{y}_i=\bar{y}_h(x_i)$, $\bar{y}_i'=\bar{y}_h'(x_i)$. Because of the known structure of the solution $\bar{y}_h$ (compare Section 3) its first derivative $\bar{y}_h'$ is explicitly available and has not to be approximated numerically.

Using the convexity of $[\bar{y}_h]_+^{3/2}$, now, we obtain

$$G\bar{y}_h\geqslant \bar{G}_h\bar{y}_h.$$

With (8), (11) this leads to

$$(L+G)\underline{y}_h \leqslant (L+\underline{G}_h)\underline{y}_h = 0 = (L+G)y = (L+\bar{G}_h)\bar{y}_h \leqslant (L+G)\bar{y}_h.$$

Because the operator $L+G$ is of monotone kind this completes the proof.

Remark. Basing on the monotonicity of $\underline{y}_h, \bar{y}_h$ shown in the proof above one can select

$$[\underline{G}_h u](x) := x^{-1/2}[u_{i-1}]_+^{3/2}, \qquad x \in \Omega_i \tag{12}$$

$$[\bar{G}_h u](x) := x^{-1/2}[u_i]_+^{3/2}, \qquad x \in \Omega_i \tag{13}$$

as simple bounding operators approximating G with the order $O(h)$.

Next, we investigate the order of approximation of the solution y of (6) by $\underline{y}_h, \bar{y}_h$.

THEOREM 2. *There exists some $c>0$ such that*

$$\|\bar{y}_h - \underline{y}_h\| \leqslant ch^2$$

holds for sufficiently small $h>0$.

Proof. From (8) we obtain

$$0 = (L+\underline{G}_h)\underline{y}_h = (L+G)\underline{y}_h + (\underline{G}_h - G)\underline{y}_h.$$

With (6) and the coercivity of the operator $L+G$ this results in

$$\|y - \underline{y}_h\| \leqslant \frac{1}{\gamma}\|(\underline{G}_h - G)\underline{y}_h\|_*. \tag{14}$$

Here $\|\cdot\|_*$ denotes the norm in V induced from $H^1(\Omega)$ defined by (4). The sequence $\{\underline{y}_h\}$ can be shown to be bounded (see [8], e.g.). Using (8), the smoothing property (5) of L^{-1}, and the continuous embedding $H^2(\Omega) \hookrightarrow C^{1+1/2}(\bar{\Omega})$ (compare [1]), we obtain

$$|\underline{y}_h'(\xi) - \underline{y}_h'(\eta)| \leqslant c_1 h^{1/2} \qquad \text{for any} \quad \xi, \eta \in \Omega_i,\ i = 1(1)N,$$

with some $c_1>0$. Taking the definition (9) of $\underline{G}_h$ into account and integrating by parts this leads to

$$|\langle(\underline{G}_h - G)\underline{y}_h, v\rangle| \leqslant c_2 h^2 \|v\| \qquad \text{for any} \quad v \in V$$

with some $c_2>0$. Thus, we have

$$\|(\underline{G}_h - G)\underline{y}_h\|_* \leqslant c_2 h^2.$$

Estimation (14) and the analogue result for the upper solution $\bar{y}_h$ prove the statement of the theorem.

3. FINITE-DIMENSIONAL REPRESENTATION

In this section we deal with the finite-dimensional representation of the solutions $\underline{y}_h, \bar{y}_h$ of the auxiliary problem (8). Let G_h, y_h denote $\underline{G}_h, \underline{y}_h$ or $\bar{G}_h, \bar{y}_h$, respectively. We abbreviate $g_h := G_h y_h$. Because of $G_h: U \to V_h$ we have $g_h \in V_h$ and due to (7) the function g_h can be represented by

$$g_h(x) = \sum_{i=1}^{N}\sum_{j=0}^{1} w_{ij}\zeta_{ij}(x). \tag{15}$$

On the other hand, the solution y_h of (6) solves the linear problem

$$Ly_h + g_h = 0. \tag{16}$$

Now we take advantage of the linearity of the operator L and its relation to local boundary value problems. We define functions ϕ_i, ψ_{ij} according to

$$\phi_i(x) = \begin{cases} (x - x_{i-1})/h_i, & x \in \Omega_i \\ (x_{i+1} - x)/h_{i+1}, & x \in \Omega_{i+1} \\ 0, & \text{otherwise,} \end{cases}$$

$i = 0(1)N$,

$$\psi_{i0}(x) = \begin{cases} \frac{4}{3}(x^{3/2} - x_{i-1}^{3/2}\phi_{i-1}(x) - x_i^{3/2}\phi_i(x)), & x \in \Omega_i \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\psi_{i1} = \begin{cases} x_i\psi_{i0} - \frac{4}{15}(x^{5/2} - x_{i-1}^{5/2}\phi_{i-1}(x) - x_i^{5/2}\phi_i(x)), & x \in \Omega_i \\ 0 & \text{otherwise,} \end{cases}$$

$i = 1(1)N$.

From these definitions we obtain

$$-\phi_i''(x) = 0 \text{ a.e. in } \Omega, \qquad \phi_i(x_k) = \delta_{ik} \tag{17}$$

and

$$-\psi_{ij}''(x) + \zeta_{ij}(x) = 0 \text{ a.e. in } \Omega, \qquad \psi_{ij}(x_k) = 0. \tag{18}$$

Taking $y_i \in R$, $i = 0(1)N$, as parameters the superposition principle results in the representation

$$y_h(x) = \sum_{i=0}^{N} y_i\phi_i(x) + \sum_{i=1}^{N}\sum_{j=0}^{1} w_{ij}\psi_{ij}(x) \tag{19}$$

of the solutions y_h of the auxiliary problems (6). Here w_{ij}

denote the coefficients of $G_h y_h$ with respect to the base $\{\zeta_{ij}\}$ of V_h, according to (15), and $y_0 = 1$, $y_N = 0$ are given by the boundary conditions (2). We obtain the remaining coefficients y_i, $i = 1(1)\,N-1$, in (19) from the differentiability of y_h in the inner grid points. Taking the supports of ϕ_i, ψ_{ij} into account this leads to

$$\frac{u_i - u_{i-1}}{h_i} + \sum_{j=0}^{1} w_{ij}\psi'_{ij}(x_i - 0) = \frac{u_{i+1} - u_i}{h_{i+1}} + \sum_{j=0}^{1} w_{i+1j}\psi'_{i+1j}(x_i + 0), \qquad (20)$$

$i = 1(1)\,N-1$.

With (9), (10), system (20) forms a nonlinear system for determining the coefficients y_i, $i = 1(1)\,N-1$. We solve this recursively by the following modifications of Newton's method:

$$\frac{y_i^{k+1} - y_{i-1}^{k+1}}{h_i} + ([y_i^k]_+^{3/2} + \frac{3}{2}[y_i^k]_+^{1/2}(y_i^{k+1} - y_i^k))\psi'_{i0}(x_i - 0) + w_{i1}^k\psi'_{i1}(x_i - 0) = \frac{y_{i+1}^{k+1} - y_i^{k+1}}{h_{i+1}} + ([y_{i+1}^k]_+^{3/2} + \frac{3}{2}[y_{i+1}^k]_+^{1/2}(y_i^{k+1} - y_i^k)) \times \psi'_{i+1,0}(x_i + 0) + w_{i+1,1}^k\psi'_{i+1,1}(x_i + 0) \qquad (21)$$

$i = 1(1)\,N-1$ with $y_0^{k+1} = 1$, $y_N^{k+1} = 0$, and

$$w_{i1}^k = \tfrac{3}{2}[y_i^k]_+^{1/2}[y_i'^k]_- \quad \text{in the case } G_h = \bar{G}_h \qquad (22)$$

or

$$w_{i1}^k = \frac{[y_i^k]_+^{3/2} - [y_{i-1}^k]_+^{3/2}}{h_i} \quad \text{in the case } G_h = \underline{G}_h, \qquad (23)$$

respectively. Problem (21) forms a system of linear equations

$$A_k y^{k+1} = b^k \qquad (24)$$

for determining the new iterates $y^{k+1} = (y_i^{k+1})_{i=1}^{N-1} \in R^{N-1}$. Due to (21) the matrices $A_k = (a_{ij}^k)$ and the vectors $b^k = (b_i^k)$ are defined by

$$a_{ij}^k = \begin{cases} -1/h_i, & j = i-1 \\ 1/h_i + 1/h_{i+1} + \alpha_i^k, & j = i \\ -1/h_{i+1}, & j = i+1 \\ 0, & \text{otherwise,} \end{cases}$$

and

$$b_i^k = \begin{cases} \beta_1^k + 1/h_1, & i = 1, \\ \beta_i^k, & i = 2(1)\,N-1, \end{cases}$$

respectively, with

$$\alpha_i^k = \tfrac{3}{2}[y_i^k]_+^{1/2}\psi'_{i0}(x_i - 0) - \tfrac{3}{2}[y_{i+1}^k]_+^{1/2}\psi'_{i+1,0}(x_i + 0),$$

$$\beta_i^k = \alpha_i^k y_i^k - [y_i^k]_+^{3/2}\psi'_{i0}(x_i - 0) + [y_{i+1}^k]_+^{3/2}\psi'_{i+1,0}(x_i + 0) - w_{i1}^k\psi'_{i1}(x_i - 0) + w_{i+1,1}^k\psi'_{i+1,1}(x_i + 0).$$

The definitions of the functions ϕ_i, ψ_{i0} lead to

$$\psi'_{i0}(x) = 2x^{1/2} - \frac{4}{3}\frac{x_i^{3/2} - x_{i-1}^{3/2}}{h_i} \quad \text{for any} \quad x \in \Omega_i.$$

With the convexity of the function $\sigma(x) = x^{3/2}$, $x > 0$ and with the mean value theorem we obtain

$$\psi'_{i0}(x_{i-1} + 0) \leqslant 0, \qquad \psi'_{i0}(x_i - 0) \geqslant 0.$$

This results in $\alpha_i^k \geqslant 0$. Thus, the matrices A_k are symmetric tridiagonal M-matrices. Therefore the linear system (24) can be solved efficiently by fast factorizations without pivoting. Furthermore, a discrete maximum principle holds and system (24) is stable w.r.t. perturbations.

Let $\{\underline{y}_h^k\}$, $\{\bar{y}_h^k\} \subset U$ denote the sequences defined by

$$y_h^k(x) = \sum_{i=0}^{N} y_i^k\phi_i(x) + \sum_{i=1}^{N}\sum_{j=0}^{1} w_{ij}^k\psi_{ij}(x) \qquad (25)$$

with y_i^k, w_{i1}^k generated according to (21), (22), (23) and with

$$w_{i0}^k := [y_i^k]_+^{3/2}.$$

Then the following theorem (compare [8]) holds.

THEOREM *The sequences $\{\underline{y}_h^k\}$, $\{\bar{y}_h^k\}$ generated due to the iteration (21)–(25) converge locally for sufficiently small $h > 0$ to $\underline{y}_h$, $\bar{y}_h$, respectively.*

Remarks. • An iteration method for solving auxiliary problems (8) with bounding operators G_h according to (12), (13) can be realized in the same way as in the case (9), (10).

• The derivative occurring in (22) can be replaced by means of the (19). This leads to the determination of w_{i1}^k from the equation

$$w_{i1}^k = \tfrac{3}{2}[y_i^k]_+^{1/2}((y_i^k - y_{i-1}^k)/h_i + [y_i^k]_+^{3/2}\psi'_{i0}(x_i) + w_{i1}^k\psi'_{i1}(x_i)).$$

• The approach given above is also applicable to neutral atoms with a Bohr radius b which results in the boundary conditions

$$y(0)=1, \qquad by'(b)-y(b)=0$$

instead of (2). The case of an isolated neutral atom characterized by

$$y(0)=1, \qquad \lim_{x\to+\infty} y(x)=0$$

requires one to estimate the asymptotic behaviour and cannot be handled directly by the proposed method.

• The investigations given in this paper can be extended to the generalized Thomas–Fermi equation

$$y''+(b/x)\,y'=cx^p y^q,$$

where b, c, p, and q are constants such that

$$0\leqslant b<1, \qquad c>0, \qquad p>-2, \qquad q>1$$

as considered in [6] by using appropriate operators L, G. However, this leads to more complicated bases ϕ_i, ψ_{i0}, ψ_{i1} satisfying the related local boundary value problem (17), (18). The non-polynomial splines considered in [12] can be used as bases in this case.

Some of the underlying ideas such as monotone discretization by bounding operators and using convexity can be applied to other types of problems as well. In the case of generalized Emden–Fowler equations considered in [7], e.g., this approach leads to a sequence of finite-dimensional nonlinear eigenvalue problems. The related eigenfunctions are represented by piecewise cubic polynomials. It should be mentioned that unlike in [7] we concentrate our attention to the construction of an adapted discretization of the eigenvalue problem. The generated finite-dimensional systems can be treated by a technique similar to [7], e.g. Here we applied some shooting method to solve the finite-dimensional problems.

Let us consider the generalized Emden–Fowler equation (superlinear case, compare [7])

$$y''(x)+\lambda y^3(x)=0 \quad \text{in} \quad \Omega:=(0,1),$$
$$y(0)=y(1)=0$$

as an example. Related to this problem we define

$$\langle Gu, v\rangle := \int_\Omega [u(x)]_+^3\, v(x)\,dx \qquad \text{for any} \quad u, v\in H_0^1(\Omega). \tag{26}$$

Similar to (9), (10) we apply piecewise secants and piecewise tangents to define bounding operators. But, unlike in the case of the Thomas–Fermi equation here the solution u is concave and the superposition u^3 neither is concave nor is convex. Thus, we use the piecewise linearizations to the function u itself. This results in the bounding operators

$$[\bar{G}_h u](x) := \left[u_i - \frac{u_i-u_{i-1}}{h_i}(x_i-x)\right]_+^3 \tag{27}$$

and

$$[\underline{G}_h u](x) := [u_i - u_i'(x_i-x)]_+^3, \tag{28}$$

for any $x\in\Omega_i$. Now, the finite-dimensional eigenvalue problem can be given in the abstract form

$$y_h\in H_0^1(\Omega), \qquad Ly_h=\lambda_h G_h y_h. \tag{29}$$

This problem can be treated in an adapted finite-dimensional space. Because $G_h y$ for any $y\in H_0^1(\Omega)$ are piecewise cubic polynomials we know the solutions y_h of (29) to be polynomials of degree five.

4. NUMERICAL SOLUTIONS

We realized the method of monotone discretization on an IBM PC in turbo Pascal. The systems (21), (22), (23) were solved by a fast Cholesky factorization. The following Tables I and II report the results obtained for different uniform grids with the second-order method (8)–(10) and first-order method (8), (12), (13), respectively. We used the

TABLE I

x	$\underline{y}_h(x)$		$y(x)$	$\bar{y}_h(x)$	
	h=0.1	h=0.025		h=0.025	h=0.1
0.0	1.00000	1.00000	1.00000	1.00000	1.00000
0.1	0.84923	0.84946	0.84947	0.84950	0.84979
0.2	0.72692	0.72721	0.72723	0.72727	0.72766
0.3	0.61896	0.61927	0.61929	0.61933	0.61977
0.4	0.52000	0.52039	0.52041	0.52045	0.52090
0.5	0.42725	0.42753	0.42755	0.42758	0.42802
0.6	0.33843	0.33867	0.33869	0.33872	0.33912
0.7	0.25218	0.25239	0.25240	0.25242	0.25277
0.8	0.16749	0.16764	0.16765	0.16767	0.16794
0.9	0.08360	0.08368	0.08369	0.08370	0.08387
1.0	0.00000	0.00000	0.00000	0.00000	0.00000

TABLE II

x	$\underline{y}_h(x)$ h=0.01	$\underline{y}_h(x)$ h=0.0025	$y(x)$	$\bar{y}_h(x)$ h=0.0025	$\bar{y}_h(x)$ h=0.01
0.0	1.00000	1.00000	1.00000	1.00000	1.00000
0.1	0.84891	0.84933	0.84947	0.84961	0.85003
0.2	0.72641	0.72703	0.72723	0.72743	0.72804
0.3	0.61837	0.61907	0.61929	0.61952	0.62020
0.4	0.51948	0.52018	0.52041	0.52065	0.52133
0.5	0.42668	0.42733	0.42755	0.42777	0.42841
0.6	0.33792	0.33850	0.33869	0.33888	0.33944
0.7	0.25178	0.25224	0.25240	0.25255	0.25301
0.8	0.16721	0.16754	0.16765	0.16776	0.16808
0.9	0.08346	0.08363	0.08369	0.08374	0.08391
1.0	0.00000	0.00000	0.00000	0.00000	0.00000

TABLE III

h	0.02	0.01	0.005	0.0025
d_h	3.7E-5	9.4E-6	2.4E-6	6.0E-7
D_h	3.7E-3	1.9E-3	9.3E-4	4.6E-4

TABLE IV

h	λ_h due to (28)	λ_h due to (27)	difference
0.1	88.0498	97.528966	9.36E-0
0.01	94.475750	94.566103	9.04E-2
0.001	94.535760	94.536652	8.92E-4
0.0001	94.536350	94.536358	8.00E-6

TABLE V

x	$y_h(x)$	$y(x)$
0.1	0.099953	0.09961
0.2	0.198497	0.19790
0.3	0.288869	0.28879
0.4	0.356025	0.35498
0.5	0.381380	0.38025

parameter $a = 1$ in both cases, but, we applied different step sizes to obtain results of appropriate accuracies.

The solution y obtained with the proposed method coincides with the results given in [6]. In Table III we show the error of the enclosures realized with the second-order monotone discretization technique.

Here d_h and D_h denote the maximal difference between upper and lower solutions in the grid points, i.e., $\max(\bar{y}_i - \underline{y}_i)$, obtained by the methods (8), (9), (10) and (8), (12), (13), respectively. The numerical results show a good coincidence with the properties derived for the proposed method of monotone discretization.

We remark that the iteration method given in Section 4 converges very fast (compare [8]). Starting from $y_h^0(x) = 1 - x/a$, the wanted accuracy 10^{-7} was reached after three iterations.

Finally, we report some numerical results obtained with the proposed method for the Emden–Fowler equation. With the scaling $y'(0) = 1$ as used in [7] we solved these problems for various stepsizes $h > 0$ on equidistributed grids by a shooting method. Our results are given in Table IV.

The bounds converge quadratically as expected. Finally, we list the solution obtained with (27), (29) at selected x-values and we compare the results with the exact solution as given in [7] (see Table V).

REFERENCES

1. R. A. Adams, *Sobolev Spaces* (Academic Press, New York, 1975).
2. N. Anderson and A. M. Arthurs, *Q. Appl. Math.* **39**, 127 (1981).
3. B. L. Burrows and P. W. Core, *Q. Appl. Math.* **42**, 73 (1984).
4. V. Bush and S. H. Caldwell, *Phys. Rev.* **38**, 1898 (1931).
5. C. Y. Chan and S. W. Du, *Q. Appl. Math.* **44**, 303 (1986).
6. C. Y. Chan and Y. C. Hon, *Q. Appl. Math.* **45**, 591 (1987).
7. S. R. Flagg, C. D. Luning, and W. L. Perry, *J. Comput. Phys.* **38**, 396 (1980).
8. C. Grossmann, in *Numerical Solution of Differential Equations and Enclosures*, edited by E. Adams *et al.*, (Akademie Verlag, Berlin, 1987), p. 99.
9. C. Grossmann, M. Krätzschmar, and H.-G. Roos, *Numer. Math.* **49**, 95 (1986).
10. C. D. Luning, *SIAM J. Math. Anal.* **9**, 515 (1978).
11. J. M. Mooney, *Q. Appl. Math.* **36**, 305 (1978).
12. M. Sakai and R. A. Usmani, *BIT* **28**, 867 (1988).
13. J. Sprekels and H. Voss, *Numer. Math.* **32**, 381 (1979).