stichting
mathematisch
centrum

MC

AFDELING NUMERIEKE WISKUNDE NW 6/75 MAY

P.W. HEMKER
THE SOLUTION OF A NONLINEAR ELLIPTIC
OPERATOR EQUATION BY FINITE ELEMENTS

2nd edition

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 65L10
ACM - Computing Reviews - category: 5.15

Contents

Abstract

A method is described for the computation of the electron density in a conducting medium in the presence of an impurity. Finite element techniques are used for the solution of the resulting one-dimensional nonlinear operator equation.
The nonlinear operator involves a linear integral operator which causes a non-sparse system of equations to be solved. Nonlinearity is dealt with by a modified Newton-Raphson process. Proper initial approximations are obtained by means of the Davidenko principle.

0. Preface

The main purpose of this report is not only to find the solution of a boundary value problem. It is also intended to show how a constructive solution of a problem can be found by only the elaboration of some conceptually simple techniques.

Often one is inclined to transform a problem into another one which is mathematically equivalent indeed, but is different conceptually. In a number of cases such a transformation enables us to find an analytical solution of the original problem, and that solution may be of use when numerical results are wanted. However, when the problem does not admit an analytical solution, often a numerical approximation remains possible.

In many cases it will be easier and more efficient to look for a constructive method that agrees with the original problem. This implies that we disregard the transformation that in some cases will make it possible to find the analytical solution.

In our example we show an electron density problem. We will disregard the differential equations and the corresponding boundary data by which the problem can be described and we will base our constructive method directly on the minimum-energy principle.

Our constructive method consists of the use of: 1. piecewise cubic Hermite polynomials for the approximation of the solution; 2. the Rayleig-Ritz-Galerkin method for the set-up of the discrete set of equations; 3. the Lagrange multiplier method for the implementation of a supplementary condition; 4. the modified Newton-Raphson method for the solution of the resulting set of nonlinear equations, and 5. the Davidenko principle for finding proper initial approximations.

1. The physical problem

The mathematical problem we solve stems from the following physical situation. An infinitely large grid of metal atoms has a density of conducting electrons $\bar{\rho}$.
The grid is disturbed by an impurity with potential $-\phi(x)$

$$\begin{aligned} \phi(x) &= -Q/r_0 \qquad r = |\vec{x}| < r_0 \\ &= -Q/r \qquad r = |\vec{x}| \geq r_0 \quad . \end{aligned} \tag{1.1}$$

This potential causes a perturbation of the electron density $\rho(x)$. It is by no means trivial to calculate this perturbation, since the electrons influence each other both directly and by means of their potential.

The energy ε of the physical system may be written as a functional of the electron density ρ

$$\varepsilon[\rho] = \int \frac{3}{10}(3\pi^2)^{2/3}\rho^{5/3} + \frac{\mu}{8}\frac{(\nabla\rho)^2}{\rho} + U(\rho) + (V(r)+\phi(r))(\rho(r)-\bar{\rho})d^3r. \tag{1.2}$$

The potential induced by the electron density is given by $-2V(r)$;
$V(r)$ being given by

$$V(r) = \frac{1}{2}\int \frac{\rho(\xi)-\bar{\rho}}{|r-\xi|}\, d^3\xi \quad . \tag{1.3}$$

This is equivalent to

$$\begin{aligned} &\rho(r) - \bar{\rho} = -\frac{1}{2\pi}\nabla^2 V(r) \ ; \\ &\lim_{r\to\infty} rV(r) = Q/2 \ . \end{aligned} \tag{1.4}$$

$U(\rho)$ is given by

$$U(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\rho^{4/3} - 0.00517\ \rho \ln\rho \quad . \tag{1.5}$$

The parameter μ takes some value between 1/9 and 1.

We remark that this model only simplifies reality as the terms $U(\rho)$ and $\frac{\mu}{8}\frac{(\nabla\rho)^2}{\rho}$ are approximate descriptions to the physical behaviour. The solution of the given equations was requested in order to test the validity of these approximations for different values of μ.

Because of the sphere symmetry of the problem we write equation (1.2) as

$$\varepsilon[\rho] = 4\pi \int [\frac{3}{10}(3\pi^2)^{2/3}\ \rho^{5/3} + \frac{\mu}{8}\frac{(\rho')^2}{\rho} + U(\rho) + \\ + (V(r)+\phi(r))(\rho(r)-\bar{\rho})]\ r^2 dr \qquad (1.6)$$

with $\rho' = d\rho/dr$

and

$$V(r) = \frac{2\pi}{r}\int_0^r (\rho-\bar{\rho})\xi^2 d\xi + 2\pi \int_r^\infty (\rho-\bar{\rho})\xi d\xi\ . \qquad (1.7)$$

The resulting electron density $\rho(r)$ will be that function of r, which minimizes the energy ε (cf. eq. 1.6) and satisfies the supplementary condition (cf. eq. 1.4)

$$4\pi \int_0^\infty (\rho-\bar{\rho})r^2 dr = Q \qquad (1.8)$$

i.e. the local surplus of charge takes a value, such that the disturbing potential $\rho(\vec{x})$ will be compensated.

2. The numerical problem

Mathematically stated, the problem reads as follows:
find a numerical approximation to the distribution $\rho(x) : \mathbb{R}^+ \to \mathbb{R}^+$ which minimizes the functional

$$\frac{\varepsilon[\rho]}{4\pi} = \int_0^\infty [g(\rho,D\rho) + (\phi(r)+W[\rho-\bar{\rho}](r))(\rho-\bar{\rho})]\, r^2 dr \tag{2.1}$$

and which satisfies the supplementary condition

$$\frac{Q}{4\pi} = \int_0^\infty (\rho-\bar{\rho})\, r^2 dr \tag{2.2}$$

where

$$g(\rho,D\rho) = \frac{3}{10}(3\pi^2)^{2/3}\, \rho^{5/3} - \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \rho^{4/3} - 0.00517\, \rho \ln\rho + \frac{\mu}{8}\frac{(D\rho)^2}{\rho}\,, \tag{2.3}$$

$$\phi(r) = \begin{cases} -Q/r_0 & \text{if } r \le r_0 \\ -Q/r & \text{if } r \ge r_0 \end{cases} \tag{2.4}$$

$$W[\rho-\bar{\rho}](r) = \frac{2\pi}{r}\int_0^r (\rho-\bar{\rho})\xi^2 d\xi + 2\pi \int_r^\infty (\rho-\bar{\rho})\xi d\xi\,, \tag{2.5}$$

and where $\bar{\rho}, Q, \mu$ and r_0 are some given positive parameters.

3. The formulation of the discrete problem

Using the basic principle of the Rayleig-Ritz-Galerkin method we try to find the solution as a linear combination of a previous given set of functions.
Hence we write

$$\rho_h(x) = \bar{\rho} + \sum_{\substack{j=0 \\ s=1,2}}^{j=N} a_j^s \, \psi_j^s(x) \tag{3.1}$$

where $\bar{\rho}$ is a known parameter (electron density at infinity) and $\{\psi_j^s(x)\}_{j,s}$ is a set of chosen functions.
The set $\{a_j^s\}_{j,s}$ has to be determined such that $\rho_h(x)$ gives a sufficient approximation to the electron density $\rho(x)$. Where no confusion will be possible, we will omit the subscript h and write $\rho(x)$ for $\rho_h(x)$.

Using the basic idea of the finite element method, we choose a finite set of functions $\{\psi_j^s\}_{j,s}$ with limited support. Since the function $\rho(x)$ is defined on the infinite interval $[0,\infty)$, we choose a sufficiently large subinterval $[0,R]$, assuming that $\rho(x) \equiv \bar{\rho}$ on $(R,\infty]$. The interval $[0,R]$ is divided into a number of subintervals $[x_i,x_{i+1}]$, $i = 0,1,2,\ldots,N-1$.
We take $x_0 = 0$, $x_M = r_0$ and $x_N = R$.

The functions $\{\psi_j^s(x)\}_{s=0,1;\ 0\le j\le N}$ are defined on $[x_0,x_N]$ such that

1) $\psi_j^s(x) = 0$ for $x \notin (x_{j-1},x_{j+1})$,

2) $\psi_j^s(x)$ is a piecewise 3rd degree polynomial on $[x_{i-1},x_i]$ and on $[x_i,x_{i+1}]$, and

3) $\psi_j^0(x_j) = 1$; $D\psi_j^0(x_j) = 0$; *)

$\psi_j^1(x_j) = 0$; $D\psi_j^1(x_j) = 1$.

---

*) D denotes the differential operator $D = \frac{d}{dx}$.

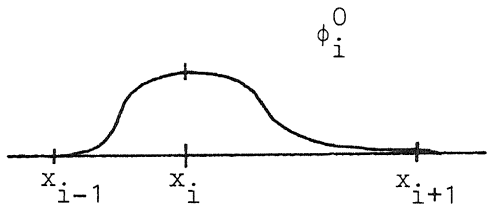
It follows that

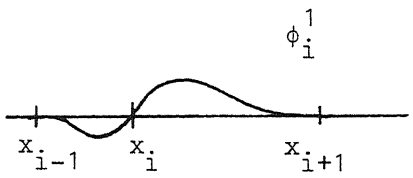
$$\psi_i^0(x) = (1-s)^2(1+2s) \qquad \text{for } x \in [x_i, x_{i+1}] \text{ with } s = (x-x_i)/(x_{i+1}-x_i)$$

$$= s^2(3-2s) \qquad \text{for } x \in [x_{i-1}, x_i] \text{ with } s = (x-x_{i-1})/(x_i-x_{i-1}) \tag{3.2}$$

$$\psi_i^1(x) = s(1-s)^2(x_{i+1}-x_i) \qquad \text{for } x \in [x_i, x_{i+1}] \text{ with } s = (x-x_i)/(x_{i+1}-x_i)$$

$$= -s^2(1-s)(x_i-x_{i-1}) \qquad \text{for } x \in [x_i, x_{i-1}] \text{ with } s = (x-x_{i-1})/(x_i-x_{i-1}) \tag{3.3}$$





The discrete problem is obtained by substituting $\rho_h(x)$ for $\rho(x)$ (cf. eq.(3.1)) in equations (2.1) and (2.2).

In order to compute the function $\rho_h(x)$ which minimizes $\varepsilon[\rho]$ we consider the variational equations

$$\frac{\partial}{\partial a_i^t} \varepsilon[\rho] = 0 , \quad \forall a_i^t . \tag{3.4}$$

In order to implement the supplementary condition (2.2), which reads in its discrete form

$$\frac{Q}{4\pi} = \int_0^\infty \sum \psi_j^s \rho_j^s r^2 dr , \tag{3.5}$$

we introduce a Lagrange multiplier λ, thus expanding the set of equations (3.4) to

$$\left.\begin{aligned} R_i^t \overset{d.}{=} \frac{\partial}{\partial a_i^t}\left[\frac{\varepsilon[\rho_h]}{4\pi} + \lambda\int_0^\infty (\rho_h-\bar{\rho})r^2dr\right] = 0 \qquad (3.6.a)\\ S \overset{d.}{=} \int_0^\infty (\rho_h-\bar{\rho})r^2dr - \frac{Q}{4\pi} = 0 \qquad (3.6.b)\end{aligned}\right\}$$

More explicitly these equations may be written as

$$R_i^t[\rho] = \int_0^\infty [D\psi_i^t\, g_{\rho'}(\rho,\rho') + \psi_i^t\, g_\rho(\rho,\rho') + \psi_i^t\{\phi(r)+\lambda\}]r^2dr +$$

$$+ \sum a_j^s \int_0^\infty \{\psi_i^t\, W[\psi_j^s](r) + \psi_j^s\, W[\psi_i^t](r)\}r^2dr = 0 \qquad (3.7)$$

and

$$S[\rho] = \sum a_j^s \int_0^\infty \psi_j^s\, r^2\, dr - Q/(4\pi) = 0 \qquad (3.8)$$

Since (3.6) is a nonlinear system - g_ρ and $g_{\rho'}$ are nonlinear functions of $\{a_j^s\}$ - we try to solve it by means of the Newton-Raphson method. Starting with sufficient accurate approximations $\tilde{\rho}$ and $\tilde{\lambda}$ to ρ and λ, respectively, and defining $\{\tilde{a}_j^s\}$ by

$$\tilde{\rho} = \bar{\rho} + \sum_{js} \tilde{a}_j^s\, \psi_j^s \quad ,$$

we compute corrections to $\{\tilde{a}_j^s\}$ and $\tilde{\lambda}$. Denoting these corrections by $\{\Delta a_j^s\}$ and $\Delta\lambda$ we have to solve the linear system

$$\begin{pmatrix} \partial R_i^t/\partial a_j^s & \partial R_i^t/\partial\lambda \\ \partial S/\partial a_j^s & \partial S/\partial\lambda \end{pmatrix} \begin{pmatrix} \Delta a_j^s \\ \Delta\lambda \end{pmatrix} + \begin{pmatrix} R_i^t \\ S \end{pmatrix} = \begin{pmatrix} \vec{0} \\ 0 \end{pmatrix}. \qquad (3.9)$$

If the initial approximations are not too bad, better approximations to ρ and λ are

$$\sum (\tilde{a}_j^s+\Delta a_j^s)\, \psi_j^s \quad \text{and} \quad \tilde{\lambda}+\Delta\lambda \ . \qquad (3.10)$$

4. The construction of the linearized system of equations

Since the support of $\psi_i^t(x)$ is the finite interval $S_i = [x_{i-1}, x_{i+1}]$ the entries of the linear system (3.9) can be computed without difficulties. The entries of the right hand side are $-R_i^t$ and $-S$, given in eqs. (3.7) and (3.8). The entries of the matrix are given by

$$\partial S/\partial\lambda = 0 \ , \tag{4.1}$$

$$\partial R_i^t/\partial\lambda = \partial S/\partial a_i^t = \int_{S_i} \psi_i^t \, r^2 dr \ , \tag{4.2}$$

$$\partial R_i^t/\partial a_j^s = \int_{S_i\cap S_j} [D\psi_i^t \, D\psi_j^s \, g_{\rho',\rho'} + (D\psi_i^t \, \psi_j^s + D\psi_j^s \, \psi_i^t) g_{\rho\rho'} + \psi_i^t \psi_j^s g_{\rho\rho}] r^2 dr +$$

$$+ \int_{S_i\cap S_j} \{\psi_i^t \, W[\psi_i^s] + \psi_j^s \, W[\psi_i^t]\} r^2 dr \ . \tag{4.3}$$

Most components of these definite integrals can easily be computed analytically or by means of numerical quadrature. However, the occurance of the operator W in R_i^t and in $\partial R_i^t/\partial a_j^s$ causes some computational inconvenience: the fact that W is defined by an integral will force us to deal with a double-integral and, as a consequence, the linear system will be non-sparse.

We will treat this in more detail.
In order to be able to compute $\partial R_i^t/\partial a_j^s$ we have to compute

$$T_{ij}^{ts} \stackrel{d}{=} \int_{S_i\cap S_j} \psi_i^t \, W[\psi_j^s](r) r^2 dr \ . \tag{4.4}$$

Using the definition of W (eq.(2.5)) we write

$$T_{ij}^{ts} = \int_0^\infty \psi_i^t(r) [\frac{2\pi}{r} \int_0^r \psi_i^s(\xi) \xi^2 d\xi + 2\pi \int_r^\infty \psi_j^s(\xi) \xi d\xi] r^2 dr \tag{4.5}$$

$$= 2\pi \iint_{0<\xi<r<\infty} \psi_i^t(r) \psi_j^s(\xi) \xi^2 r \, d\xi dr + 2\pi \iint_{0<r<\xi<\infty} \psi_i^t(r) \psi_j^s(\xi) \xi r^2 \, d\xi dr \ .$$

We note that $T_{ij}^{ts} = T_{ji}^{st}$, and, because of the finite support of ψ_i^t and ψ_j^s , it follows that

$$T_{ij}^{ts} = 2\pi \iint_{\substack{r\in S_i \\ \xi\in S_j}} \psi_i^t(r)\psi_j^s(\xi) \quad \{\underline{\text{if}}\ \xi<r\ \underline{\text{then}}\ \xi^2 r\ \underline{\text{else}}\ r^2\xi\}\ d\xi dr$$

$$= 2\pi \sum_{\substack{k=i-1,i \\ \ell=j-1,j}} K_{k\ell ij}^{ts} \ . \tag{4.6}$$

with

$$K_{k\ell ij}^{ts} = \int_{r=x_k}^{x_{k+1}} \int_{\xi=x_\ell}^{x_{\ell+1}} \psi_i^t(r)\psi_j^s(\xi)\ \{\underline{\text{if}}\ \xi<r\ \underline{\text{then}}\ \xi^2 r\ \underline{\text{else}}\ r^2\xi\}\ d\xi dr. \tag{4.7}$$

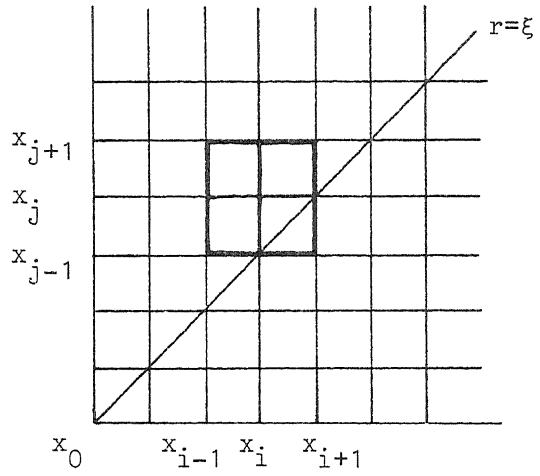


Figure 1.

The domain of integration of T_{ij}^{ts} .

Thus we split the double-integral into four parts, each of which can be compute separately.

In the case $k \neq \ell$ the calculation can be reduced to the computation of line-integrals:

$$K_{k\ell ij}^{ts} = M_{ki}^{t} \times L_{\ell j}^{s} \qquad \text{for } k < \ell$$

$$= L_{ki}^{t} \times M_{\ell j}^{s} \qquad \text{for } k > \ell \tag{4.8}$$

where

$$M_{ki}^{t} = \int_{r=x_k}^{x_{k+1}} \psi_i^t(r)\, r^2 dr \ ,$$

$$L_{ki}^{t} = \int_{r=x_k}^{x_{k+1}} \psi_i^t(r)\, r dr \ .$$

We note that M_{ki}^{t} can also be used for the computation of $\partial R_i^t/\partial\lambda$, since

$$\partial R_i^t/\partial\lambda = \partial S/\partial a_i^t = M_{(i-1)i}^{t} + M_{ii}^{t} \ . \qquad (4.9)$$

Now equations (3.7) and (4.3) may be written

$$R_i^t = \int_{S_i} [D\psi_i\, g_{\rho}, + \psi_i^t\, g_{\rho} + \psi_i^t\{\phi(r)+\lambda\}]\, r^2 dr + 2 \sum_{js} a_j^s T_{ij}^{ts} \qquad (4.10)$$

and

$$\partial R_i^t/\partial a_j^s = Q_{ij}^{ts} + 2T_{ij}^{ts} \qquad (4.11)$$

where

$$Q_{ij}^{ts} = \int_{S_i \cap S_j} [D\psi_i^t\, D\psi_j^s\, g_{\rho,\rho}, + (D\psi_i^t\, \psi_j^s + D\psi_j^s\, \psi_i^t) g_{\rho\rho}, + \psi_i^t\, \psi_j^s\, g_{\rho\rho}] r^2 dr \, .$$

Inspection of (4.11) shows that

1. $(\partial R_i^t/\partial a_j^s)$ is a symmetric positive definite matrix.
2. Q_{ij}^{ts} - with rows identified by (i,t) and columns identified by (j,s) - is a bandmatrix.
3. $(\partial R_i^t/\partial a_j^s)$ is non-sparse, since T_{ij}^{ts} is a full matrix.
4. The first term, Q_{ij}^{ts}, of the right hand side of equation (4.11) is dependent on ρ - and hence on $\{a_j^s\}$ - but T_{ij}^{ts} is independent of ρ.

Hence, we have to solve the non-sparse symmetric positive semi-definite system (3.9) at each Newton iteration step, and each time the corrected

value ρ

$$\rho = \tilde{\rho} + \sum_{js} \Delta a_j^s \psi_j^s$$

enables us to update R_i^t and Q_{ij}^{ts} .
However, in order to save computing time, the modified Newton-Raphson method is used, in which the matrix Q_{ij}^{ts} is evaluated only a small number of times.

5. The physical meaning of the Lagrange multiplier.

Of course it would be possible to solve the system (3.9) and to consider the Lagrange multiplier as an unknown quantity indeed. However, in the present problem we are able to give an alternative treatment of the supplementary condition and to compute the Lagrange multiplier in advance.

Let us first consider the simple problem to find the vector x that minimizes the vectorfunction f(x) under the supplementary condition $\ell(x) = c$. The Lagrange multiplier method leads us to solve the system

$$\begin{aligned} \text{grad } (f(x)) + \lambda \text{ grad } (\ell(x)) &= 0 \\ \ell(x) &= c \ . \end{aligned} \tag{5.1}$$

Thus λ denotes a constant ratio such that

$$\text{grad } (f(x_0)) = -\lambda \text{ grad } (\ell(x_0))$$

where x_0 is the solution of (5.1).

The same arguments applied to the functional $\varepsilon[\rho]$ (cf. eq.(2.1)) and the supplementary condition eq. (2.2) show that

$$\text{grad } (\varepsilon[\rho]) = -\lambda \text{ grad } \left(\int_0^\infty (\rho-\bar{\rho})\, r^2 dr\right) \tag{5.2}$$

with

$$\text{grad } (\varepsilon[\rho]) = \lim_{h\to 0} \frac{\varepsilon[\rho+h\times\delta\rho] - \varepsilon[\rho]}{h} \ .$$

The constant ratio $-\lambda$, which is independent of the shape of the perturbation $\delta\rho$, is easily calculated for some $\delta\rho$ with its support at infinity and $D\delta\rho \ll \delta\rho$. A simple calculation yields

$$\text{grad } (\varepsilon[\rho]) = \int_0^\infty [g_\rho\ \delta\rho + g_{\rho'}\ D\delta\rho] r^2 dr \ .$$

For this special choice of $\delta\rho$

$$\text{grad } (\varepsilon[\bar{\rho}]) = g_\rho(\bar{\rho},0) \int_0^\infty \delta\rho \; r^2 dr$$

$$= g_\rho(\bar{\rho},0) \text{ grad } \left(\int_0^\infty (\rho-\bar{\rho}) \; r^2 dr\right) .$$

Which yields

$$\lambda = -g_\rho(\bar{\rho},0) . \tag{5.3}$$

This consideration enables us to compute the Lagrange multiplier in advance and to solve the linear system (3.9) with disregard of the last row and the last column. The remaining matrix directly stems from our minimizing problem and, consequently, is symmetric and positive definite. In order to solve the system Choleski's method is used.

Actually we replaced a global condition (eq.(2.9))

$$\frac{Q}{4\pi} = \int_0^\infty (\rho-\bar{\rho}) \; r^2 dr \tag{2.2}$$

by an equivalent local one

$$\text{grad } (\varepsilon[\rho]) = g_\rho(\bar{\rho},0) \quad \text{grad } \left(\int_0^\infty (\rho-\bar{\rho}) \; r^2 dr\right) . \tag{5.4}$$

We may compute $\int_0^R (\rho-\bar{\rho}) \; r^2 dr$ and check relation (2.2), which yields some information about the relevancy of truncating the infinite interval $[0,\infty)$ to the interval $[0,R]$.

6. The solution

The boundary conditions

The treatment of the boundary conditions is simple.
It can be done in two ways.

1. Since the boundary conditions at r = 0 and r = ∞ are natural, the non-linear system (3.6) can be solved, just as it stands. In this case all parameters $\{a_j^s\}$ j = 0,1,...,N ; s = 0,1 are computed.
   It is expected that $a_0^1 \approx 0$, $a_N^0 \approx 0$ and $a_N^1 \approx 0$, as they are the computed values of $\rho'(0)$, $\rho(x_N)-\bar{\rho}$ and $\rho'(x_N)$, respectively. The discrepancy can be used as an indication of the accuracy obtained.
2. It is also possible to keep one or more of the values a_0^1, a_N^0, a_N^1 fixed at zero.
   In this case we disregard the equations (3.6.a) for (i,t) = (0,1),(N,0) and (N,1).
   In most computations we kept only a_0^1 fixed at zero in order to preserve continuity of the first derivative at r = 0.

The convergence of the Newton-Raphson iteration

The value of the parameter $\bar{\rho}$ has a large influence on the convergence of the Newton-Raphson iteration process. This is caused by the nonlinearity of the function $g(\rho,\rho')$. In figure 2 the behaviour of $g(\rho,0)$ is sketched.

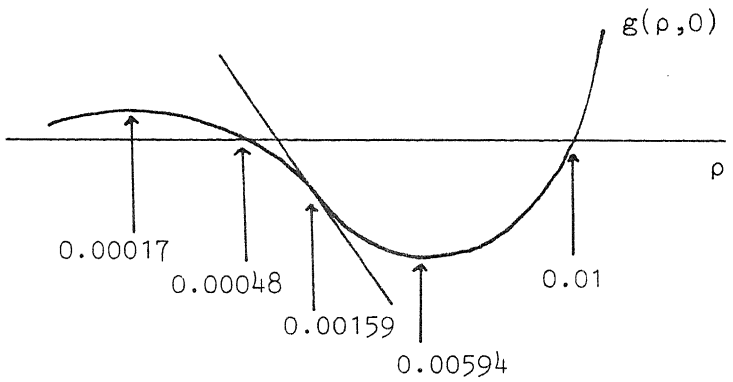


Figure 2.
The qualitative behaviour of $g(\rho,0)$.

Since $g_\rho(\rho,0) > 0$ and $g_{\rho\rho}(\rho,0) > 0$ for $\rho > 0.06$ the functional $\varepsilon[\rho]$ is easily minimized for these values of ρ and, started with the initial approximation $\rho = \bar{\rho}$, good convergence is obtained for $\bar{\rho} > 0.06$. However, values down to $\rho = 0.001$ are of physical interest and in this range convergence is very poor. We expect bad convergence since $g_{\rho\rho}(\rho,0) < 0$ for $\rho < 1.592_{10}-3$. Actually, for $\bar{\rho} = 0.001$ the process diverges when started with $\rho = \bar{\rho}$. This difficulty is dealt with by the Davidenko principle as follows.
The solution $\rho(r)$ is continuously dependent on the parameter $\bar{\rho}$. Hence, for reasonable small values of Δc, the solution with $\bar{\rho} = c + \Delta c$ will be a good approximation to the solution with $\bar{\rho} = c$. An approximate solution for $\bar{\rho} = 0.01$ is obtained in one Newton iteration step. This approximation is used as an initial approximation in another Newton iteration step that computes the solution for $\bar{\rho} = 0.01 - \Delta c$ (e.g. $\bar{\rho} = 0.008$). In this way $\bar{\rho}$ is decreased each time when an iteration step has been executed.
Thus $\bar{\rho}$ is changed during the iteration process and a good approximation is obtained when the process goes to an end with the required $\bar{\rho}$. In this way convergence is obtained. E.g. with $\mu = 0.5$ and $Q = r_0 = 1$ the sequence $\bar{\rho} = 0.01$, 0.008, 0.006, 0.004, 0.002 is sufficient to obtain convergence for $\bar{\rho} = 0.001$.

The number of iteration steps which is necessary also depends on the number of nodes (N) used in the discretization of the continuous problem.
The larger the number of nodes, the larger the number of iteration steps to obtain the same accuracy. Hence, we increased the number of nodes during the iteration process. A large number of numerical experiments have been done and, in all computations, 15 iteration steps were sufficient in order to obtain the solution.

The convergence with respect to the discretization

A comprehensive literature exists on the convergence of piece-wise polynomial approximations to the solution of continuous problems. Important papers on this subject are published by Bramble and Hilbert [1970, 1971], Ciarlet and Raviart [1972], Ciarlet, Schultz and Varge [1967, 1969] and

Strang [1972].

The most striking feature of the 3rd degree Hermite polynomial which we used is the order of accuracy [cf. Strang, 1972]. The discretization error in the computation of ρ is $O(h^4)$ and the error of ρ' is $O(h^3)$, where h denotes the mesh width. Moreover, this property not only holds for a uniform mesh, but also for a non-uniform one. This means that, given a number of nodes, the discretization error is decreased by a factor of 16 when an equal number of nodes is added. The only condition is that no extreme mesh-ratios occur.

These features enable us (1) to place the nodes in an efficient way, i.e. dense in those regions where the solution has large higher derivatives, and (2) to estimate the discrepancy between the continuous and the discrete problem by analyzing the difference between the discrete problems.

In order to give an impression of the accuracy obtained and of the amount of computational work involved, we report a numerical experiment with $\mu = 0.5$, $\bar{\rho} = 0.006$, $Q = r_0 = 1$.

number of nodes	max. rel. err. in ρ	number of iterations
14	$1_{10}-4$	8N
4	$4_{10}-2$	1N + 1MN
9	$3_{10}-3$	+ 1N + 1MN
28	$2_{10}-6$	+ 3N + 3MN

N : Newton iteration step (with evaluation of a matrix $\partial R_i^t/\partial a_j^s$)

MN: Modified Newton iteration step (without evaluation of a matrix $\partial R_i^t/\partial a_j^s$)

During the second iteration process the number of nodes was increased from 4 to 9 to 28.

References

J.H. Bramble and S.R. Hilbert (1970).
Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation.
SIAM J. Numer. Anal. 17 (1970) 112.

-- (1971).
Bounds for a class of linear functionals with application to Hermite interpolation.
Numer. Math. 16 (1971) 362.

P.G. Ciarlet and P.A. Raviart (1972).
General Lagrange and Hermite interpolation in R^n with applications to finite element methods.
Arch. Rational Mech. Anal. 46 (1972) 177.

G. Strang (1972).
Approximation in the finite element method.
Numer. Math. 19 (1972) 81.

P.G. Ciarlet, M.A. Schultz and R.S. Varga (1967-1969).
Numerical methods of high order accuracy for nonlinear boundary value problems.
I. One dimensional problem.
Numer. Math. 9 (1967) 394.
V. Monotone operator theory.
Numer. Math. 13 (1969) 51.