Miroslav Pospíšek

Nonlinear boundary value problems with application to semiconductor device equations

Applications of Mathematics, Vol. 39 (1994), No. 4, 241–258

Persistent URL: http://dml.cz/dmlcz/134255

NONLINEAR BOUNDARY VALUE PROBLEMS WITH APPLICATION TO SEMICONDUCTOR DEVICE EQUATIONS

MIROSLAV POSPÍŠEK, Praha

(Received February 14, 1992)

Summary. The paper deals with boundary value problems for systems of nonlinear elliptic equations in a relatively general form. Theorems based on monotone operator theory and concerning the existence of weak solutions of such a system, as well as the convergence of discretized problem solutions are presented.
As an example, the approach is applied to the stationary Van Roosbroeck's system, arising in semiconductor device modelling.
A convergent algorithm suitable for solving sets of algebraic equations generated by the discretization procedure proposed will be described in a forthcoming paper.

Keywords: boundary value problems for systems of nonlinear elliptic equations, semiconductor device equations

AMS classification: 65N30, 35J65, 65P05

1. INTRODUCTION

Very fast progress in many current technologies has also brought forth an increasing interest in mathematical modelling of the undergoing physical processes, which are often described by (usually nonlinear) systems of partial differential equations. In this paper, an approach to the analysis of a boundary value problem for a system of nonlinear, elliptic type equations in rather general form is described.

In Section 3, conditions on the problem data that are sufficient to define weak solutions of the problem and to prove their existence are given. Similar conditions can also be found in many other publications, see e.g. Fučík, Kufner [4], Nečas [14] and Franců [3], but the form of the so-called coercivity conditions presented here seems to be new.

Then, a discretization scheme based on the numerical integration of the lower order terms only is proposed. Existence and convergence results for this procedure are proved. Moreover, Hackbusch's results [8] are used to show that the proposed discretization scheme has some properties of the box integration method (Varga [20]).

In Section 4, the theory is applied to the well-known Van Roosbroeck's system of three coupled nonlinear partial differential equations describing the function of a semiconductor device in stationary state.

Most theorems on the existence of (weak) solutions of this problem use directly the Schauder fixed point theorem (Mock [13], Gajewski [5], Gröger [6]), or the theory of variational inequalities (Jerome [9]).

The results of Section 3 are based on monotone operator theory. To be able to apply them to the semiconductor device problem, a modified problem, solutions of which are also solutions of the original one, is formulated. Then, using theorems of Section 3, the existence of the modified problem solutions, and also existence/convergence results for the discretized problem solutions, are proved. This procedure is similar to that of Gröger [6], the difference being in the technique used to prove existence of the modified problem solutions. In [6], Schauder fixed point theorem is used directly, leading to a nonlinear block Gauss-Seidel type algorithm, but without a proof of its convergence. We shall apply Theorem 3.2, use its assertions also in the analysis of discretization procedure and in the related paper [16] prove the convergence of a solution algorithm, based on fully coupled Newton's method.

The current continuity equations (4.2) and (4.3) are often discretized by the box integration method in conjuction with the so called Scharfetter-Gummel approximation of current densities, see [18]. However, other techniques also have been developed recently, see Markowich, Zlámal [11], Miller [12], Bürgler et al. [1], Shigyo, Wada, Yasuda [19], Chen [2] and others. Good numerical properties of the discretization scheme proposed in this paper should be guaranteed by the fact that it is actually a box integration method applied to some closely related differential equation.

As far as the author knows, no similar approach to the semiconductor device equations resulting in theoretically convergent multigrid based algorithm (see [16]) has been published yet. Moreover, some results from the more general part of the paper also seem to be new—the form of the coercivity condition (3.10) and Theorem 3.3 on the convergence of discretized problem solutions.

2. Basic notation

We shall use the following notation:

$\mathbb{N}$	the set of non-negative integers,
$\mathbb{R}$	the set of real numbers,
$\dot{\forall}$	almost everywhere,
$\vec{n} = (n_1, \ldots, n_N)$	vector of outward normal,
$\rightarrow$	strong convergence,
$\rightharpoonup$	weak convergence.

Let X be a real reflexive separable Banach space, equipped with a norm $\|\cdot\|_X$. The dual space of X will be denoted by X^* and the value of a continuous linear functional $F \in X^*$ on an element $v \in X$ will be denoted as

$$\langle F, v \rangle_X .$$

Let $N \geqslant 1$, $m \geqslant 1$ be integers, $\kappa = m(N+1)$, and let $\Omega \subset \mathbb{R}^N$ be a bounded domain with a Lipschitz boundary divided into two disjoint measurable subsets Γ_D and Γ_N. Suppose that $\mu_{N-1}(\Gamma_D)$—the $(N-1)$-dimensional Lebesgue measure of Γ_D—is nonzero.

For a given vector function $u = (u_1, \ldots, u_m)$ with sufficiently smooth components $u_i \colon \Omega \rightarrow \mathbb{R}$, $1 \leqslant i \leqslant m$, we write

$$\nabla u = \left(\frac{\partial u_1}{\partial x_1}, \ldots, \frac{\partial u_m}{\partial x_1}, \frac{\partial u_1}{\partial x_2}, \ldots, \frac{\partial u_m}{\partial x_N} \right)$$

and

$$D^j u_i = \begin{cases} u_i, & j = 0, \\ \partial u_i / \partial x_j, & 1 \leqslant j \leqslant N. \end{cases}$$

For $\xi \in \mathbb{R}^\kappa$, we denote its components in the following way:

$$\xi = (\xi_{10}, \ldots, \xi_{m0}, \xi_{11}, \ldots, \xi_{mN}),$$

so that they correspond to the components of $(u, \nabla u)$.

We also introduce here an abstract function space V, which will be referred to throughout the paper:

Let $1 < p < \infty$. The closure of the set

$$\{v \in C^\infty(\bar{\Omega}) \colon v = 0 \text{ on } \Gamma_D\}$$

in the norm of $W_0^{1,p}(\Omega)$ [1] will be denoted as V^p. The space V is defined by

$$V = \prod_{i=1}^{m} V^{p_i}, \quad 1 < p_i < \infty,\ 1 \leqslant i \leqslant m, \tag{2.1}$$

and equipped with the norm

$$\|v\|_V = \Big(\sum_{i=1}^{m} \|v_i\|_{V^{p_i}}^{p_{\min}}\Big)^{\frac{1}{p_{\min}}} = \Bigg(\sum_{i=1}^{m}\Big(\sum_{j=1}^{N}\int_\Omega |D^j v_i|^{p_i}\,\mathrm{d}x\Big)^{\frac{p_{\min}}{p_i}}\Bigg)^{\frac{1}{p_{\min}}}, \tag{2.2}$$

where $p_{\min} = \min\{p_1, \dots, p_m\}$.

3. General Theory

3.1. Weak formulation and existence theorem. Let us first recall some important definitions and an abstract theorem, which plays principal role in our analysis.

Definition 3.1. Let X be a real reflexive and separable Banach space. A mapping $A\colon X \to X^*$ is said to be
demicontinuous, if

$$(\forall u_0 \in X)(\forall \{u_n\}_{n\geqslant 1} : u_n \in X)(u_n \to u_0 \text{ in } X) \Rightarrow (Au_n \rightharpoonup Au_0 \text{ in } X^*), \tag{3.1}$$

coercive, if

$$\lim_{\|v\|_X \to \infty} \frac{\langle Av, v\rangle_X}{\|v\|_X} = \infty, \tag{3.2}$$

strictly monotone, if

$$(\forall v, w \in X, v \neq w)(\langle Av - Aw, v - w\rangle_X > 0), \tag{3.3}$$

[1] Recall that this norm can be defined as follows:

$$(\forall u \in W_0^{1,p}(\Omega))\Big(\|u\|_{W_0^{1,p}(\Omega)} = \Big(\sum_{j=1}^{N}\int_\Omega |D^j u|^p\,\mathrm{d}x\Big)^{1/p}\Big).$$

satisfying the so-called condition $(M)_0$, if

$$(v_n \rightharpoonup v) \wedge (Av_n \stackrel{*}{\rightharpoonup} \varphi) \wedge (\langle Av_n, v_n\rangle_X \to \langle \varphi, v\rangle_X) \Rightarrow (Av = \varphi \text{ in } X^*). \tag{3.4}$$

Definition 3.2. Let X be a real reflexive and separable Banach space, $A\colon X \to X^*$, $F \in X^*$, X_l a finite-dimensional subspace of X, and let the following problem be given:

$$\begin{aligned} &\text{Find } u \in X \text{ such that}\\ &(\forall v \in X)(\langle Au, v\rangle_X = \langle F, v\rangle_X). \end{aligned} \tag{3.5}$$

The problem

$$\begin{aligned} &\text{Find } u_l \in X_l \text{ such that}\\ &(\forall v \in X_l)(\langle Au_l, v\rangle_X = \langle F, v\rangle_X) \end{aligned} \tag{3.6}$$

is called *the Galerkin approximation of the problem* (3.5) *on the subspace* X_l.

Theorem 3.1. *Let X be a real reflexive and separable Banach space and let a mapping $A\colon X \to X^*$ which is demicontinuous, bounded, coercive and satifying the condition $(M)_0$ be given. Then the problem* (3.5) *has a solution for all $F \in X^*$. The problems* (3.6), *$l \in N$ are also solvable and if*

$$\overline{\bigcup_{l=1}^{\infty} X_l} = X,$$

then $u_l \rightharpoonup u$. Moreover, if A is strictly monotone, then the solution of (3.5) *is unique.*

Proof. See e.g. Franců [3, Th.5.2. and Th.7.2] or Pospíšek [15, Th.4.7]. □

Let N, m and Ω be as in Section 2 and let functions

$$\begin{aligned} &a_{ij}\colon \Omega \times \mathbb{R}^{\kappa} \to \mathbb{R}, \ 1 \leqslant i \leqslant m, \ 0 \leqslant j \leqslant N,\\ &f_i : \Omega \to \mathbb{R}, \qquad 1 \leqslant i \leqslant m,\\ &d_i : \Omega \cup \Gamma_D \to \mathbb{R}, \ 1 \leqslant i \leqslant m,\\ &h_i : \Gamma_N \to \mathbb{R}, \qquad 1 \leqslant i \leqslant m \end{aligned}$$

be given. (Recall that $\kappa = m(N+1)$.) We are interested in boundary value problems in the following form:

$$\begin{aligned} -\sum_{j=1}^{N} D^j a_{ij}(x; u, \nabla u) + a_{i0}(x; u, \nabla u) &= f_i, \quad i = 1, \ldots, m, \ x \in \Omega,\\ u_i &= d_i, \quad i = 1, \ldots, m, \ x \in \Gamma_D,\\ \sum_{j=1}^{N} n_j a_{ij}(x; u, \nabla u) &= h_i, \quad i = 1, \ldots, m, \ x \in \Gamma_N. \end{aligned} \tag{3.7}$$

Let us introduce some useful properties of the functions a_{ij}:

Definition 3.3. Let functions $a_{ij}\colon \Omega \times \mathbf{R}^{\kappa} \to \mathbf{R}$, $1 \leqslant i \leqslant m$, $0 \leqslant j \leqslant N$, be given. We say that they satisfy *Carathéodory conditions*, if

$$(3.8) \qquad (\forall \xi \in \mathbf{R}^{\kappa})(a_{ij}(\cdot;\xi) \text{ is measurable in } \Omega), \\ (\dot\forall x \in \Omega)(a_{ij}(x;\cdot) \text{ is continuous in } \mathbf{R}^{\kappa}),$$

growth conditions with the coefficients $p_1, \ldots, p_m$, if

$$(3.9) \quad (\exists p_k \geqslant 1,\ 1 \leqslant k \leqslant m)(\forall i,\ 1 \leqslant i \leqslant m)(\forall j,\ 0 \leqslant j \leqslant N)(\dot\forall x \in \Omega)(\forall \xi \in \mathbf{R}^{\kappa}) \\ |a_{ij}(x;\xi)| \leqslant \sum_{k=1}^{m} \Big(g_{ij}(x) + c_{ij} \sum_{l=0}^{N} |\xi_{kl}|^{\frac{p_k}{p_i}(p_i-1)} \Big),$$

where c_{ij} are non-negative constants and $g_{ij} \in L_{q_i}(\Omega)$, $1/p_i + 1/q_i = 1$,
coercivity condition with the coefficients $p_1, \ldots, p_m$, if

$$(3.10) \qquad (\exists p_k \geqslant 1, 1 \leqslant k \leqslant m)(\exists C_c > 0)(\dot\forall x \in \Omega)(\forall \xi \in \mathbf{R}^{\kappa}) \\ \sum_{i=1}^{m}\sum_{j=0}^{N} a_{ij}(x;\xi)\xi_{ij} \geqslant C_c \sum_{i=1}^{m}\sum_{j=1}^{N} |\xi_{ij}|^{p_i} + \sum_{i=1}^{m} \theta_i(x)\xi_{i0},$$

where $C_c > 0$ is a constant and $\theta_i \in L_\infty(\Omega)$, $1 \leqslant i \leqslant m$,
strict monotonicity condition, if

$$(3.11) \qquad (\dot\forall x \in \Omega)(\forall \xi, \eta \in \mathbf{R}^{\kappa}, \xi \neq \eta) \\ \Big(\sum_{i=1}^{m}\sum_{j=0}^{N} [a_{ij}(x;\xi) - a_{ij}(x;\eta)](\xi_{ij} - \eta_{ij}) > 0 \Big),$$

condition of strict monotonicity in principal part, if

$$(3.12) \qquad (\dot\forall x \in \Omega)(\forall \xi \in \mathbf{R}^{m})(\forall \eta, \nu \in \mathbf{R}^{mN}, \eta \neq \nu) \\ \Big(\sum_{i=1}^{m}\sum_{j=1}^{N} [a_{ij}(x;\xi,\eta) - a_{ij}(x;\xi,\nu)](\eta_{ij} - \nu_{ij}) > 0 \Big).$$

The next theorem summarizes first results. It shows conditions on boundary value problem data that are sufficient for defining weak solutions of the problem and for proving their existence.

Theorem 3.2. *Let N, m and Ω be given. Consider the boundary value problem*

$$(3.13)\qquad -\sum_{j=1}^{N} D^j a_{ij}(x; u, \nabla u) + a_{i0}(x; u, \nabla u) = f_i, \quad i = 1, \ldots, m, \ x \in \Omega,$$

$$u_i = d_i, \quad i = 1, \ldots, m, \ x \in \Gamma_D,$$

$$\sum_{j=1}^{N} n_j a_{ij}(x; u, \nabla u) = h_i, \quad i = 1, \ldots, m, \ x \in \Gamma_N,$$

where the functions $a_{ij}\colon \Omega \times \mathbb{R}^{\kappa} \to \mathbb{R}$, $1 \leqslant i \leqslant m$, $0 \leqslant j \leqslant N$, satisfy

(A1) *Carathéodory conditions,*

(A2) *growth conditions with some coefficients $p_i > 1$, $1 \leqslant i \leqslant m$,*

(A3) *coercivity condition with the same coefficients as in* (A2),

(A4) *condition of strict monotonicity in principal part*

and

(*D*) $d_i \in W^{1,p_i}(\Omega)$, $f_i \in L_{q_i}(\Omega)$, $g_i \in L_{q_i}(\Gamma_N)$, $1/p_i + 1/q_i = 1$, $1 \leqslant i \leqslant m$.

Let the space V be defined as in (2.1)–(2.2), *with p_i from* (A2). *Then:*

The expression

$$(3.14)\qquad (\forall u, v \in V)\Big(\langle Au, v\rangle_V = \sum_{i=1}^{m}\sum_{j=0}^{N}\int_\Omega a_{ij}(x; u + d, \nabla(u + d)) D^j v_i \,\mathrm{d}x\Big)$$

defines a mapping $A\colon V \to V^*$ which is bounded, continuous, coercive and satisfies the condition $(M)_0$.

A functional $F \in V^*$ can be defined by

$$(3.15)\qquad (\forall v \in V)\bigg(\langle F, v\rangle_V = \sum_{i=1}^{m}\Big(\int_\Omega f_i v_i \,\mathrm{d}x + \int_{\Gamma_N} h_i v_i \,\mathrm{d}S\Big)\bigg)$$

and thus the weak formulation of the problem (3.13) *can be obtained (with A from* (3.14) *and F from* (3.15)*):*

$$(3.16)\qquad \text{Find } u \in V \text{ such that } (\forall v \in V)\big(\langle Au, v\rangle_V = \langle F, v\rangle_V\big).$$

The problem (3.16) *has a solution.*

If the strict monotonicity condition is fulfilled, then the solution of (3.16) *is unique.*

Any solution of the problem (3.16) can be expressed as a weak limit of solutions of its Galerkin approximations on the subspaces V_l, supposing that

$$\overline{\bigcup_{l=1}^{\infty} V_l} = V$$

is valid.

Proof. For the case $m = 1$, i.e. single partial differential equation, the proof can be found e.g. in Fučík, Kufner [4], Nečas [14] or Franců [3]. For the case $m \geqslant 1$, only few corrections in the proof are to be made, as shown in Pospíšek [15]. Here, only these differences summarized in Lemma 3.1 will be proved. □

Lemma 3.1. *Let, as in Theorem 3.2, the boundary value problem* (3.13) *be given. The assumptions* (A1), (A2) *and* (D) *are sufficient for the mapping A from* (3.14) *to be bounded. Moreover, if* (A3) *is also valid, then the mapping A is coercive.*

Proof of Lemma 3.1. *Boundedness.* Applying the theorem on Nemyckij operators (see e.g. Franců [3, Th.8.9]) to the functions a_{ij}, we see that the mappings

$$(u(\cdot), \nabla u(\cdot)) \mapsto a_{ij}(\cdot; u(\cdot), \nabla u(\cdot)) \tag{3.18}$$

are bounded and continuous mappings from $\prod_{i=1}^{m} \prod_{j=0}^{N} L_{p_i}(\Omega)$ to $L_{q_i}(\Omega)$, where p_i and q_i are given by (3.9). Let $v, w \in V$. Using the Hölder inequality, we estimate

$$\begin{aligned} |\langle Av, w\rangle_V| &= \left| \sum_{i=1}^{m} \sum_{j=0}^{N} \int_\Omega a_{ij}(x; v+d, \nabla(v+d)) D^j w_i \,\mathrm{d}x \right| \\ &\leqslant \sum_{i=1}^{m} \sum_{j=0}^{N} \|a_{ij}(\cdot; v+d, \nabla(v+d))\|_{q_i} \|D^j w_i\|_{p_i}, \end{aligned} \tag{3.19}$$

where $\|\cdot\|_p$, $p \in \mathbb{N}$, denotes the norm in the space $L_p(\Omega)$. The term $\|D^j w_i\|_{p_i}$ can be estimated as follows. By the Friedrichs inequality we have

$$(\forall i,\ 1 \leqslant i \leqslant m)(\exists c_i > 0)(\forall v_i \in V^{p_i})(\|v_i\|_{p_i} \leqslant c_i \|v_i\|_{V^{p_i}}). \tag{3.20}$$

Denoting $C = \max\{1, c_1, \ldots, c_m\}$, we obtain

$$(\forall i, 1 \leqslant i \leqslant m)(\forall j,\ 0 \leqslant j \leqslant N)(\|D^j w_i\|_{p_i} \leqslant C \|w_i\|_{V^{p_i}}).$$

Now, using the Hölder inequality, we see that

$$\|D^j w_i\|_{p_i} \leqslant C \sum_{i=1}^{m} \|w_i\|_{V^{p_i}} \leqslant C m^{p_{\min}/(p_{\min}-1)} \|w\|_V. \tag{3.21}$$

From (3.19) and (3.21) we obtain

$$|\langle Av, w\rangle_V| \leqslant m(N+1)Cm^{p_{\min}/(p_{\min}-1)} \sum_{j=1}^{m}\sum_{j=0}^{N} \|a_{ij}(\cdot; v+d, \nabla(v+d))\|_{q_i}\|w\|_V.$$

The boundedness of A now follows from the same property of the mappings in (3.18) as above and from the fact that $d \in \prod_{i=1}^{m} W^{1,p_i}(\Omega)$.

Coercivity. Integration of (3.10) yields

$$(\forall v \in V)\Big(\langle Av, v\rangle_V \geqslant \sum_{i=1}^{m}\Big(C_c\|v_i\|_{V_{p_i}}^{p_i} + \int_\Omega \theta_i v_i \,\mathrm{d}x\Big)\Big). \tag{3.22}$$

We shall estimate the two terms on the right-hand side of this inequality. First, note that

$$(\forall i,\ 1 \leqslant i \leqslant m)(\forall v_i \in V^{p_i})(\|v_i\|_{V_{p_i}}^{p_i} > \|v_i\|_{V_{p_i}}^{p_{\min}} - 1)$$

holds. Thus, we have the estimate

$$\sum_{i=1}^{m}\|v_i\|_{V_{p_i}}^{p_i} > \sum_{i=1}^{m}\|v_i\|_{V_{p_i}}^{p_{\min}} - m = \|v\|_V^{p_{\min}} - m. \tag{3.23}$$

Consider now the term $\int_\Omega \theta_i v_i \,\mathrm{d}x$. We shall show that it is bounded for $\|v\|_V \to \infty$. Denoting $\max_{1\leqslant i\leqslant m}\{\|\theta_i\|_\infty\}$ as θ_∞, we have

$$-\theta_\infty \int_\Omega |v_i|\,\mathrm{d}x \leqslant \int_\Omega \theta_i v_i\,\mathrm{d}x \leqslant \theta_\infty \int_\Omega |v_i|\,\mathrm{d}x,\ 1 \leqslant i \leqslant m, \tag{3.24}$$

hence we are interested in the integral $\int_\Omega |v_i|\,\mathrm{d}x$. By the Hölder inequality,

$$\int_\Omega |v_i|\,\mathrm{d}x \leqslant \mu_2(\Omega)^{\frac{1}{q_i}}\|v_i\|_{p_i}.$$

Combining this inequality with the Friedrichs inequality (3.20) and using the same technique as in (3.21), we obtain

$$\int_\Omega |v_i|\,\mathrm{d}x \leqslant c\|v\|_V, \tag{3.25}$$

where $c = \max\{\mu_2(\Omega)^{1/q_i}\}_{1\leqslant i\leqslant m} Cm^{p_{\min}/(p_{\min}-1)}$. Now, we can see from (3.22)–(3.25) that

$$(\forall v \in V)\Big(\frac{\langle Av, v\rangle_V}{\|v\|_V} \geqslant \|v\|_V^{p_{\min}-1} - \frac{m}{\|v\|_V} - h_\infty c \to \infty\Big)$$

and hence A is coercive. □

3.2. Discretization. Now, we could immediately start to look for an algorithm solving the Galerkin approximations (3.6) of the problem (3.16) on some finite-dimensional space V_l. In practice, however, problems defined by introducing some kind of numerical integration into (3.6) are solved. We shall consider these modified problems. Starting from this point, we restrict ourselves to the case $N = 2$ and in addition we shall suppose that Ω is a polygon.

First, let us construct a sequence $\{V_l\}_{l\geqslant 0}$ of finite-dimensional subspaces of the space V, such that (3.17) is valid:

1. Let T_0 be a conforming triangulation of Ω, i.e. a set of triangles such that their union is $\overline{\Omega}$ and the intersection of any two distinct triangles is either a common edge or a common vertex or is empty. Assume that the points $\overline{\Gamma_D} \cap \overline{\Gamma_N}$ are vertices of some triangles from T_0. Then triangulations T_l, $l > 0$, are constructed by induction: For each $t \in T_{l-1}$ we generate four triangles in T_l by pairwise connecting the midpoints of the edges.
2. For $l \geqslant 0$, $p > 1$ we define V_l^p as the set of continuous and piecewise linear functions

$$\{v \in V^p\colon (\forall t \in T_l)(v \text{ is linear on } t)\}, \tag{3.26}$$

equipped with the norm of the space V^p. Then, given p_i, $1 \leqslant i \leqslant m$, we set

$$V_l = \prod_{i=1}^{m} V_l^{p_i}.$$

(Note that the condition (3.17) is fulfilled.)

We introduce the following abbreviations:

$$\begin{aligned} \Omega_l^* &= \{P \in \overline{\Omega}\colon P \text{ is a vertex of some } t \in T\}, \\ \Omega_l &= \{P \in \Omega \dot{-} \overline{\Gamma_D}\colon P \text{ is a vertex of some } t \in T\}, \\ N_l &= \operatorname{card} \Omega_l. \end{aligned} \tag{3.27}$$

We now construct a dual mesh B_l for T_l: for each triangle $t \in T_l$, connect its centre of gravity by straight line segments to the edge midpoints of t. This subdivides t into three subregions with the same area. With each vertex $P \in \Omega_l^*$ we will associate a region ω_P consisting of those triangles $t \in T_l$ which have P as a vertex, and the so-called box $b_P \in B_l$, $b_P \subset \omega_P$ which consists of the union of the subregions in ω_P which again have P as a vertex.

Theorem 3.3. *Consider the problem* (3.13) *and suppose that, in addition to the assumptions of Theorem* 3.2, *the following is valid:*

$N = 2$ *and* Ω *is a polygon,*

$a_{i0} \in C(\overline{\Omega} \times \mathbb{R}^m)$, $1 \leqslant i \leqslant m$, and they do not depend on ∇u,
$f_i \in C(\overline{\Omega})$, $1 \leqslant i \leqslant m$,
$d_i \in C^1(\overline{\Omega})$, $1 \leqslant i \leqslant m$,
$h_i \in C(\Gamma_N)$, $1 \leqslant i \leqslant m$.
Then, in addition to the assertions of Theorem 3.2:

1. *The following problem on the space V_l, $l \geqslant 0$, is well-defined:*

$$\text{(3.28)} \qquad \text{Find } u^l \in V_l \text{ such that } (\forall v \in V_l)(\langle A_l u^l, v\rangle_V = \langle F, v\rangle_V),$$

where the mapping $A_l\colon V_l \to V_l^*$ and the functional $F_l \in V_l^*$ are defined by $(u, v \in V_l)$

$$\text{(3.29)} \quad \langle A_l u, v\rangle_V = \sum_{i=1}^m \sum_{j=1}^N \int_\Omega a_{ij}(x; u+d, \nabla(u+d)) D^j v_i \,\mathrm{d}x + \sum_{i=1}^m \sum_{P \in \Omega_l} \mu_2(b_P) a_{i0}(P; (u+d)(P)) v_i(P),$$

$$\langle F_l, v\rangle_V = \sum_{i=1}^m \sum_{P \in \Omega_l} (\mu_2(b_P) f_i(P) v_i(P) + \mu_1(b_P \cap \Gamma_N) h_i(P) v_i(P)) .$$

2. *For every $l \geqslant 0$, the problem (3.28) has a solution.*
3. *The sequence $\{u^l\}_{l\geqslant 0}$ defined by (3.28) weakly converges to the solution of the problem (3.14)–(3.16).*

Proof. 1. The assertion follows from the properties (A1), (A2) and (D), the continuity of the functions a_{i0}, $1 \leqslant i \leqslant m$ and from Theorem 3.2.

2. This also follows from Theorem 3.2. Note that the problems are formulated in finite-dimensional spaces, hence only continuity and coercivity of mappings A_l are sufficient for those problems to be solvable (see e.g. Francû [3]).

3. Consider a sequence of problems from (3.28),

$$\text{(3.31)} \qquad \langle A_l u^l, v\rangle_V = \langle F_l, v\rangle_V, \qquad v \in V_l,\ l \in \mathbb{N}.$$

As discussed above, the problems (3.31) are solvable for all $l \in \mathbb{N}$. Moreover, there exists a constant $M_1 > 0$ such that

$$\frac{\langle A_l u^l, u^l\rangle_V}{\|u^l\|_V} \leqslant \|F_l\|_{V^*} \leqslant \|F\|_{V^*} + \|F_l - F\|_{V^*} \leqslant M_1.$$

Then, taking the coercivity of A, A_l, $l \geqslant 0$, into account, we see that the sequence $\{\|u^l\|_V\}_{l\in\mathbb{N}}$ is uniformly bounded, i.e.

$$\text{(3.32)} \qquad (\exists M_2 > 0)(\forall l \in \mathbb{N})(\|u^l\|_V \leqslant M_2)$$

and thus there exists a subsequence of $\{u^l\}_{l\in\mathbb{N}}$ weakly converging to some $u^B \in V$. Denote this sequence again by $\{u^l\}_{l\in\mathbb{N}}$. Denote also

$$\Phi \equiv \prod_{i=1}^{m} \{\varphi_i \in C^\infty(\overline{\Omega}) \colon \varphi_i = 0 \text{ on } \Gamma_D\}$$

and let $\varphi \in \Phi$. We have

$$\lim_{l\to\infty} \langle A_l u^l - F, \varphi\rangle_V = \lim_{l\to\infty} \langle F_l - F, \varphi\rangle_V = 0. \tag{3.33}$$

On the other hand,

$$\begin{aligned} \lim_{l\to\infty} \langle A_l u^l - F, \varphi\rangle_V &= \lim_{l\to\infty} \langle A_l u^l - Au^l, \varphi\rangle_V + \lim_{l\to\infty} \langle Au^l - F, \varphi\rangle_V \\ = \lim_{l\to\infty} \sum_{i=1}^{m} \Big[\sum_{P\in\Omega_l} & \mu_2(b_P) a_{i0}(P; (u^l + d)(P))\varphi_i(P) - \int_\Omega a_{i0}(x; u^l + d)\varphi_i(x)\,\mathrm{d}x \Big] \\ &+ \lim_{l\to\infty} \langle Au^l - F, \varphi\rangle_V. \end{aligned}$$

The first limit in the last expression is zero while the second equals to

$$\langle Au^B - F, \varphi\rangle_V.$$

Hence we have from (3.33)

$$(\forall \varphi \in \Phi)(\langle Au^B - F, \varphi\rangle_V = 0).$$

The set Φ being dense in V, we thus conclude that u^B (which is the weak limit of the subsequence $\{u^l\}_{l\in\mathbb{N}}$) is a solution of the problem (3.14)–(3.16). □

Remark 3.1. Clearly, the set

$$\{\varphi_{lP} \in V_l^p \colon (\forall P, Q \in \Omega_l, \varphi_{lP}(Q) = \delta_{PQ})\} \tag{3.34}$$

(δ_{PQ} is the Kronecker symbol) forms a basis of the space V_l^p, i.e. any $v \in V_l$ can be expressed in the form

$$v = \sum_{i=1}^{m} \sum_{P\in\Omega_l} v_{i,lP}\varphi_{lP}. \tag{3.35}$$

Let $\prec$ be a complete ordering of the set Ω_l. Define a mapping $\nu_l \colon \{1, 2, \ldots, N_l\} \to \Omega_l$ by

$$(\forall k_1, k_2, 1 \leqslant k_1,\ k_2 \leqslant N_j)(k_1 < k_2 \Leftrightarrow \nu_l(k_1) \prec \nu_l(k_2)).$$

Then the coefficient vector $v_l^H \equiv ((v_{1,lQ})_{Q\in\Omega_l}, \ldots, (v_{m,lQ})_{Q\in\Omega_l})$ can be understood as a point in a linear space $H_l = \mathbb{R}^{mN_l}$, with its $[(i-1)\cdot N_l + j]$-th component being equal to $v_{i,l\nu(j)}$, $i = 1, \ldots, m$, $j = 1, \ldots, N_l$. The isomorphism between the space H_l and V_l is given by a mapping

$$P_l^G\colon H_l \to V_l, \text{ with } P_l^G v_l^H = \sum_{i=1}^{m} \sum_{Q\in\Omega_l} v_{i,lQ}\varphi_{lQ}. \tag{3.36}$$

If $(\cdot,\cdot)$ is the scalar product in the space H_l and $g_l\colon H_l \to H_l$ and $f_l^H \in H_l$ are defined by

$$(\forall u, v \in H_l)(g_l(u), v) = \langle A_l P_G u_l^H, P_G v_l^H\rangle_V, \tag{3.37}$$

$$(\forall v \in H_l)(f_l^H, v) = \langle F_l, P_G v_l^H\rangle_V, \tag{3.38}$$

then the problem (3.28) in the space V_l is equivalent to the problem

$$g_l(u_l^H) = f_l^H \text{ in } H_l. \tag{3.39}$$

Remark 3.2. Consider the problem (3.13) and suppose, in addition to the assumptions of Theorem 3.3, that the functions $a_{ij}, 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant N$, have the form

$$a_{ij} = a_i(x; u) D^j u_i, \tag{3.40}$$

where $a_i \in C^1(\overline{\Omega} \times \mathbb{R}^m)$, $1 \leqslant i \leqslant m$, and

$$(\exists a_0 > 0,\ a_0' > 0)(\forall i\colon i = 1, \ldots, M)(\forall x \in \Omega)(\forall \xi \in \mathbb{R}^M)(a_0' \geqslant a_i(x;\xi) \geqslant a_0).$$

Then, obviously,

$$(\forall u, w \in V_l)\left(\int_\Omega a_i(x;u) D^j u_i(x) D^j w_i(x)\,\mathrm{d}x = \int_\Omega \overline{a_i}(x;u) D^j u_i(x) D^j w_i(x)\,\mathrm{d}x\right),$$

where the functions $\overline{a_i}\colon \Omega \to \mathbb{R}$ are defined by

$$(\forall t \in T_l)\left(\overline{a_i}\big|_t = \frac{1}{\mu_2(t)} \int_t a_i(x;u)\,\mathrm{d}x\right)$$

As shown in Hackbusch [8, Propositon 3.1.2],

$$(\forall P \in \Omega_l)(\forall v_i \in V^{p_i})\left(\int_\Omega D^j \varphi_{lP}(x) D^j v_i(x)\,\mathrm{d}x = \int_{b_P} \frac{\partial v_i}{\partial n_i}\,\mathrm{d}S\right)$$

and hence the system (3.39) derived from our procedure is the same as that generated by the so-called box integration method (see e.g. Varga [20]), applied to the system (3.13) with the functions $\overline{a_i}(x;u)$ in the place of $a_i(x;u)$. (Note that in the case $a_i = \text{const.}$ the functions a_i and $\overline{a_i}$ are identical.)

4.1. Model problem. In 1950, Van Roosbroeck [17] proposed a system of three coupled nonlinear partial differential equations as a basic mathematical model describing electro-physical behaviour of semiconductor devices. We shall be interested in the following, rather simplified form of these equations, ignoring complications like variable mobilities, oxide regions and avalanche generation rate. Our problem, however, captures some of the difficulties that occur in practice and its satisfactory solution still represents a great challenge to numerical analysis:

$$-\operatorname{div}(\operatorname{grad} u) + \mathrm{e}^{u-v} - \mathrm{e}^{w-u} = D_C, \tag{4.1}$$

$$-\operatorname{div}(\mathrm{e}^{u-v}\operatorname{grad} v) - Q(u,v,w)(\mathrm{e}^{w-v} - 1) = 0 \qquad \text{in } \Omega, \tag{4.2}$$

$$-\operatorname{div}(\mathrm{e}^{w-u}\operatorname{grad} w) + Q(u,v,w)(\mathrm{e}^{w-v} - 1) = 0, \tag{4.3}$$

$$u = u_D,\ v = v_D,\ w = w_D \qquad \text{on } \Gamma_D, \tag{4.4}$$

$$\frac{\partial u}{\partial \vec{n}} = \mathrm{e}^{u-v}\frac{\partial v}{\partial \vec{n}} = \mathrm{e}^{w-u}\frac{\partial w}{\partial \vec{n}} = 0 \qquad \text{on } \Gamma_N. \tag{4.5}$$

where

$$D_C \in L_\infty(\Omega), \quad Q \in C(\mathbb{R}^3) \text{ and } (u_D, v_D, w_D) \in [L_\infty(\overline{\Omega}) \cap C^1(\overline{\Omega})]^3. \tag{4.6}$$

4.2. Weak formulation and existence theorem. First, we reformulate the problem (4.1)–(4.5) in terms of the so-called Slotboom variables u, η, ν, defined by

$$\eta = \mathrm{e}^{-v}, \qquad \nu = \mathrm{e}^{w}.$$

We obtain a boundary value problem in the form (3.13), where $N = 2$, $m = 3$ and (writing a^S_{ij} instead of a_{ij})

$$a^S_{1j}(x;\xi) = \begin{cases} \mathrm{e}^{\xi_{10}}\xi_{20} - \mathrm{e}^{-\xi_{10}}\xi_{30}, & j = 0, \\ \xi_{1j}, & j = 1,2, \end{cases} \tag{4.7}$$

$$a^S_{2j}(x;\xi) = \begin{cases} Q(\xi_{10},\xi_{20},\xi_{30})(\xi_{20}\xi_{30} - 1), & j = 0, \\ \mathrm{e}^{\xi_{10}}\xi_{2j}, & j = 1,2, \end{cases} \tag{4.8}$$

$$a^S_{3j}(x;\xi) = \begin{cases} Q(\xi_{10},\xi_{20},\xi_{30})(\xi_{20}\xi_{30} - 1), & j = 0, \\ \mathrm{e}^{-\xi_{10}}\xi_{3j}, & j = 1,2, \end{cases} \tag{4.9}$$

$$f_1 = D_C, f_2 = f_3 = 0, d_1 = u_D, d_2 = \mathrm{e}^{-v_D}, d_3 = \mathrm{e}^{w_D}, h_1 = h_2 = h_3 = 0. \tag{4.10}$$

We refer to this problem as to the problem (S) and use the notation

$$U_D \equiv (u_D, \mathrm{e}^{-v_D}, \mathrm{e}^{w_D}).$$

However, due to exponentials in (4.7)–(4.9), no p_i, $1 < p_i < \infty$, $i = 1, 2, 3$, satisfy the condition (A2), so direct application of Theorem 3.2 is not possible. Nevertheless, note that if $V^\infty = [V^2 \cap L_\infty(\Omega)]^3$ and $V = [V^2]^3$, then the following expressions define a mapping $A^S \colon V^\infty \to V^*$ and a functional $f^S \in V^*$:

$$(\forall U \in V^\infty, U \equiv (u, \eta, \nu))(\forall \Phi \in V, \Phi \equiv (\varphi_1, \varphi_2, \varphi_3))$$

$$\langle\langle A^S U, \Phi \rangle\rangle_V = \sum_{i=1}^{3} \sum_{j=0}^{2} \int_\Omega a^S_{ij}(x; U + U_D, \nabla(U + U_D)) D^j \varphi_i \, \mathrm{d}x), \tag{4.11}$$

$$(\forall \Phi \in V) \quad \langle\langle f^S, \Phi \rangle\rangle_V = \int_\Omega D_C(x) \varphi_1(x) \, \mathrm{d}x). \tag{4.12}$$

Definition 4.1. Let $W^\infty \equiv [W^{1,2}(\Omega) \cap L_\infty(\Omega)]^3$. We say that $U_S \in W^\infty$ *is a solution of the problem (S) in the space* W^∞, if

$$U_S = U_S^* + U_D, \tag{4.13}$$

where $U_S^* \in V^\infty$ and

$$A^S U_S^* = f^S \text{ in } V^*. \tag{4.14}$$

Theorem 4.1. *There exists at least one solution of the problem* (S) *in the space* W^∞.

P r o o f. of this theorem will be divided into two steps. First, as in Gröger [6], we shall formulate a modified problem such that its solutions are also solutions of (S). Then we shall apply Theorem 3.2 to this modified problem. □

Definition 4.2. Let $r < s$ be real numbers and $g \colon \mathbb{M} \to \mathbb{R}$ ($\mathbb{M}$ is an arbitrary set) any real function. We define $P_{rs}g \colon \mathbb{M} \to \mathbb{R}$ by

$$(\forall x \in \mathbb{M})(P_{rs}g)(x) = \begin{cases} r & \text{if } g(x) \leqslant r, \\ g(x) & \text{if } r < g(x) < s, \\ s & \text{if } s \leqslant g(x). \end{cases}$$

For $r = -s$ we write only P_s.

Definition 4.3. Let $F \geqslant \max\{\|v_D\|_\infty, \|w_D\|_\infty\}$, $G = \mathrm{e}^{-F}$, $H = \mathrm{e}^F$. We choose $E \geqslant \|u_D\|_\infty$ such that

$$\mathrm{e}^{F-E} - \mathrm{e}^{E-F} + D_C \leqslant 0, \quad \mathrm{e}^{E-F} - \mathrm{e}^{F-E} + D_C \geqslant 0$$

and define $A^{Sr}: V \to V^*$ by

$$(\forall U \in V, U \equiv (u,\eta,\nu))(\forall \Phi \in V, \Phi \equiv (\varphi_1,\varphi_2,\varphi_3))$$

$$(4.15) \qquad (\langle A^{Sr}U, \Phi\rangle_V = \sum_{i=1}^{3}\sum_{j=0}^{2}\int_\Omega a_{ij}^{Sr}(x; U+U_D, \nabla(U+U_D))D^j\varphi_i \, \mathrm{d}x),$$

where

$$(\forall x \in \Omega)(\forall i, i=1,2,3)(\forall j, j=0,1,2)(\forall U = (u,\eta,\nu)\colon \Omega \to \mathbb{R}^3)$$

$$(4.16) \qquad (a_{ij}^{Sr}(x; U, \nabla U) = a_{ij}^{S}(x; P_E u, P_{GH}\eta, P_{GH}\nu, \nabla U)).$$

We say that $U \in W$ *is a weak solution of the problem (Sr),* if

$$(4.17) \qquad U = U^* + U_D,$$

where $U^* \in V$ and

$$(4.18) \qquad A^{Sr}U^* = f^S \text{ in } V^*.$$

Lemma 4.1. (Gröger [6, Lemma 1]) *Let U be a weak solution of the problem* (Sr). *Then U is also a solution of the problem* (S) *in the space W^∞.*

Lemma 4.2. *There exists at least one weak solution of the problem* (Sr).

Proof. We shall verify the assumptions of Theorem 3.2:

(A1) is obvious.

(A2) with the coefficients $p_1 = p_2 = p_3 = 2$ can be verified easily. The only arguments of the functions a_{ij}^{Sr} that can cause a_{ij}^{Sr} to grow to infinity are those with ∇U. But the dependence of a_{ij}^{Sr} on ∇U is linear, in the worst case.

(A3). In (3.10), we can choose

$$c = \min(1, \mathrm{e}^{-E}),\ \theta_1 = -\mathrm{e}^{-F},\ \theta_2 = \theta_3 = -\sup_{x\in\Omega, \xi\in\mathbb{R}^9} a_{20}^{Sr}(x;\xi).$$

(A4) is also easy, because of

$$(\forall x \in \Omega)(\forall \xi \in \mathbb{R}^3)(\forall \eta, \nu \in \mathbb{R}^6)$$
$$\sum_{i=1}^{3}\sum_{j=1}^{2}(a_{ij}^{Sr}(x;\xi,\eta) - a_{ij}^{Sr}(x;\xi,\nu))(\eta_{ij} - \nu_{ij})$$
$$\geqslant \sum_{j=1}^{2}(\eta_{1j} - \nu_{1j})^2 + \mathrm{e}^{-E}(\eta_{2j} - \nu_{2j})^2 + \mathrm{e}^{-E}(\eta_{3j} - \nu_{3j})^2.$$

(D) follows from (4.6). □

The assertion of Lemma 4.2 now follows from Theorem 3.2.

Discretization.

Theorem 4.2. *Consider the problem* (S) *with* (4.10) *and* $D_C \in C(\overline{\Omega})$. *Then the assertions of Theorem* 3.3 *hold.*

P r o o f. Taking Lemma 4.2 and continuity of D_C into account, the assumptions of Theorem 3.3 can be easily verified. □

R e m a r k 4.1. Note that if we formulate the problem (3.28) for some $l \geqslant 0$, then the integrals of the form

$$\int_t \mathrm{e}^{P_E u}\,\mathrm{d}x$$

($t \in T_l$) are to be evaluated. But this can be done easily, because u is a piecewise linear function. If, for example, $|u(x)| \leqslant E$ for all $x \in t$, and U_1, U_2, U_3 ($U_1 \neq U_2 \neq U_3 \neq U_1$) denote the values of $u(x)$ at the vertices of the triangle $t \in T_l$, we obtain after some calculation

$$\int_t \mathrm{e}^{P_E u}\,\mathrm{d}x = \int_t \mathrm{e}^{u}\,\mathrm{d}x = \frac{2\mu(t)}{U_2 - U_1}\left[\frac{\mathrm{e}^{U_3} - \mathrm{e}^{U_2}}{U_3 - U_2} - \frac{\mathrm{e}^{U_3} - \mathrm{e}^{U_1}}{U_3 - U_1}\right].$$

Conclusion

In this paper, we stop at the point when a system of (possibly nonlinear) algebraic equations is generated. In the forthcoming paper Pospíšek [16], a convergent algorithm suitable for solving such a system is proposed.

References

[1] *J. F. Bürgler, R. E. Bank, W. Fichtner, R. K. Smith*: A new discretization scheme for the semiconductor current continuity equations. IEEE Trans. on CAD *8* (1989), 479–489.
[2] *Z. Chen*: Hybrid variable finite elements for semiconductor devices. Comput. Math. Appl. *19* (1990), 65–73.
[3] *J. Franců*: Monotone operators. A survey directed to applications to differential equations. Apl. Mat. *35* (1990), 257–301.
[4] *S. Fučík, A. Kufner*: Nonlinear Differential Equations. Czech edition – SNTL, Prague, 1978; English translation – Elsevier, Amsterdam. 1980.
[5] *H. Gajewski*: On uniqueness and stability of steady-state carrier distributions in semiconductors. Proc. Equadiff Conf. 1985. Springer, Berlin, 1986, pp. 209–219.
[6] *K. Gröger*: On steady-state carrier distributions in semiconductor devices. Apl. Mat. *32* (1987), 49–56.
[7] *H. K. Gummel*: A self-consistent iterative scheme for one-dimensional steady state transistor calculations. IEEE Trans. on Electron Devices *ED-11* (1964), 455–465.

[8] *W. Hackbusch*: On first and second order box schemes. Computing *41* (1989), 277–296.
[9] *J. W. Jerome*: Consistency of semiconductor modelling: An existence/stability analysis for the stationary Van Roosbroeck system. SIAM J.Appl.Math. *45* (1985), 565–590.
[10] *P. A. Markowich*: The Stationary Semiconductor Device Equations. Springer-Verlag, Wien–New York 1986.
[11] *P. A. Markowich, M. Zlámal*: Inverse-average-type finite element discretizations of self--adjoint second order elliptic problems. Math. Comp. *51* (1988), 431–449.
[12] *J. Miller*: Mixed FEM for semiconductor devices. In: Numerical Mathematics. Singapore 1988. Proc. Int. Conf. (R.P. Agarwal, Y.M. Chow, S.J. Wilson, eds.). Basel, Birkhäuser Verlag, 1988, pp. 349–356.
[13] *M. S. Mock*: Analysis of Mathematical Models of Semiconductor Devices. Boole Press, Dublin, 1983.
[14] *J. Nečas*: Introduction to the Theory of Nonlinear Elliptic Equations. Teubner Texte zur Math. 52, Leipzig, 1987.
[15] *M. Pospíšek*: Mathematical Methods in Semiconductor Device Modelling. PhD Thesis, MÚ ČSAV, Prague, 1991. (In Czech.)
[16] *M. Pospíšek*: Convergent algorithms suitable for the solution of the semiconductor device equations. To be published.
[17] *W.V. Van Roosbroeck*: Theory of flow of electrons and holes in germanium and other semiconductors. Bell Syst. Tech. J. *29* (1950), 560–607.
[18] *D. L. Scharfetter, H. K. Gummel*: Large signal analysis of a silicon Read diode oscillator. IEEE Trans. Electron Devices *ED-16* (1969), 64–77.
[19] *N. Shigyo, T. Wada, S. Yasuda*: Discretization problem for multidimensional current flow. IEEE Trans. on CAD *8* (1989), 1046–1050.
[20] *R. S. Varga*: Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, NJ, 1962.

Author's address: *Miroslav Pospíšek*, T-SOFT, spol. s r.o., Novodvorská 994, 140 00 Praha 4.