

## NUMERICAL SIMULATION OF SEMICONDUCTOR DEVICES

Franco BREZZI\*, Luisa Donatella MARINI, Paola PIETRA

*Istituto di Analisi Numerica del C.N.R., Pavia, Italy*

The theoretical analysis of the mixed and hybrid exponential fitting methods, introduced in [1], is extended to the case of a non-zero source term. The methods are then studied from the numerical point of view and their upwinding features are analysed.

### 1. Introduction

The equations commonly used for semiconductor device simulation are the Van Roosbroeck equations [2–4]. They are considered to be a reasonably satisfactory model for materials like silicon or germanium and within a certain range of applications. For the use of other possible models better suited to describe the behaviour of other semiconductor materials such as gallium arsenide, we refer for example to [5–7]. The Van Roosbroeck equations involve a number of physical quantities and parameters with enormously different orders of magnitude (from the  $10^{-18}$  A of the electron charge to the  $10^{17}$  cm $^{-3}$  of a realistic doping profile). It is clear that in this situation a scaling is convenient (and recommended). We refer to [3] and [8] for two different types of possible scalings and we present here the Van Roosbroeck scaled equations obtained following [3] in the special case of a stationary problem with constant coefficients:

$$\begin{aligned} l \Delta \psi &= n - p + D, & \text{in } \Omega \subset \mathbb{R}^2, \\ \operatorname{div}(\nabla n - n \nabla \psi) &= R, & \text{in } \Omega, \\ \operatorname{div}(\nabla p + p \nabla \psi) &= R, & \text{in } \Omega, \\ &+ \text{boundary conditions.} \end{aligned} \tag{1.1}$$

The unknown  $\psi$ ,  $n$  and  $p$  represent the (scaled) electric potential and the (scaled) concentration of negative and positive charges, respectively.  $D$  describes the doping profile and  $R$  is the so-called generation-recombination term which takes into account the possible ‘birth’ and ‘death’ of charges. The domain  $\Omega$  is scaled to have diameter equal to 1. The first equation of (1.1) is the classical Poisson’s equation for the potential, and the second and third equations are the continuity equations for negative and positive current densities  $J^n, J^p$ :

$$J^n = \nabla n - n \nabla \psi, \quad J^p = \nabla p + p \nabla \psi.$$

The choice of boundary conditions depends on the applications. As an example, we can assume the boundary to be split into two disjoint parts:  $\partial \Omega = \Gamma_0 \cup \Gamma_1$ . On  $\Gamma_0$ ,  $\psi$ ,  $n$  and  $p$  are

\* Also at Dipartimento di Meccanica Strutturale dell’Università, Pavia, Italy.

prescribed, while on  $\Gamma_1$  vanishing outward normal components for  $\mathbf{J}^n, \mathbf{J}^p$ , and for the electric field  $\mathbf{E} = -\nabla\psi$  are assumed. The parameter  $l$  is the dimensionless scaling parameter, whose value lies typically between  $10^{-3}$  and  $10^{-6}$ . For results and additional references on the limit behaviour of the problem for different scalings we refer to [3] and [9]. Existence results for the stationary problems hold under fairly general assumptions [10, 3, 4], while uniqueness is not true in general. It holds only in some special cases [3, 4, 10, 9].

When dealing with system (1.1) from a numerical point of view, one key point is the coupling of the equations. The unknowns are strongly coupled and the choice of the iterative method to decouple them is a difficult task. We recall here that the most used iterative algorithms are variants of the Newton method, plus some other tricks, such as continuation in some parameter. We refer to [11, 12, 3] for detailed descriptions of possible algorithms.

A second problem to be faced is the choice of a proper discretization for the single equations. The Poisson equation is not a source of problems; every reasonable approximation scheme will work. Difficulties arise when discretizing the current equations, which are advection dominated (due to the large size of  $\nabla\psi$  in a consistent portion of the domain). The most popular and effective method for one-dimensional problems is the so-called Scharfetter–Gummel method [13], better known in other fields as the exponential fitting method [14, 15]. The main feature of this scheme is to be current-preserving. This property is particularly important in the applications, as the current is possibly the most relevant unknown of the problem. In two dimensions many methods have been proposed: Mock's method [16], the Box-scheme (which is the most widely used in the codes) [17] and, more recently, Markowich–Zlámal [18] and the modified method of characteristics (the latter being essentially introduced for transient problems) [19]. All except the latter, are attempts to extend the Scharfetter–Gummel method to 2-D problems. In the present paper we concentrate our analysis on the so-called 2-D exponential fitting methods recently introduced in [20, 1]. These new schemes, essentially based on mixed or hybrid finite element methods, are both current preserving, and appear to be a very natural and satisfactory extension of the Scharfetter–Gummel method.

In order to describe the methods, we shall focus our attention on one equation only, the current equation for positive charges. We shall denote by  $u$  the unknown and assume  $\psi$  to be a given function (piecewise linear in the computations).

$$\begin{aligned} &\text{Find } u \in H^1(\Omega), \text{ such that} \\ &\operatorname{div}(\nabla u + u \nabla \psi) = f, \quad \text{in } \Omega \subset \mathbb{R}^2, \\ &u = g, \quad \text{on } \Gamma_0 \subset \partial\Omega, \\ &\frac{\partial u}{\partial n} + u \frac{\partial \psi}{\partial n} = 0, \quad \text{on } \Gamma_1 = \partial\Omega \setminus \Gamma_0. \end{aligned} \tag{1.2}$$

The structure behind the methods is the following. First, we introduce the classical change of variables from  $u$  to the Slotboom variable  $\rho$

$$u = \rho e^{-\psi}. \tag{1.3}$$

In terms of  $\rho$ , problem (1.2) can be written in the symmetric form:

$$\begin{aligned}
& \text{Find } \rho \in H^1(\Omega), \text{ such that} \\
& \operatorname{div}(e^{-\psi} \nabla \rho) = f, \quad \text{in } \Omega, \\
& \rho = \chi := e^{\psi} g, \quad \text{on } \Gamma_0, \\
& \frac{\partial \rho}{\partial n} = 0, \quad \text{on } \Gamma_1.
\end{aligned} \tag{1.4}$$

Then, we discretize problem (1.4) with mixed or hybrid schemes. Finally, we go back to the variable  $u$  by means of (1.3) and we solve in  $u$ . If a proper decomposition is introduced, the discretization of (1.4) provides a symmetric positive-definite M-matrix and the M-matrix property is preserved in the final problem for the variable  $u$ .

An outline of the paper is as follows. In Sections 2 and 3 we introduce the mixed and hybrid approximations of problem (1.4), respectively and we extend the analysis of [1] to the case  $f \neq 0$ . Such an extension is straightforward for the mixed method, while it needs some care when dealing with the hybrid approach. In Section 4 we quit the theoretical framework of [1] and we go into more numerical details. In particular, we analyse the qualitative behaviour of the schemes for large  $|\nabla \psi|$  and underline their upwinding features. We also point out some numerical difficulties that can arise. Section 5 is devoted to the presentation of numerical results obtained using both schemes on a number of examples.

## 2. Mixed methods

Let us consider (see [21]) a regular sequence  $\{T_h\}_h$  of decompositions of  $\Omega$  into triangles  $T$  (for the sake of simplicity we shall assume that  $\Omega$  is a polygon). According to [22], we define the following set of polynomial vectors:

$$\operatorname{RT}(T) = \{\tau = (\tau_1, \tau_2) \mid \tau_1 = \alpha + \beta x, \tau_2 = \gamma + \beta y, \alpha, \beta, \gamma \in \mathbb{R}\}, \tag{2.1}$$

and, for every triangulation  $T_h$ , we set:

$$V_h = \{\tau \in [L^2(\Omega)]^2 \mid \operatorname{div} \tau \in L^2(\Omega), \tau \cdot n = 0 \text{ on } \Gamma_1, \tau|_T \in \operatorname{RT}(T) \forall T \in T_h\}, \tag{2.2}$$

$$W_h = \{\varphi \in L^2(\Omega) \mid \varphi|_T \in P_0(T) \forall T \in T_h\}. \tag{2.3}$$

Then, the mixed discretization of (1.4) is defined as follows:

$$\begin{aligned}
& \text{Find } J_h \in V_h \text{ and } \rho_h \in W_h, \text{ such that} \\
& \int_{\Omega} e^{\psi} J_h \cdot \tau \, dx + \int_{\Omega} \rho_h \operatorname{div} \tau \, dx = \int_{\Gamma_0} \chi \tau \cdot n \, d\Gamma \quad \forall \tau \in V_h, \\
& \int_{\Omega} \varphi \operatorname{div} J_h \, dx = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in W_h.
\end{aligned} \tag{2.4}$$

It is clear that  $\rho_h$  will be an approximation of the solution  $\rho$  of (1.4) while  $J_h$  will be an approximation of the current

$$\mathbf{J} = e^{-\psi} \nabla \rho = \nabla u + u \nabla \psi. \quad (2.5)$$

In particular, the first equation of (2.4) is a discretized version of (2.5), and the second equation of (2.4) is a discretized version of  $\operatorname{div} \mathbf{J} = f$ . In the computations one usually assumes that  $\psi$  is piecewise linear in each  $T$ , so that the integrals in (2.4) can be computed exactly. Uniqueness results for problem (2.4) follow easily from the general theory of [23].

We remark that the condition  $\operatorname{div} \boldsymbol{\tau} \in L^2(\Omega)$  in the definition (2.2) of  $V_h$  implies that every  $\boldsymbol{\tau}$  in  $V_h$  has a continuous normal component when passing from one element to another. This means, in particular, that the current  $\mathbf{J}_h$  is preserved.

The algebraic treatment of the system (2.4) needs some care. The matrix associated with (2.4) has the form

$$\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix} \quad (2.6)$$

and it is not positive-definite. A way to avoid this inconvenience was introduced by Fraeijs de Veubeke [24] for a similar problem. The idea is to relax the continuity requirement in the space definition (2.2) and to enforce it back by using interelement Lagrange multipliers. We give now a more detailed description of the procedure. For this, we first set

$$\tilde{V}_h = \{ \boldsymbol{\tau} \in [L^2(\Omega)]^2 \mid \boldsymbol{\tau}|_T \in \mathbf{RT} \, \forall T \in T_h \}. \quad (2.7)$$

Then, calling  $E_h$  the set of edges of  $T_h$ , we define, for every function  $\xi \in L^2(\Gamma_0)$ :

$$\Lambda_{h,\xi} = \{ \mu \mid \mu \in L^2(E_h); \mu|_e \in P_0 \, \forall e \in E_h; \int_e (\mu - \xi) \, ds = 0 \, \forall e \subset \Gamma_0 \}. \quad (2.8)$$

We consider now the mixed-equilibrium discretization of problem (1.4):

$$\begin{aligned} & \text{Find } \tilde{\mathbf{J}}_h \in \tilde{V}_h, \tilde{\rho}_h \in W_h, \tilde{\lambda}_h \in \Lambda_{h,\chi}, \text{ such that} \\ & \int_{\Omega} e^{\psi} \tilde{\mathbf{J}}_h \cdot \boldsymbol{\tau} \, dx + \sum_T \int_T \tilde{\rho}_h \operatorname{div} \boldsymbol{\tau} \, dx - \sum_T \int_{\partial T} \tilde{\lambda}_h \boldsymbol{\tau} \cdot \mathbf{n} \, ds = 0 \quad \forall \boldsymbol{\tau} \in \tilde{V}_h, \\ & \sum_T \int_T \varphi \operatorname{div} \tilde{\mathbf{J}}_h \, dx = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in W_h, \\ & \sum_T \int_{\partial T} \mu \tilde{\mathbf{J}}_h \cdot \mathbf{n} \, ds = 0 \quad \forall \mu \in \Lambda_{h,0}. \end{aligned} \quad (2.9)$$

It is easy to see that the linear system (2.9) has a unique solution and that

$$\tilde{\mathbf{J}}_h \equiv \mathbf{J}_h, \quad \tilde{\rho}_h \equiv \rho_h. \quad (2.10)$$

Moreover,  $\tilde{\lambda}_h$  is a good approximation of  $\rho$  at the interelements (see [25] for detailed proofs). The matrix associated with (2.9) now has the structure

$$\begin{bmatrix} A & B & C \\ B^* & 0 & 0 \\ C^* & 0 & 0 \end{bmatrix}, \quad (2.11)$$

where  $A$  is block-diagonal (each block corresponding to a single element). In other words,  $\tilde{\mathbf{J}}_h$  can be eliminated, element by element, by static condensation. This leaves us with the matrix

$$\begin{vmatrix} B^* A^{-1} B & B^* A^{-1} C \\ C^* A^{-1} B & C^* A^{-1} C \end{vmatrix}, \quad (2.12)$$

where  $B^* A^{-1} B$  is block-diagonal (actually, in the present case, diagonal) and non-singular. Hence,  $\tilde{\rho}_h$  can also be eliminated, at the element level, by static condensation. This leaves us, finally, with a matrix (acting only on the interelement multiplier  $\tilde{\lambda}_h$ ) which can be proved to be a symmetric positive-definite matrix, which is an M-matrix if the triangulation is of weakly acute type (each angle of every  $T \in T_h$  is  $\leq \pi/2$ ). This general strategy for dealing with problems of the type (2.4) can be simplified in the present case. Let us see it in more detail. Denoting by  $\bar{f}_T$  the mean value of  $f$  in  $T$  and using definition (2.1), the second equation in (2.9) implies that, for every  $T$  in  $T_h$  we have

$$\operatorname{div} \tilde{\mathbf{J}}_h|_T = 2\beta = \bar{f}_T := \frac{1}{|T|} \int_T f \, dx. \quad (2.13)$$

Hence,  $\tilde{\mathbf{J}}_h$  can be split in two parts

$$\tilde{\mathbf{J}}_h = \mathbf{J}^0 + \mathbf{J}^f, \quad (2.14)$$

where

$$\mathbf{J}^0|_T = (\alpha, \gamma), \quad \mathbf{J}^f|_T = \frac{\bar{f}_T}{2} \mathbf{x}, \quad (2.15)$$

( $\mathbf{x}$  being the coordinate vector  $\mathbf{x} = (x_1, x_2)$ ). Taking now in the first equation of (2.9)  $\boldsymbol{\tau} = (1, 0)$  and  $\boldsymbol{\tau} = (0, 1)$  in  $T$ , ( $\boldsymbol{\tau} = \mathbf{0}$  elsewhere), we obtain, with the notation of Fig. 1

$$\mathbf{J}^0 \int_T e^\psi \, dx = \sum_{k=1}^3 |e_k| \tilde{\lambda}_k \mathbf{n}^{(k)} - \int_T e^\psi \mathbf{J}^f \, dx. \quad (2.16)$$

We make for the basis functions  $\mu$  in  $\Lambda_{h,0}$  the obvious choice  $\mu = 1$  on one edge and  $\mu = 0$  on the others. Then, inserting (2.16) in the third equation of (2.9), one can easily compute the contributions of the triangle  $T$  to the stiffness matrix and to the right-hand side:

$$a_{ij}^T = \frac{\boldsymbol{\nu}^{(i)} \cdot \boldsymbol{\nu}^{(j)}}{\int_T e^\psi \, dx}, \quad b_i^T = \frac{\int_T e^\psi \mathbf{J}^f \cdot \boldsymbol{\nu}^{(i)} \, dx}{\int_T e^\psi \, dx} - \int_{e_i} \mathbf{J}^f \cdot \mathbf{n}^{(i)} \, ds, \quad (2.17)$$

where we set

$$\boldsymbol{\nu}^{(k)} = |e_k| \mathbf{n}^{(k)}, \quad k = 1, 3. \quad (2.18)$$

As already pointed out, the integrals appearing in (2.17) can be computed exactly, since  $\psi$  has been assumed linear in each triangle. It is now easy to check that the final matrix, obtained by

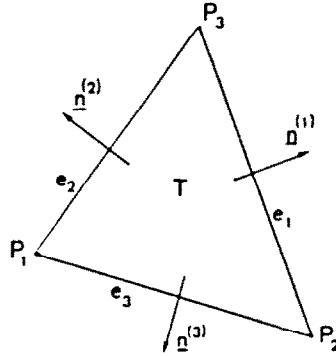


Fig. 1.

assembling the elementary contributions (2.17), is a symmetric positive-definite matrix and it is an M-matrix if the triangulation is of weakly acute type. Note that the matrix (2.17) corresponds to a linear non-conforming finite element approximation of problem (1.4) where the harmonic average of the coefficient  $e^{-\psi}$  is taken instead of the usual average.

We are now ready to introduce an approximation of our original unknown  $u$ . For this we define, for every  $\zeta \in L^2(E_h)$ ,  $\zeta^I$  to be the  $L^2$ -projection of  $\zeta$  onto  $\Lambda_{h,\zeta}$ , that is:

$$\zeta^I|_e = \frac{1}{|e|} \int_e \zeta \, ds \quad \forall e \in E_h. \quad (2.19)$$

Using now the change of variable  $\tilde{\lambda}_h = (e^\psi)^I u_h$ , we can rewrite (2.9) in the form:

$$\begin{aligned} &\text{Find } J_h \in \tilde{V}_h, \rho_h \in W_h, u_h \in \Lambda_{h,g}, \text{ such that} \\ &\int_{\Omega} e^\psi J_h \cdot \tau \, dx + \sum_T \int_T \rho_h \operatorname{div} \tau \, dx - \sum_T \int_{\partial T} e^\psi u_h \tau \cdot n \, ds = 0 \quad \forall \tau \in \tilde{V}_h, \\ &\sum_T \int_T \varphi \operatorname{div} J_h \, dx = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in W_h, \\ &\sum_T \int_{\partial T} \mu J \cdot n \, ds = 0 \quad \forall \mu \in \Lambda_{h,0}, \end{aligned} \quad (2.20)$$

where we went back to the original names for the variables  $J_h$  and  $\rho_h$  in virtue of (2.10). If one eliminates  $J_h$  and  $\rho_h$  in (2.20) by static condensation, the resulting matrix, acting on the unknown  $u_h$  only, is an M-matrix, if the decomposition is of weakly acute type. Actually, it can be obtained from the matrix (2.17), after assembling, by multiplying each column by the value of  $(e^\psi)^I$  on the corresponding edge.

**REMARK 2.1.** Different choices of mixed finite element methods can be made. We refer to [1] for the use of quadrilateral decompositions of  $\Omega$  in this context and to [26] for general references on mixed methods.

We conclude this section with an error estimate result on the current. The proof is a slight modification of that given in [1, Theorem 3.1].

**THEOREM 2.1.** *Let  $u$  be the solution of (1.2), let  $\mathbf{J} := e^{-\psi} \nabla(e^\psi u) = \nabla u + u \nabla \psi$  and  $(\mathbf{J}_n, \rho_h)$  be the solution of (2.4). Then*

$$\|\mathbf{J} - \mathbf{J}_h\|_a \leq \|\mathbf{J} - \Pi_h \mathbf{J}\|_a, \quad (2.21)$$

where  $\|\mathbf{J}\|_a^2 = \int_\Omega e^\psi \mathbf{J} \cdot \mathbf{J} \, dx$ , and  $\Pi_h \mathbf{J} \in V_h$  is defined locally as

$$\int_e (\Pi_h \mathbf{J} - \mathbf{J}) \cdot \mathbf{n} \, ds = 0 \quad \forall e \in E_h. \quad (2.22)$$

### 3. Hybrid methods

We consider now another class of non-standard methods for dealing with problem (1.4), namely, the class of hybrid methods. Let again  $\{T_h\}_h$  be a regular sequence of decompositions of  $\Omega$  into triangles  $T$  [21]. For each  $T_h$  in the sequence and for every function  $\xi \in C^0(\Gamma_0)$  we introduce the sets

$$V_h = \{\boldsymbol{\tau} \in [L^2(\Omega)]^2 \mid \boldsymbol{\tau}|_T = (\alpha, \gamma) \quad \forall T \in T_h, \alpha, \gamma \in \mathbb{R}\}, \quad (3.1)$$

$$W_{h,\xi} = \{\varphi \in C^0(\bar{\Omega}) \mid \varphi|_T \in P_1(T) \quad \forall T \in T_h, \varphi = \xi \text{ at nodes on } \Gamma_0\}. \quad (3.2)$$

Then the hybrid discretization of (1.4) can be written as [26]:

$$\begin{aligned} &\text{Find } (\mathbf{J}_h, \rho_h) \in V_h \times W_{h,\chi}, \text{ such that} \\ &\int_\Omega e^\psi (\mathbf{J}_h + \mathbf{J}^f) \cdot \boldsymbol{\tau} \, dx - \int_\Omega \boldsymbol{\tau} \cdot \nabla \rho_h \, dx = 0 \quad \forall \boldsymbol{\tau} \in V_h, \\ &\int_\Omega \mathbf{J}_h \cdot \nabla \varphi \, dx = - \sum_{T \in T_h} \int_{\partial T} \varphi \mathbf{J}^f \cdot \mathbf{n} \, ds \quad \forall \varphi \in W_{h,0}. \end{aligned} \quad (3.3)$$

In (3.3)  $\mathbf{J}^f$  should be a solution of the equation  $\operatorname{div} \mathbf{J}^f = f$  in  $T$  for all  $T$ . Without substantial loss of accuracy, one can choose instead  $\mathbf{J}^f$  satisfying  $\operatorname{div} \mathbf{J}^f = \bar{f}_T$ , where  $\bar{f}_T$  is the mean value of  $f$  in  $T$ , so that  $\mathbf{J}^f|_T$  can be chosen, for instance, as in (2.15). The integrals appearing in (3.3) can be computed exactly, since  $\psi$  is piecewise linear. However, for reasons that will be made clear in the next section, it is convenient to use a suitable numerical integration. For the moment, let us denote by  $I_T$  the quadrature formula

$$I_T(\varphi) \approx \int_T \varphi \, dx. \quad (3.4)$$

Note that  $I_T$  is still to be chosen. However, we can already assume (3.4) to be exact for a constant  $\varphi$ . In order to simplify the notation, we can introduce the piecewise constant function  $\bar{\psi}$  defined in each  $T$  by

$$\int_T e^{\bar{\psi}} = I_T(e^\psi), \quad (3.5)$$

and then use  $\bar{\psi}$  instead of  $\psi$  in the stiffness matrix of (3.3).

It is easy to check that problem (3.3) has a unique solution [23]. It is also clear that  $\rho_h$  will be an approximation of the solution  $\rho$  of (1.4), while  $\mathbf{J}_h + \mathbf{J}^f$  will be an approximation of the current  $\mathbf{J}$ . In particular, the first equation of (3.3) is the discrete version of (2.5). The second equation of (3.3) ensures a weak conservation property of the current. After integration by parts, it implies

$$\sum_T \int_{\partial T} (\mathbf{J}_h + \mathbf{J}^f) \cdot \mathbf{n} \varphi \, ds = 0 \quad \forall \varphi \in W_{h,0}. \quad (3.6)$$

Then, with the obvious choice for the basis-functions in  $W_{h,0}$ :  $\varphi = 1$  at a node (=vertex)  $P$  and  $\varphi = 0$  at the other nodes, one obtains from (3.6) that the jumps of the flux of  $\mathbf{J}_h + \mathbf{J}^f$  across the edges having  $P$  in common sum up to zero.

For hybrid schemes, the solution of the algebraic system associated with (3.3) is easier than for mixed methods. As before, the matrix form of system (3.3) is the following:

$$\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix} \begin{bmatrix} \mathbf{J}_h \\ \rho_h \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}, \quad (3.7)$$

where here  $A$  is a block-diagonal matrix, since no continuity requirement is made in  $V_h$ . One can then eliminate  $\mathbf{J}_h$  by static condensation. The resulting matrix  $H = B^* A^{-1} B$ , acting on the nodal values of  $\rho_h$ , is symmetric and positive-definite. Also, it will be an M-matrix if the triangulation is of weakly acute type. For the convenience of the reader, we write explicitly the contributions of an element  $T$  to the stiffness matrix and to the right-hand side

$$a_{ij}^T = \frac{\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)}}{4 \int_T e^{\bar{\psi}} \, dx}, \quad (3.8)$$

$$b_i^T = - \frac{\int_T e^{\psi} \mathbf{J}^f \cdot \mathbf{v}^{(i)} \, dx}{2 \int_T e^{\bar{\psi}} \, dx} + \frac{\int_T \mathbf{J}^f \cdot \mathbf{v}^{(i)} \, dx}{2|T|} - \frac{1}{3} |T| \bar{f}_T, \quad (3.9)$$

where  $\mathbf{v}^{(k)}$  is defined as in (2.18). One can then go back to the unknown  $u_h$  via the discrete analogue of (1.3); this corresponds to multiply the  $j$ th column of the matrix  $H$  by  $e^{\psi_j}$  and does not change the M-property of the matrix.

**REMARK 3.1.** Note that the matrix (3.8) corresponds to a linear conforming finite element approximation of problem (1.4), where a sort of harmonic average of the coefficient  $e^{-\psi}$  is taken instead of the usual average. More precisely, since (3.5) implies:

$$e^{-\bar{\psi}}|_T = \frac{|T|}{I_T(e^{\psi})}, \quad (3.10)$$

a linear conforming approximation of problem (1.4) with ‘harmonic’ average can be written as follows:



Find  $\tilde{\rho}_h \in W_{h,x}$ , such that

$$\int_{\Omega} e^{-\tilde{\psi}} \nabla \tilde{\rho}_h \cdot \nabla \varphi \, dx = - \sum_T \int_T \bar{f}_T \varphi \, dx \quad \forall \varphi \in W_{h,0}. \quad (3.11)$$

It is clear that the matrix associated with (3.11) coincides with the matrix  $H$  given by (3.8), as we said before, while the contribution of each triangle to the right-hand side of (3.11) is given by

$$\tilde{b}_i^T = -\frac{1}{3} |T| \bar{f}_T, \quad (3.12)$$

which is much simpler than (3.9). On the other hand, (3.11) will produce an approximation of the current

$$\tilde{J}_h|_T = (e^{-\tilde{\psi}} \nabla \tilde{\rho}_h)|_T, \quad (3.13)$$

which will no longer have continuity properties. However, since the difference between the two problems only consists of a different handling of the right-hand side, the two solutions will coincide if we can choose  $J^f$  in (3.9) such that

$$-\frac{\int_T e^{\psi} J^f \cdot \nu^{(i)} \, dx}{2 \int_T e^{\tilde{\psi}} \, dx} + \frac{\int_T J^f \cdot \nu^{(i)} \, dx}{2|T|} = 0 \quad \forall T \in T_h \quad \forall \nu^{(i)}. \quad (3.14)$$

This is equivalent to choosing  $J^f$  such that

$$-\frac{\int_T e^{\psi} J^f \, dx}{2 \int_T e^{\tilde{\psi}} \, dx} + \frac{\int_T J^f \, dx}{2|T|} = 0 \quad \forall T \in T_h. \quad (3.15)$$

Note that in the hybrid formulation (3.3)  $J^f$  is any solution of the equation  $\operatorname{div} J^f = \bar{f}_T$  in  $T$  for all  $T$ . Hence, we can look for  $J^f = (J_1^f, J_2^f)$  of the form

$$\begin{aligned} J_1^f &= \frac{\bar{f}_T}{2} \tilde{x}_1 + \theta_1 \tilde{x}_1 + \theta_2 \tilde{x}_2, \\ J_2^f &= \frac{\bar{f}_T}{2} \tilde{x}_2 - \theta_1 \tilde{x}_2 + \theta_2 \tilde{x}_1, \end{aligned} \quad (3.16)$$

where  $\tilde{x} = x - x_B$ ,  $x_B$  = coordinates of the barycenter of  $T$ . In (3.16)  $\theta_1$  and  $\theta_2$  are to be chosen in order to verify (3.15). By inserting (3.16) in (3.15) we obtain a system of two equations in the unknowns  $\theta_1$  and  $\theta_2$  which is always solvable. Also, we point out that

$$\theta_1^2 + \theta_2^2 = (\bar{f}_T)^2/4. \quad (3.17)$$

With this choice for  $\mathbf{J}^f$  in (3.3) one clearly obtains

$$\begin{aligned} \rho_h &= \tilde{\rho}_h, \\ \mathbf{J}_h &= \tilde{\mathbf{J}}_h - \frac{\int_T e^{\psi} \mathbf{J}^f dx}{\int_T e^{\tilde{\psi}} dx} = \tilde{\mathbf{J}}_h - \frac{\int_T \mathbf{J}^f dx}{|T|}. \end{aligned} \quad (3.18)$$

In view of these considerations, we therefore suggest the following approach: to use in the computations the simplified form (3.12) for the right-hand side, instead of the original one (3.9). Then, if one is interested in having an approximation of  $\mathbf{J}$  with reasonable continuity properties (of the type (3.6)), use (3.18) as a post-processing. We explicitly point out that the computational cost of (3.18) matches closely the difference in computational cost between (3.9) and (3.12). The advantage here is that, in many applications, one would use (3.18) only in a few triangles. Besides, when using an iterative procedure, one might choose to post-process only the last iteration and not the intermediate ones.

**REMARK 3.2.** For hybrid methods over quadrilateral decompositions of  $\Omega$  in this context, we refer to [1].

We conclude this section with some error estimates on the current density  $\mathbf{J}$ . First, let us note that from (3.16), (3.17) and (3.18) one easily obtains

$$\|\mathbf{J}_h - \tilde{\mathbf{J}}_h\|_{L^2(T)} \leq ch \tilde{f}_T \quad \forall T \in \mathcal{T}_h. \quad (3.19)$$

Hence, we shall prove estimates for  $\mathbf{J} - \tilde{\mathbf{J}}_h$  only. For that, we assume  $f$  to be piecewise constant and, in order to simplify the notation, we introduce the bilinear form:

$$a(\mathbf{J}, \boldsymbol{\tau}) = \int_{\Omega} e^{\tilde{\psi}} \mathbf{J} \cdot \boldsymbol{\tau} dx, \quad (3.20)$$

and the associated norm

$$\|\boldsymbol{\tau}\|_a^2 = a(\boldsymbol{\tau}, \boldsymbol{\tau}). \quad (3.21)$$

With this notation, (3.11) can be written as

$$\begin{aligned} &\text{Find } (\tilde{\mathbf{J}}_h, \rho_h) \in V_h \times W_{h,\chi}, \quad \text{such that} \\ &a(\tilde{\mathbf{J}}_h, \boldsymbol{\tau}) - \int_{\Omega} \boldsymbol{\tau} \cdot \nabla \rho_h dx = 0 \quad \forall \boldsymbol{\tau} \in V_h, \\ &\int_{\Omega} \tilde{\mathbf{J}}_h \cdot \nabla \varphi dx = - \int_{\Omega} f \varphi dx \quad \forall \varphi \in W_{h,0}. \end{aligned} \quad (3.22)$$

Note that the first equation of (3.22) is nothing but (3.13). Then, we have the following theorem.

**THEOREM 3.1.** Let  $u$  be the solution of (1.2) and let  $\mathbf{J} := e^{-\psi} \nabla(e^{\psi} u) = \nabla u + u \nabla \psi$ . Let  $(\tilde{\mathbf{J}}_h, \rho_h)$  be the solution of (3.22) and let  $P_h \mathbf{J}$  and  $\tilde{\mathbf{J}}$  be the approximations of  $\mathbf{J}$  defined by

$$P_h \mathbf{J} \in V_h \quad \text{and} \quad (\mathbf{J} - P_h \mathbf{J}, \boldsymbol{\tau}) = 0 \quad \forall \boldsymbol{\tau} \in V_h, \quad (3.23)$$

$$\tilde{\mathbf{J}} \in V_h \quad \text{and} \quad \tilde{\mathbf{J}} = e^{-\tilde{\psi}} \nabla(e^{\tilde{\psi}} u)^I (= e^{-\tilde{\psi}} \nabla \rho^I), \quad (3.24)$$

where  $z^I \in W_{h,\chi}$  denotes the piecewise linear interpolant of  $z$ . Then

$$\|\mathbf{J} - \tilde{\mathbf{J}}_h\|_a \leq \frac{3}{2} \{ \|\mathbf{J} - P_h \mathbf{J}\|_a + \|\mathbf{J} - \tilde{\mathbf{J}}\|_a \}. \quad (3.25)$$

**PROOF.** We have

$$\begin{aligned} \|\mathbf{J} - \tilde{\mathbf{J}}_h\|_a^2 &= a(\mathbf{J} - \tilde{\mathbf{J}}_h, \mathbf{J} - \tilde{\mathbf{J}}_h) \\ &= a(\mathbf{J} - \tilde{\mathbf{J}}_h, \mathbf{J} - P_h \mathbf{J}) + a(\mathbf{J} - \tilde{\mathbf{J}}_h, P_h \mathbf{J} - \tilde{\mathbf{J}}_h) \\ &= \text{I} + \text{II} \end{aligned} \quad (3.26)$$

$$\begin{aligned} \text{II} &= a(\mathbf{J} - \tilde{\mathbf{J}}_h, P_h \mathbf{J} - \tilde{\mathbf{J}}_h) \\ &= a(\mathbf{J} - \tilde{\mathbf{J}}, P_h \mathbf{J} - \tilde{\mathbf{J}}_h) + a(\tilde{\mathbf{J}} - \tilde{\mathbf{J}}_h, P_h \mathbf{J} - \tilde{\mathbf{J}}_h) \\ &= a(\mathbf{J} - \tilde{\mathbf{J}}, P_h \mathbf{J} - \tilde{\mathbf{J}}_h). \end{aligned} \quad (3.27)$$

In fact

$$\begin{aligned} a(\tilde{\mathbf{J}} - \tilde{\mathbf{J}}_h, P_h \mathbf{J} - \tilde{\mathbf{J}}_h) &= \int_{\Omega} e^{\tilde{\psi}} (e^{-\tilde{\psi}} \nabla \rho^I - e^{-\tilde{\psi}} \nabla \rho_h) (P_h \mathbf{J} - \tilde{\mathbf{J}}_h) \, dx \\ &= \int_{\Omega} (\nabla \rho^I - \nabla \rho_h) (P_h \mathbf{J} - \tilde{\mathbf{J}}_h) \, dx \\ &= \int_{\Omega} (\nabla \rho^I - \nabla \rho_h) (\mathbf{J} - \tilde{\mathbf{J}}_h) \, dx \\ &= - \int_{\Omega} f(\rho^I - \rho_h) \, dx - \int_{\Omega} (\nabla \rho^I - \nabla \rho_h) \tilde{\mathbf{J}}_h \, dx = 0. \end{aligned} \quad (3.28)$$

In (3.28) we used the definition (3.23) of  $P_h \mathbf{J}$ , the fact that  $\text{div } \mathbf{J} = f$  and  $\rho^I - \rho_h \in W_{h,0}$ . From (3.26) and (3.27) we have then, via the Cauchy–Schwartz inequality

$$\begin{aligned} \|\mathbf{J} - \tilde{\mathbf{J}}_h\|_a^2 &\leq \|\mathbf{J} - \tilde{\mathbf{J}}_h\|_a \|\mathbf{J} - P_h \mathbf{J}\|_a + \|\mathbf{J} - \tilde{\mathbf{J}}\|_a \|P_h \mathbf{J} - \tilde{\mathbf{J}}_h\|_a \\ &\leq \|\mathbf{J} - \tilde{\mathbf{J}}_h\|_a (\|\mathbf{J} - P_h \mathbf{J}\|_a + \|\mathbf{J} - \tilde{\mathbf{J}}\|_a) + \|\mathbf{J} - \tilde{\mathbf{J}}\|_a \|\mathbf{J} - P_h \mathbf{J}\|_a \end{aligned} \quad (3.29)$$

from which (3.25) follows easily.  $\square$

#### 4. Numerical aspects

As already pointed out, in the applications to semiconductor device simulation,  $\nabla\psi$  can be quite large in a part of the domain. This is clearly a potential source of numerical difficulties. More generally, the presence of a big  $\nabla\psi$  can change the qualitative behaviour of the scheme. In the present section we shall deal with some numerical aspects related to this problem. To this aim, it will be more convenient to set

$$\psi = \frac{\psi_0}{l} \quad (4.1)$$

and assume that  $\nabla\psi_0$  is smooth everywhere and  $l$  is small. Accordingly, (1.2) becomes

$$\operatorname{div}\left(\nabla u + \frac{\nabla\psi_0}{l} u\right) = f. \quad (4.2)$$

In what follows, we shall first justify, in (a), the need of a numerical integration already pointed out for the hybrid scheme of Section 3. Then, in (b) and (c) we shall assume  $\nabla\psi_0$  to be constant and show the qualitative behaviour of the mixed and hybrid scheme, respectively, for  $l \ll h$ . Finally, in (d) and (e) we shall briefly discuss the numerical difficulties that can arise in the two schemes when  $\nabla\psi_0$  changes abruptly from one triangle to another.

(a) Let us first consider the problem of the choice of the integration formula (3.4). We recall that the contributions of each triangle  $T$  to the stiffness matrix, acting on the final unknown  $u_h$ , and to the right-hand side are respectively

$$a_{ij}^T = \frac{\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)}}{4I_T(e^\psi)} e^{\psi_j} \quad (4.3)$$

and

$$b_i^T = -\frac{1}{3} |T| \bar{f}_T. \quad (4.4)$$

A quick glance at (4.2) suggests that, as  $l \rightarrow 0$ , the behaviour of  $a_{ij}^T$  should be of the order of magnitude of  $l^{-1}$ . Assume now that  $\psi_0$  is linear in  $T$  and assume that  $l \ll |\nabla\psi_0| h_T$ . If you neglect the very special cases in which  $\psi$  reaches its maximum  $\psi^M$  on one edge, you are left with the generic case in which  $\psi$  reaches its maximum only at one vertex. A simple computation shows that, in this case,

$$\int_T e^\psi dx = \int_T e^{\psi_0/l} dx \simeq |T| l^2 e^{\psi_0^M/l} = |T| l^2 e^{\psi^M}. \quad (4.5)$$

As a consequence, the use of exact integration in (4.3) would give

$$a_{ij}^T \simeq \begin{cases} \mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)} / 4l^2 |T|, & \text{if } \psi_j = \psi^M, \\ 0, & \text{otherwise,} \end{cases} \quad (4.6)$$

which is not very satisfactory. Hence we need an integration formula such that

$$I_T(e^{\psi_0/l}) \simeq |T| l e^{\psi_0^{M/l}}. \quad (4.7)$$

For that, we made the following choice: let  $V_m$  and  $V_M$  be (any) two vertices where  $\psi$  reaches its minimum and (respectively) its maximum (you can have a non-unique choice for them, but this does not matter). Let, accordingly,  $e$  be the edge of  $T$  joining  $V_M$  to  $V_m$ . Then, we set

$$I_T(e^\psi) = \frac{|T|}{|e|} \int_e e^\psi dx. \quad (4.8)$$

It is easy to check that choice (4.8) fulfills requirement (4.7). On the other hand, in the triangles  $T$  where  $|\nabla\psi|h$  is not too big, formula (4.8) gives a reasonably good approximation of the exact integral over  $T$ , so that we suggest to use it for every triangle  $T$  in  $T_h$ , as we did in our computations.

In order to convince the reader that the use of exact integration in (4.3) leads (for  $f \neq 0$ ) to the wrong result, we refer to the numerical experiments of the next section.

(b) We now want to analyse the qualitative behaviour of the mixed scheme (2.20) (in the unknown  $u_h$ ) as  $l \rightarrow 0$ . For this, we consider as a model case the situation of Fig. 2, where  $P$  is the midpoint of the common edge of the two equilateral triangles  $T_1$  and  $T_2$ . It is clear that the row of the matrix associated with the node  $P$  will have only five coefficients that (a priori) are different from zero. We denote them by  $a_{P,P}$  and  $a_{P,P_k}$  ( $k = 1, 2, 3, 4$ ) with the notation of Fig. 3. For every  $\theta$  in  $[0, 2\pi[$  we consider a constant value of  $\nabla\psi_0 \equiv (\cos \theta, \sin \theta)$  on  $T_1 \cup T_2$ , and we compute the coefficients of the row associated with  $P$  for  $l \ll h$ . In Fig. 4 we can see the behaviour (as functions of  $\theta$ ) of the coefficients  $a_{P,P}$  and  $a_{P,P_1}$  for  $h = 1$  and  $l = 10^{-4}$ . The other three coefficients can be obtained from  $a_{P,P_1}$  by the formulae

$$a_{P,P_1}(\theta) = a_{P,P_2}\left(\theta + \frac{\pi}{3}\right) = a_{P,P_3}(\theta + \pi) = a_{P,P_4}\left(\theta + \frac{4}{3}\pi\right). \quad (4.9)$$

Varying  $h$ , the corresponding coefficients vary linearly with  $h$ , as far as  $l \ll h$ . We can see that the scheme has always at least one coefficient equal to zero, corresponding to the downwind node (wind =  $-\nabla\psi_0$ ). Figure 3 helps in understanding which coefficient is zero and which one is different from zero for the various values of  $\theta$ . In this figure, we split  $T_1 \cup T_2$  into four dashed regions, plus four undashed ones. We label the undashed ones by  $R_1, \dots, R_4$ .

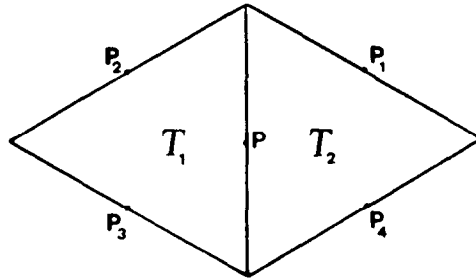


Fig. 2.

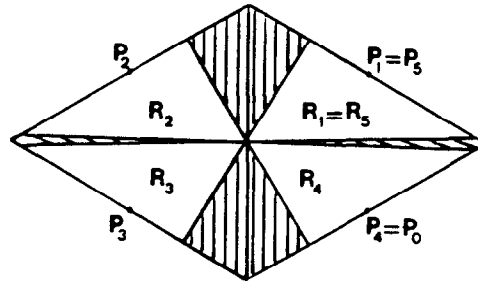


Fig. 3.

Now, if  $\nabla\psi_0$  points into a region  $R_i$ , only the nodes  $P_{i-1}$ ,  $P_i$ ,  $P_{i+1}$  will be active (i.e. the corresponding coefficient will be different from 0). If instead,  $\nabla\psi_0$  points into a dashed region which is between  $R_i$  and  $R_{i+1}$ , then only the nodes  $P_i$  and  $P_{i+1}$  will be active. Hence the scheme produces a sort of upwinding effect. It can be seen that the matrix produced by the scheme corresponds roughly to the use of an artificial viscosity of the order of  $h/4$ .

(c) An analogous upwind behaviour is exhibited by the hybrid scheme of Section 3. To see that more clearly, let us consider the simplified situation of Fig. 5, where six equilateral triangles join at a node  $P$ . The row of the matrix associated with  $P$  will have seven coefficients (a priori) different from zero:  $a_{P,P}$ ,  $a_{P,P_k}$  ( $k = 1, \dots, 6$ ). As before for every  $\theta \in [0, 2\pi[$  we consider a constant value of  $\nabla\psi_0 \equiv (\cos \theta, \sin \theta)$  and we compute the coefficients of the row corresponding to  $P$  for  $l \ll h$ . Figure 7 shows the behaviour of the coefficients  $a_{P,P}$  and  $a_{P,P_1}$  as functions of  $\theta$  for  $h = 1$  and  $l = 10^{-4}$ . The other coefficients are obtainable from  $a_{P,P_1}$  by the formula

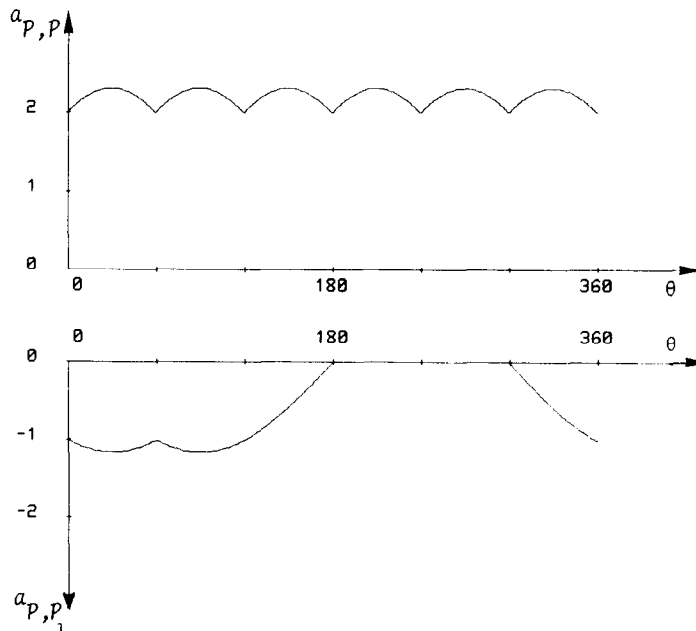


Fig. 4.

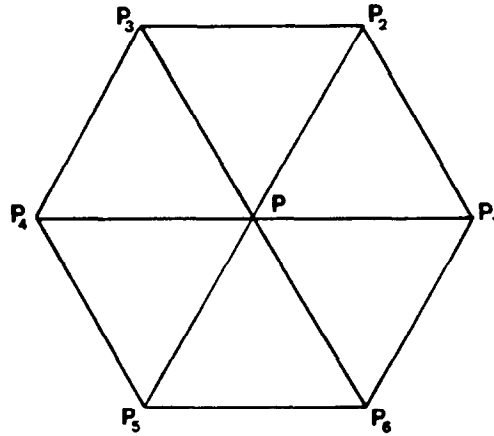


Fig. 5.

$$a_{P,P_1}(\theta) = a_{P,P_k} \left( \theta + (k-1) \frac{\pi}{3} \right), \quad k = 2, \dots, 6. \quad (4.10)$$

We can see that the scheme has always at least three coefficients equal to zero, corresponding to the three downwind nodes. Figure 6 helps in understanding how the upwind nodes change with  $\nabla\psi_0$ . If  $\nabla\psi_0$  points into a region  $R_i$ , then the upwind nodes are  $P_{i-1}$ ,  $P_i$ ,  $P_{i+1}$ . If instead  $\nabla\psi_0$  lies along  $R_i \cap R_{i+1}$ , then the upwind nodes are  $P_i$  and  $P_{i+1}$ . In this respect, the scheme exhibits the same upwind nodes of other upwind schemes [27], although with a different conception and different numbers. As such, it can be considered to have an artificial viscosity roughly of the order of  $h/3$ .

(d) We assumed so far that  $\nabla\psi$  does not change from one triangle to another. As we have seen, with this assumption both schemes produce, even for very small  $l$ , a quite reasonable matrix that corresponds roughly to the use of some kind of artificial viscosity. We shall see now that the situation is less comfortable if we admit the possibility of a  $\nabla\psi$  which changes abruptly from one triangle to the neighbouring one. In order to see what can happen it is

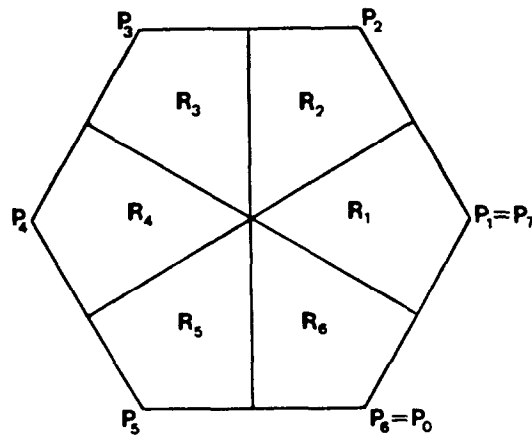


Fig. 6.

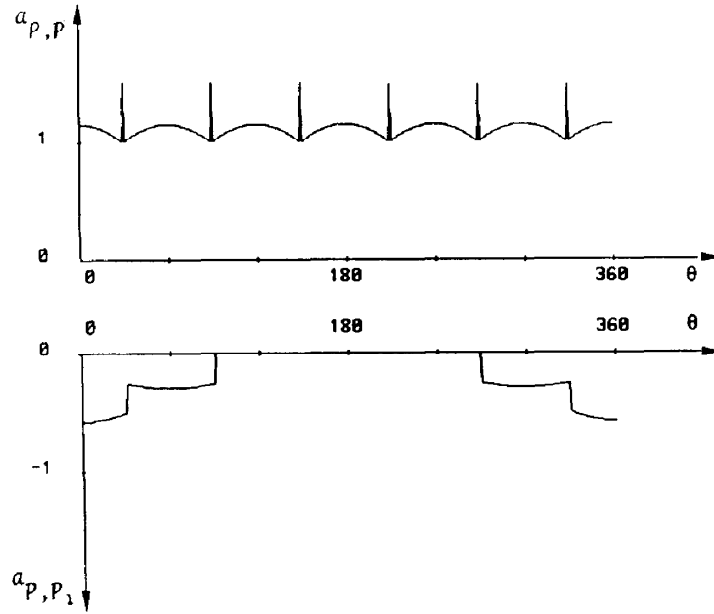


Fig. 7.

convenient to have one more look at the contributions of each triangle to the final matrix in the unknown  $u_h$ . Let us consider the mixed scheme first. Remember that each of the final degrees of freedom is associated with one edge of  $T_h$  (or, if you prefer, with the midpoint of the edge). The contributions of  $T$  to the coefficients  $a_{ij}$  of the final matrix are

$$a_{ij}^T = \frac{\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)}}{\int_T \mathbf{e}^\psi \, dx} (\mathbf{e}^\psi)'|_{e_j}, \quad (4.11)$$

where we recall that here

$$(\mathbf{e}^\psi)'|_{e_j} = \frac{1}{|e_j|} \int_{e_j} \mathbf{e}^\psi \, ds. \quad (4.12)$$

A simple computation shows that, if  $\psi^{M_j}$  is the maximum value of  $\psi = \psi_0/l$  over  $e_j$  (and if  $\psi$  is not constant on  $e_j$ ), then

$$(\mathbf{e}^\psi)'|_{e_j} = l \mathbf{e}^{\psi^{M_j}}. \quad (4.13)$$

It is now clear from (4.11)–(4.13) and (4.5) that

$$a_{ij}^T \approx \begin{cases} \mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)} / l |T|, & \text{if } \psi^{M_j} = \psi^M, \\ 0, & \text{otherwise,} \end{cases} \quad (4.14)$$



where  $\psi^M$  denotes the maximum value of  $\psi$  in  $T$  as in (4.5). Hence, the scheme annihilates the contributions  $a_{ij}^T$  corresponding to the (nodes on) the edges  $e_j$  where  $\psi$  does not reach its maximum. In particular, when  $\nabla\psi$  changes from one triangle to another, the annihilation effect can lead to a column (corresponding to an edge  $e_j$  of  $T_h$ ) of very small coefficients. This happens if  $\psi^{M_i} \ll \psi^{M_k}$  for all edges  $e_k$  ( $k \neq j$ ) sharing a same triangle with  $e_j$ . On the other hand, a row of very small coefficients can never occur. Actually, it is easy to see that equality  $\psi^{M_j} = \psi^M$  occurs, in each triangle, for at least two different  $j$ s. This implies, using (4.14), that  $a_{ij}^T$  will be non-negligible for at least one value of  $j$ . (Clearly,  $\nu^{(i)} \cdot \nu^{(j)}$  cannot vanish for two different values of  $j$ ).

(e) Analogous considerations hold for the hybrid scheme of Section 3, if  $\nabla\psi$  changes from one triangle to another. Looking at the contributions (4.3) of a triangle  $T$  to the stiffness matrix in the unknown  $u_h$ , and using the integration formula (4.8), it is easy to see that

$$a_{ij}^T = \begin{cases} \nu^{(i)} \cdot \nu^{(j)} / 4l |T|, & \text{if } \psi_j = \psi^M, \\ 0, & \text{otherwise.} \end{cases} \quad (4.15)$$

It is clear from (4.15) that the scheme annihilates the contributions  $a_{ij}^T$  corresponding to the nodes  $j$  where  $\psi_j \ll \psi^M$ . This can lead to a final matrix having either a row or a column of very small coefficients. In particular, assume that a node  $P$  is such that  $\psi(P) \ll \max_{x \in T} \psi(x)$  for every triangle  $T$  having a vertex in  $P$  (i.e.  $\psi$  has a very deep minimum in  $P$ , compared with the local mesh size). Then, the scheme will produce a column of very small coefficients. In order to see what can happen to a row, we first remark that, in general, for a given triangle  $T$  equality  $\psi_j = \psi^M$  will hold for one value of  $j$  only. If the decomposition is of weakly acute type, the corresponding contribution  $a_{ij}^T$  can be zero if  $\nu^{(i)} \cdot \nu^{(j)} = 0$ . Suppose now that you have a node  $P_i$  in the following unlucky situation: for every triangle  $T$  having  $P_i$  as a vertex  $|\nabla\psi|$  is big and the maximum of  $\psi$  over  $T$  is reached at a point  $P_j$  such that  $\nu^{(i)} \cdot \nu^{(j)} = 0$ . Then, the  $i$ th row will be made of very small coefficients.

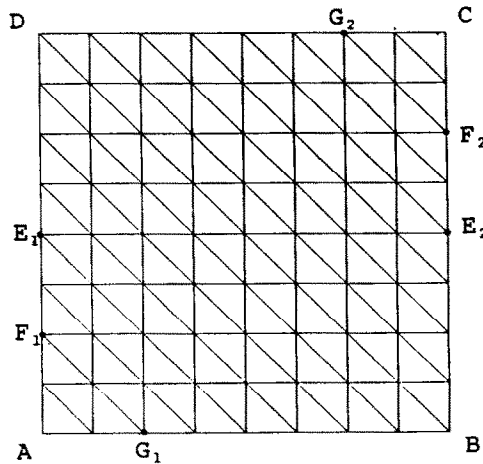


Fig. 8.

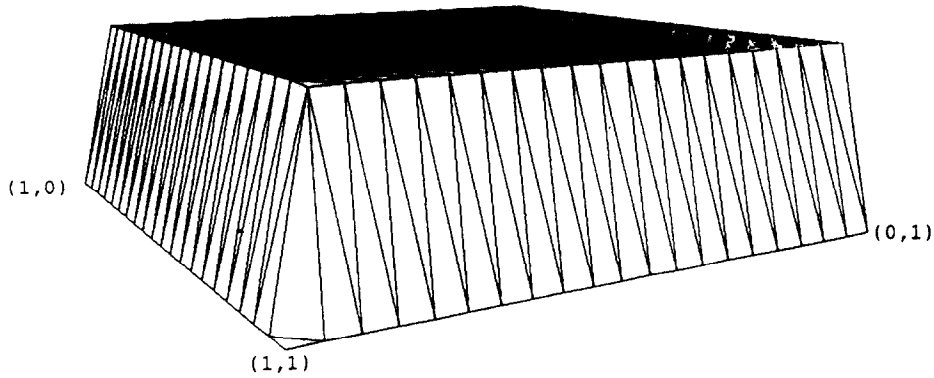


Fig. 9.

## 5. Numerical results

We present here some numerical results obtained using the methods described in the previous sections over uniform triangular decompositions of the unit square (see Fig. 8). The function  $\psi(x)$  is assumed to be piecewise linear and given by the expression  $\psi(x) = \psi_0(x)/l$ , with  $l = 10^{-4}$  in all the examples.

In the first example we show the resolution of a boundary layer (see Figs. 9, 10). The function  $\psi_0(x)$  has the constant gradient  $\nabla\psi_0 = -(2, 1)$ , and  $f = 0$ . The boundary conditions for  $u$  are:  $u = 1$  on the edges  $AB$  and  $AD$ ,  $u = 0$  elsewhere. In Fig. 9 the solution has been computed with the hybrid method (3.3) on a  $20 \times 20$  mesh, and in Fig. 10 with the mixed method (2.20) on a  $10 \times 10$  mesh. The boundary layer is caught in one element and no smoothing effect is introduced. Figure 11 shows how an internal layer is transported:  $\psi_0$  is chosen as in the previous case, and  $f = 0$ . The boundary conditions for  $u$  are:  $u = 0$  on  $AB$ ,  $AF_1$  and  $BF_2$ ,  $u = 1$  elsewhere. The solution is computed with the hybrid method on a  $40 \times 40$  mesh. (For the mixed method the situation is the same). In this case, a smoothing effect in the crosswind direction appears, as pointed out in Section 4. The example given in Fig. 12 refers

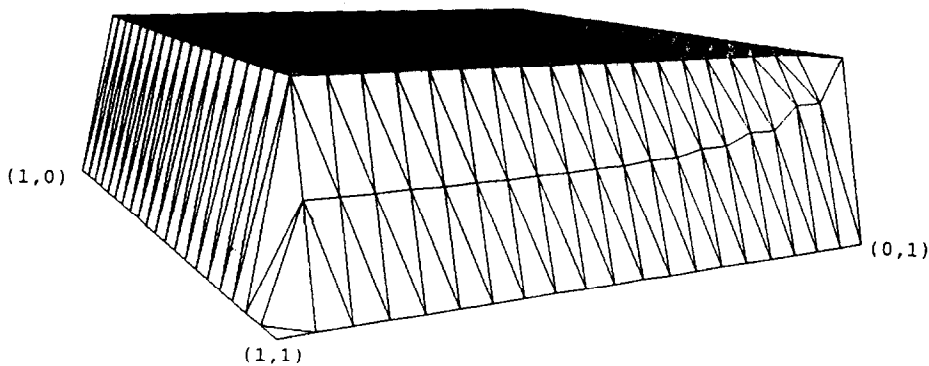


Fig. 10.

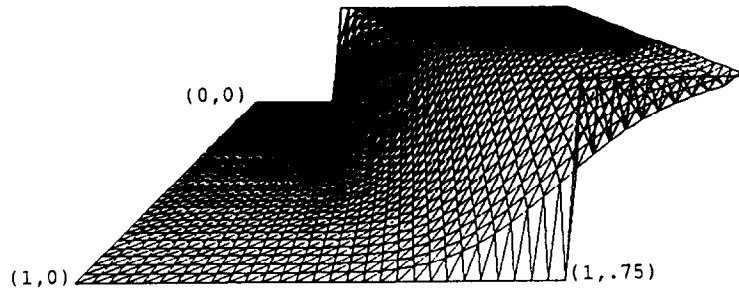


Fig. 11.

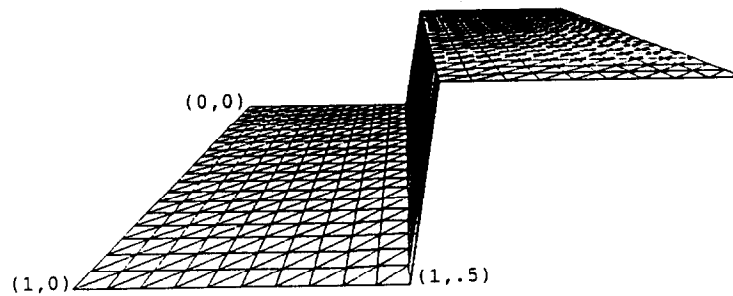


Fig. 12.

to a favourable situation where the mesh is well oriented with respect to the transport. The function  $\psi_0$  has the constant gradient  $\nabla\psi_0 = -(1, 0)$ ,  $f = 0$  and the boundary conditions for  $u$  are:  $u = 0$  on  $AB$ ,  $AE_1$  and  $BE_2$ ,  $u = 1$  elsewhere. Figure 12 shows the mixed solution on a  $10 \times 10$  mesh. In this case the internal layer is as sharp as it can be. An analogous behaviour is exhibited by the hybrid scheme, if exact integration is used in (3.4). On the contrary, the use of a numerical integration as in (4.8) produces, even in this favourable case, a smoothing effect as that of Fig. 11.

Figures 13 and 14 show the results obtained using the hybrid scheme when the source term  $f$  is different from zero. The function  $\psi_0$  has the constant gradient  $\nabla\psi_0 = (-2, 0)$  and  $f = 2$ . The

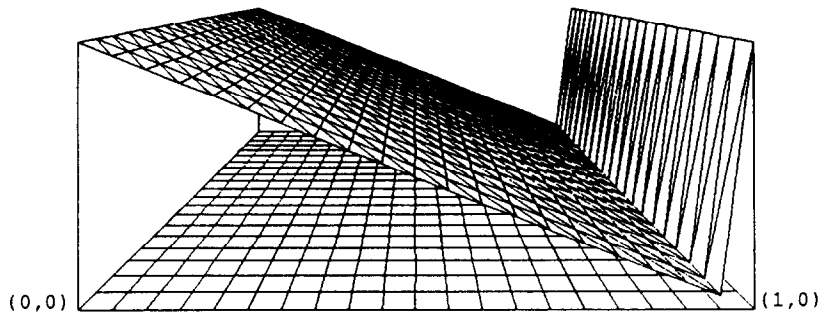


Fig. 13.

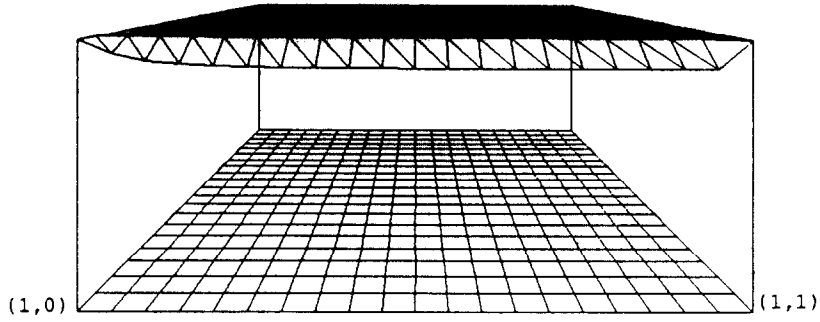


Fig. 14.

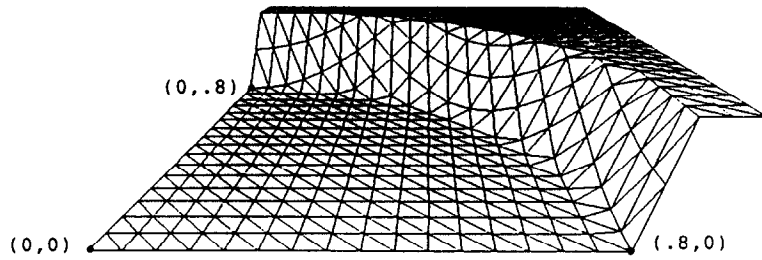


Fig. 15.

boundary conditions for  $u$  are:  $u = 1$  on  $AD$  and  $BC$ , and homogeneous Neumann boundary conditions elsewhere. In the first case, the quadrature rule (4.8) is used, while in the second one the exact integrals are computed. As already pointed out in Section 4, the second choice leads to a discrete solution which is not an approximation of the solution of the continuous problem.

In the last example,  $\psi_0$  is chosen with the expression

$$\psi_0(x) = \begin{cases} 0, & \text{if } 0 \leq \rho \leq 0.8, \\ 2(\rho - 0.8), & \text{if } 0.8 \leq \rho \leq 0.9, \\ 0.2, & \text{if } 0.9 \leq \rho, \end{cases}$$

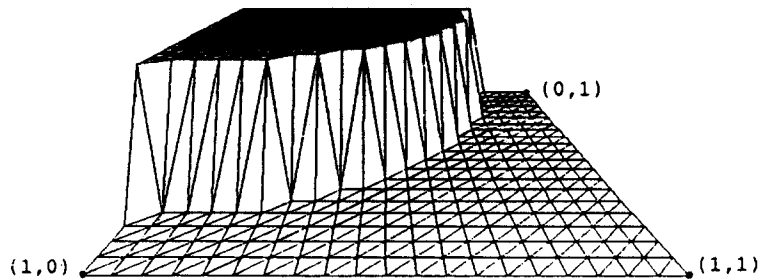


Fig. 16.

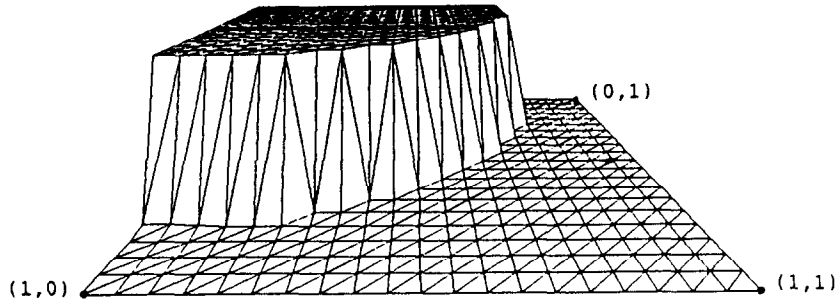


Fig. 17.

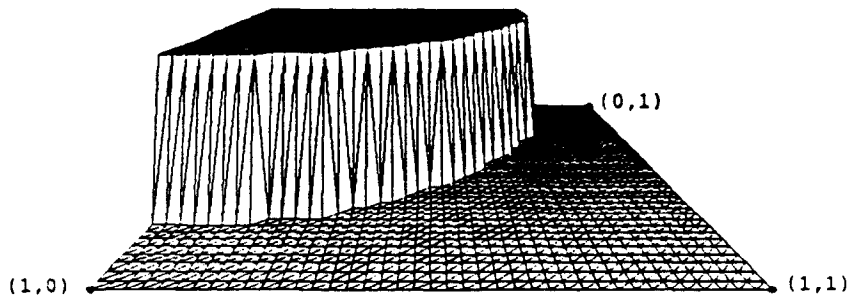


Fig. 18.

and  $\rho = (x_1^2 + x_2^2)^{1/2}$  (see Fig. 15). The boundary conditions for  $u$  are:  $u = 1$  on  $AF_1$  and on  $AG_1$ ,  $u = 0$  on  $F_2C$  and  $G_2C$ . Homogeneous Neumann boundary conditions are assumed elsewhere. In Figs. 16–18 we show  $u_h$  as computed, respectively, with hybrid methods (3.3) on a  $20 \times 20$  mesh, mixed methods (2.20) on a  $10 \times 10$  and  $20 \times 20$  mesh.

**REMARK 5.1.** For graphical purposes, the solutions computed with the mixed scheme on the  $N \times N$  mesh have been interpolated at the vertices and presented on a  $2N \times 2N$  grid.

## References

- [1] F. Brezzi, L.D. Marini and P. Pietra, Two-dimensional exponential fitting and applications to drift-diffusion models, *SIAM J. Numer. Anal.* 26 (1989).
- [2] W.V. Van Roosbroeck, Theory of flow of electrons and holes in germanium and other semiconductors, *Bell Syst. Tech. J.* 29 (1950) 560–607.
- [3] P.A. Markowich, *The Stationary Semiconductor Device Equations* (Springer, Berlin, 1986).
- [4] M.S. Mock, *Analysis of Mathematical Models of Semiconductor Devices* (Boole Press, Dublin, 1983).
- [5] H.U. Baranger, Ballistic electrons in a submicron semiconducting structure: A Boltzmann equation approach, Ph.D. Thesis, Cornell University, 1986.
- [6] W. Fawcett, A.D. Boardman and S. Swain, Monte Carlo determination of electron transport properties in GaAs, *J. Phys. Chem. Solids* 31 (1970).
- [7] B. Niclot, P. Degond and F. Poupaud, Deterministic particle simulation of the Boltzmann transport equations of semiconductors, *J. Comp. Phys.* 78 (1988) 313–349.

- [8] F. Brezzi, A.C. Capelo and L.D. Marini, Singular perturbation problems in semiconductor devices, in: *Numerical Analysis, Lecture Notes in Mathematics*, Vol. 1230 (Springer, Berlin, 1986) 191–198.
- [9] F. Brezzi, A.C. Capelo and L. Gastaldi, A singular perturbation analysis for reverse-biased diodes, to appear in *SIAM J. Math. Anal.* 19 (5) (1989).
- [10] J.W. Jerome, Consistency of semiconductor modeling: an existence/stability analysis for the stationary Van Roosbroeck System, *SIAM J. Appl. Math.* 45 (1985) 565–590.
- [11] H.K. Gummel, A self-consistent iterative scheme for one dimensional steady state transistor calculations, *IEEE Trans. Electron. Devices* 11 (1964) 455–465.
- [12] S.J. Polak, C. Den Heijer, W.H.A. Schilders and P.A. Markowich, Semiconductor device modelling from the numerical point of view, *Internat. J. Numer. Methods Engrg.* 24 (1987) 763–838.
- [13] D. Scharfetter and H. Gummel, Large-signal analysis of a silicon Read diode oscillator, *IEEE Trans. Electron. Devices* 16 (1969) 64–77.
- [14] K.W. Morton, Generalised Galerkin methods for steady and unsteady problems, in: K.W. Morton and M.J. Baines, eds., *Numerical Methods for Fluid Dynamics* (Academic Press, New York, 1982) 1–32.
- [15] D. Allen and R. Southwell, Relaxation methods applied to determining the motion, in two dimensions, of a viscous fluid past a fixed cylinder, *Q. J. Mech. Appl. Math.* 8 (1955) 129–145.
- [16] M.S. Mock, Analysis of a discretisation algorithm for stationary continuity equations in semiconductor device models II, *COMPEL* 3 (1984) 137–149.
- [17] R.E. Bank, D.J. Rose and W. Fichtner, Numerical Methods for Semiconductor Device Simulation, *IEEE Trans. Electron. Devices* 30 (9) (1983) 1031–1041.
- [18] P.A. Markowich and M. Zlámal, Inverse-average-type finite element discretisations of self-adjoint second order elliptic problems, *Math. Comp.* 51 (1988) 431–449.
- [19] J. Douglas, jr., I. Martinez Gamba and M.C.J. Squeff, Simulation of the transient behavior of a one-dimensional semiconductor device, *Mat. Apl. Comput.* 5 (1986) 103–122.
- [20] F. Brezzi, L.D. Marini and P. Pietra, Méthodes d'éléments finis mixtes et schéma de Scharfetter–Gummel, *C.R. Acad. Sci. Paris Série I* 305 (1987) 599–604.
- [21] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems* (North-Holland, Amsterdam, 1978).
- [22] P.A. Raviart and J.M. Thomas, A mixed finite element method for second order elliptic problems, in: *Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics*, Vol. 606 (Springer, Berlin, 1977) 292–315.
- [23] F. Brezzi, On the existence uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers, *RAIRO 8-R2* (1974) 129–151.
- [24] B.X. Fraeijs de Veubeke, Displacement and equilibrium models in the finite element method, in: O.C. Zienkiewicz and G. Hollister, eds., *Stress Analysis* (Wiley, New York, 1965).
- [25] D.N. Arnold and F. Brezzi, Mixed and non-conforming finite element methods: implementation, post-processing and error estimates, *M<sup>2</sup>AN* 19 (1985) 7–32.
- [26] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, submitted to Springer.
- [27] K. Baba and M. Tabata, On a conservative upwind finite element scheme for convective diffusion equations, *RAIRO Anal. Numer.* 15 (1981) 3–25.