UDK 621.3:(53+54+621+66), ISSN0352-9045 Informacije MIDEM 23(1993)3, Ljubljana

ALGORITHMS FOR MULTI-DIMENSIONAL ANALYSIS OF SEMICONDUCTORS DERIVED FROM FIRST PRINCIPLES

Zvonko Fazarinc

KEYWORDS: semiconductors, semiconductor structure, computer analysis, algorithms, C-language, charge carriers, transport equation, electron distribution, hole distribution, electric field, basic principles, multi-dimensional

ABSTRACT: Expressions for multi-dimensional analysis of semiconductor structures in the discrete domain are derived from first principles. A simple structure is analyzed and the algorithms are cast in C-language.

Algoritmi za večdimenzionalno računalniško analizo polprevodnikov iz osnovnih fizikalnih principov

KLJUČNE BESEDE: polprevodniki, strukture polprevodniške, računalniška analiza, algoritmi, C jezik, nosilci nabojev, enačba transportna, porazdelitev elektronov, porazdelitev vrzeli, polje električno, principi osnovni, večdimenzionalnost

POVZETEK: Izrazi za večdimenzionalno računalniško analizo polprevodniških struktur so izvedeni iz osnovnih fizikalnih zakonov. Uporaba je ilustrirana na enostavnem primeru in algoritmi so prikazani v C-jeziku.

1. Introduction

Multi-dimensional analysis of semiconductor structures is commonly deferred to pre-canned computer programs [1] which often drape a veil of mystery over the inner workings of such design tools. The Poisson's equation is usually taken as the basis for evaluation of Fermi levels which then control the distribution of charged carriers. When the transport equation is used instead, the carrier distributions are computed from its discrete counterpart. This is prone to producing wrong answeres and makes the imposition of boundary conditions quite difficult.

The purpose of this paper is to derive the relevant equations for multi-dimensional analysis of semiconductors in discrete form directly from first physical principles. Such approach effectively avoids the hazards of discretization of partial differential equations [2], makes the imposition of boundary conditions. intuitive and, most importantly, it provides the practicing engineer and the novice with an insight which enables them to independently access the analytical powers of computers. The limitations of a paper prevent us from developing anything resembling a complete source code. Nevertheless, we will address the crucial ideas and make them understandable so that they can be embellished with refinements when needed.

We will be making use of the classical physics principles which are applicable to semiconductor structures larger than a few tenths of microns. For smaller structures quantum mechanics must be invoked and the reader should take note of this.

2. The Method of Approach

Before we address the general case we introduce the methodology with a simple example. Fig. 1 illustrates three points in space separated by Δx. With each location we associate a particle count $C(x,t)$ at time t.

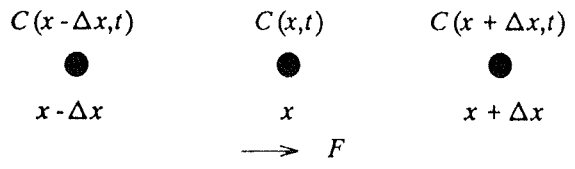

Fig. 1: Illustration of three points in one-dimensional space.

Particles are assumed to be in random thermal agitation which implies that they are equally likely to move to the left or to the right. We denote the likelyhood of their motion in one and the other direction by l. This means

that *l*-times the number of particles in a given position will move to the left and the same number will move to the right. We make a simplifying assumption that particles which do move make one single space step Δx in one time increment Δt. Armed with this information we can entertain the following question: "What will the concentration be in position x at time $t+\Delta t$ given the status at time *t*?" The answer proceeds along the following line of reasoning. If the likelyhood of motion is *l* then *l*-times the number of particles in position $x - \Delta x$ will move into position x during one time interval Δt. During the same time *l*-times the number of particles will move into position *x* form $x + \Delta x$. 2*l*-times the number of particles initially residing at *x* will have moved out of this position. What we have then left at *x* is

$$C(x,t+\Delta t) = C(x,t)+l\,C(x-\Delta x,t)$$

$$+l\,C(x+\Delta x,t) - 2l\,C(x,t) \qquad (1)$$

The mathematical manipulation below is intended to show that (1) is the discrete form of the diffusion equation with diffusivity *D* given by

$$D = l\,\Delta x^2/\Delta t \qquad (2)$$

First we subtract $C(x,t)$ on both sides of equation (1) and divide by Δt. Then we multiply and divide the RHS of the resulting equation by Δx^2 and obtain

$$\frac{C(x,t+\Delta t) - C(x,t)}{\Delta t} =$$

$$\frac{l\,\Delta x^2}{\Delta t}\,\frac{C(x-\Delta x,t)+C(x+\Delta x,t) - 2C(x,t)}{\Delta x^2}$$

We recognize the numerator of the LHS of the above equation as the temporal difference of $C(x,t)$ and the numerator of the RHS as the second spatial difference of $C(x,t)$. Upon taking the limit as Δx and Δt go to zero we end up with the familiar basic diffusion equation for which the relationship (2) applies.

$$\frac{\partial C(x,t)}{\partial t} = D\frac{\partial^2 C(x,t)}{\partial x^2}$$

A temporal sequence of plots produced by (1) when the initial distribution is a δ-function in the center and a unit step at the left is shown in Fig.2.

An implementation of our example in C-language is shown below

```
// Initialization:
 for (x = 0; x<200; x+ +) C[x] = 0;
 for (x=0; x<20; x+ +) C[x] = 1.0;
 C[100] = 10.0;

//Time loop of N passes:
  for (t = 0; t< N; t+ +)
   { // Space loop:
      L = 1.0;
```

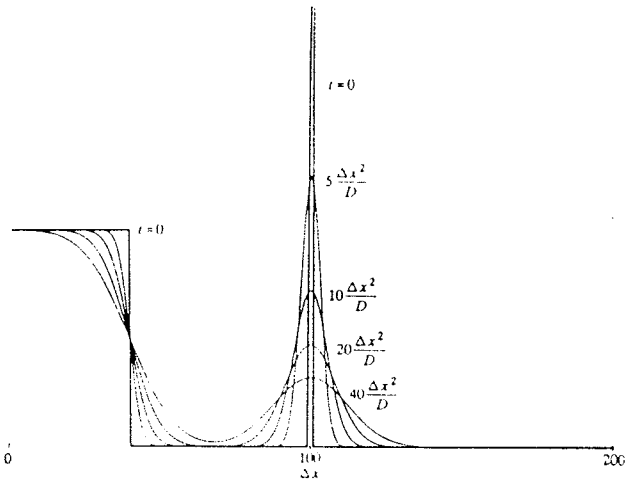

Fig. 2: Solution of discrete diffusion equation in one dimension

```
      for (x =  0; x< 200; x+ +)
{ temp =  0.25*C[x-1] + 0.25*C[x+1] + 0.5*C[x];
C[x-1] =  L; L =  temp;
        // Plot temp vs x
        }
  }
```

The timing required by (1) has been implemented in the above algorithm by reliance on memory management of computers. The values on the RHS of an equation are always taken to be the old values and those on the LHS of equation as the new ones. But we update C[x] by the new value only after we have computed C[x+1] which requires the old value of C[x]. This is accomplished by swapping the temp variable with L every time we pass through the space loop. It is obvious that the *l* value has been taken to be 0.25 in the example. One may wonder if such simple algorithm truly represents the solution of the diffusion equation. A quantitative comparison of (1) with the continuous diffusion equation is given in reference [2].

Next we allow a force *F*, indicated in Fig.1, to act uniformly on all particles. It is not difficult to conclude that a positive force - one that pushes to the right - will increase the likelyhood of particles moving to the right. At the same time it will decrease their chances of moving to the left. We can modify equation (1) for this case by introducing a skew factor *f* which is related to the force *F* and which biases the likelyhood *l* in the direction of force *F*.

$$C(x,t+\Delta t) = C(x,t) + (l+f)\,C(x-\Delta x,t)$$

$$+ (l-f)\,C(x+\Delta x,t) - 2l\,C(x,t) \qquad (3)$$

The effect of force on $C(x,t)$ cancels out in this simple example as the reader may verify. Expression (3) is the transport equation in one dimension with constant diffu-

sivity and force. To prove this we subtract $C(x,t)$ on both sides, divide by Δt and separate the l and f contributions

$$\frac{C(x,t+\Delta t) - C(x,t)}{\Delta t} = \frac{l}{\Delta t}[C(x-\Delta x,t) + C(x+\Delta x,t)$$

$$- 2C(x,t)] + \frac{f}{\Delta t}[C(x-\Delta x,t) - C(x+\Delta x,t)]$$

Now we need a relationship between the factor f and the force F. This is best done by comparing the energies involved during the move. The thermal energy kT is related to l in the same manner as the energy derived from the force field is related to f. A formal expression to that effect is

$$l : f = kT : \frac{F\Delta x}{2}$$

It yields the following for the force factor

$$f = l\frac{F\Delta x}{2kT} \qquad (4)$$

We substitute (4) into our last equation, multiply and divide the first term on the right by Δx^2, and the second term by $2\Delta x$ and end up with the following difference equation

$$\frac{C(x,t+\Delta t) - C(x,t)}{\Delta t} =$$

$$\frac{l\Delta x^2}{\Delta t}\frac{C(x-\Delta x,t) + C(x+\Delta x,t) - 2C(x,t)}{\Delta x^2}$$

$$- \frac{2l\Delta x^2}{\Delta t}\frac{F}{2kT}\frac{C(x+\Delta x,t) - C(x-\Delta x,t)}{2\Delta x}$$

The second term on the right is readily recognized as the first central difference in x of $C(x,t)$. It becomes the first derivative with respect to x when the limit is taken. The above equation then assumes the form

$$\frac{\partial C(x,t)}{\partial t} = D\frac{\partial^2 C(x,t)}{\partial x^2} - D\frac{F}{kT}\frac{\partial C(x,t)}{\partial x}$$

In the above we have substituted (2) for $l\Delta x^2/\Delta t$. Fig.3 shows a plot of expression (3) as function of x for a uniform force F with time being a parameter. Initial conditions are identical to those in Fig.2.

The numeric values for the plot of Fig.3 were generated by a source code identical to that shown earlier. The only difference is that the skew factor is introduced. We have chosen its value to be $f = 0.4$ $l = 0.1$. Consequently the only modification of the algorithm is in the space loop which now reads: *temp* = 0.35*C[x-1] + .15*C[x+1] + .5*C[x]; Everything else remains the same. The reader, familiar with problems arising in computer solutions of transport equation may find (3) to be of considerable interest. Its simplicity and the physical basis from which

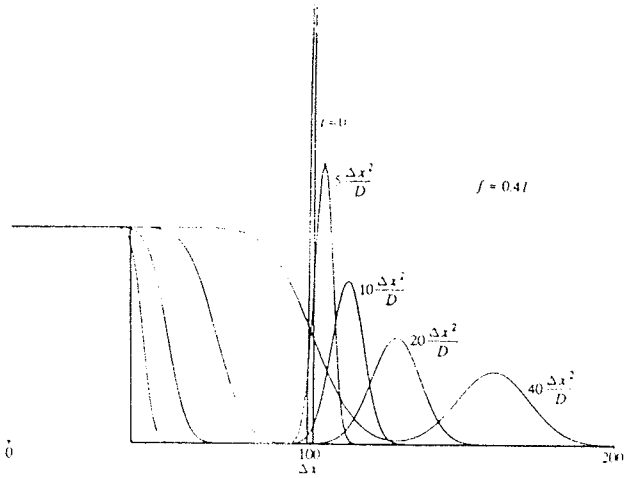

Fig. 3: Solution of transport equation for one-dimensional case.

(3) has been derived make the introduction of boundary conditions particularly easy.

We will now exploit the method just introduced for answering some of the more profound questions. One of them involves the diffusivity gradient. If D is a function of space, does it appear under the first or the second derivative sign? What if the diffusivity and force have different gradients along the spatial directions? How do we handle multidimensional analysis of semiconductors in general? These and some other questions will be addressed as we proceed.

3. Transport of Particles in Two Dimensions

The motion of charge carriers in semiconductors is governed by thermal energy, by electrical forces* and by properties of material through which they are moving. Their number depends on influx and outflow and on generation and recombination of oppositely charged pairs. We will derive the relevant equations taking into consideration the spatial variability of these effects and will allow, in addition, the temporal variation of electric fields and of carrier concentrations. In order to shield the derivation from excessive notational complexity we will limit it to the two-dimensional case. The extension to three dimensions will become self evident as we proceed.

We start with the illustration in Fig.4 which shows two points in x-space separated by Δx and a third point displaced in y-space by Δy. We assign the likelyhood of thermal motion in x-direction by l_x and that in y-direction

* Gravitational forces are negligible in comparison

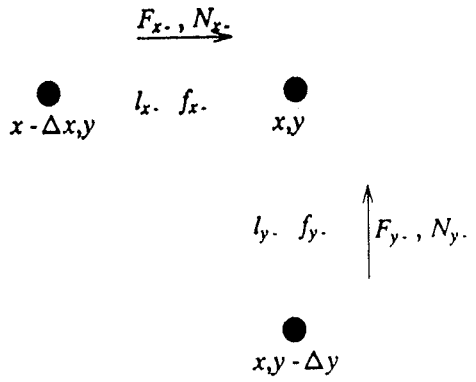

Fig. 4: Illustration of three points in two-dimensional space.

by l_y. We also allow a force F_x to act in the x-direction and a force F_y in the *y*-direction. These forces are responsible for the respective skew factors f_x and f_y through the relationship (4).

The motion likelyhood *l* and its skew *f* are assigned to the space between discrete spatial locations rather than to the locations themselves. The reason for this can be understood physically if we consider the case where the point $(x-\Delta x,y)$ belongs to the conducting region and the point (x,y) to an insulating layer characterized by $l = 0$. The particles can move into the insulator but they could never escape if the property of space $l = 0$ were assigned to (x,y) itself. On the other hand, if $l = 0$ is assigned to the space between the two points, no particle exchange can take place across the boundary. Force factor *f*, being a bias to *l* must share the same space. The rule is implemented in Fig.4 by making the subscript "x_-" to mean "half Δx to the left of x" and "y_-" to mean "half Δy below y". Using this notation we can state that during one time interval Δt there will be $[l_x(x_-,y) + f_x(x_-,y)]C(x-\Delta x,y,t)$ particles flowing from left to right and $[l_x(x_-,y)-f_x(x_-,y)]C(x,y,t)$ particles flowing from right to left. The net number of particles $N_x(x_-,y)$ flowing from $(x-\Delta x,y)$ to (x,y) is the difference of these two terms

$$N_x(x_-,y) = [l_x(x_-,y)+f_x(x_-,y)]C(x-\Delta x,y,t) - [l_x(x_-,y) - f_x(x_-,y)]C(x,y,t) \quad (6)$$

Similarly we get for the vertical or *y*-direction the following expression for the net number of particles $N_y(x,y_-)$ flowing from $(x,y-\Delta y)$ towards (x,y)

$$N_y(x,y_-) = [l_y(x,y_-)+f_y(x,y_-)]C(x,y-\Delta y,t) - [l_y(x,y_-) - f_y(x,y_-)]C(x,y,t) \quad (7)$$

Equations (6) and (7) contain the particle flux information which we intend to address later. At this time we remain focussed on our goal to obtain the complete transport equation of the form (3) in two dimensions and for

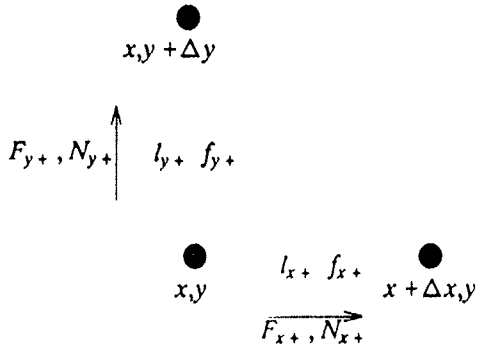

Fig. 5: The other points in two-dimensional space

variable diffusivities and forces. This requires that we know all the influxes and outflows from (x,y). Fig.4 guided us to obtain two of them, expressed in (6) and (7). With the aid of Fig.5 we obtain the other two which flow to the right and upward of (x,y) respectively.

We denote the net number of particles flowing out of (x,y) towards $(x+\Delta x,y)$ by $N_x(x_+,y)$. Their count is

$$N_x(x_+,y) = [l_x(x_+,y)+f_x(x_+,y)]C(x,y,t) - [l_x(x_+,y) - f_x(x_+,y)]C(x+\Delta x,y,t) \quad (8)$$

The net number of particles $N_y(x,y_+)$, flowing from (x,y) upward towards (x,y+ Δy) is

$$N_y(x,y_+) = [l_y(x,y_+)+f_y(x,y_+)]C(x,y,t) - [l_y(x,y_+) - f_y(x,y_+)]C(x,y+\Delta y,t) \quad (9)$$

We can write an expression for the particle count at time $t+\Delta t$ at the point (x,y) by adding all inflowing particles to $C(x,y,t)$ and subtracting all outflowing particles from it. We will allow for the possibility that in addition to particle exchange some of them are being generated at the rate $G(x,y)$ while some are being lost at the rate $R(x,y)$. This produces a net particle increase of $[G(x,y) - R(x,y)]\,\Delta t$ during the time interval Δt. The total particle count at (x,y) at $t+\Delta t$. The total particle count at (x,y) at $t+\Delta t$ is then

$$C(x,y,t+\Delta t) = C(x,y,t) + N_x(x_-,y)+N_y(x,y_-)-N_x(x_+,y)-N_y(x,y_+) + [G(x,y)-R(x,y)]\Delta t \quad (10)$$

Equation (10) in association with (6) through (9) is the discrete form of the two dimensional transport equation and we will make use of it shortly. But first we will perform a limiting process on it to obtain the equivalent differential equation. Rather than blindly substituting expressions (6) through (9) into (10) by noting that the difference $N_x(x_+,y,t) - N_x(x_-,y,t)$ refers to the midpoint between (x_-,y) and (x_+,y). But this is excatly (x,y) by our definition of x_- and x_+. Consequently we can write the difference $N_x(x_+,y,t) - N_x(x_-,y,t)$ as $\Delta_x N_x(x,y,t)$ where Δ_x signifies that the difference is with respect to x. When we do the same with the y-difference we get for (10)

$$C(x,y,t+\Delta t) = C(x,y,t) - \Delta_x N_x(x,y,t) - \Delta_y N_y(x,y,t) + [G(x,y) - R(x,y)]\Delta t \qquad (11)$$

The value for $N_x(x,y,t)$ can be formally derived from (6) or (8) by respectively incrementing or decrementing all x-arguments by x/2. The result is

$$N_x(x,y,t) = [l_x(x,y) + f_x(x,y)]C(x_-,y,t) - [l_x(x,y) - f_x(x,y)]C(x_+,y,t) \qquad (12)$$

In the above we have used the established notation for half-Δx values as, for example, $C(x - \Delta x/2,y,t) = C(x_-,y,t)$.

Similarly we get by either incrementing (7) or decrementing (9) by $\Delta y/2$ the expression for $N_y(x,y)$

$$N_y(x,y,t) = [l_y(x,y) + f_y(x,y)]C(x,y_-,t) - [l_y(x,y) - f_y(x,y)]C(x,y_+,t) \qquad (13)$$

Substitution of (12) and (13) into (11) yields

$$C(x,y,t+\Delta t) = C(x,y,t) - \Delta_x\left[l_x(x,y)[C(x_-,y,t) - C(x_+,y,t)]\right] - \Delta_x\left[f_x(x,y)[C(x_-,y,t) + C(x_+,y,t)]\right] - \Delta_y\left[l_y(x,y)[C(x,y_-,t) - C(x,y_+,t)]\right] - \Delta_y\left[f_y(x,y)[C(x,y_-,t) + C(x,y_+,t)]\right] + [G(x,y) - R(x,y)]\Delta t$$

This time we recognize the term $C(x_-,y,t) - C(x_+,y,t)$ to be the negative difference centered on (x,y). Consequently we can denote it by $-\Delta_x C(x,y,t)$. Similarly for the y-direction. Furthermore the sum of $C(x_-,y,t) + C(x_+,y,t) \approx 2C(x,y,t)$ and becomes exact when the increments go to zero which we are just about to do. We have now

$$C(x,y,t+\Delta t) = C(x,y,t) + \Delta_x[l_x(x,y)\Delta_x C(x,y,t)] - \Delta_x[f_x(x,y)2C(x,y,t)] + \Delta_y[l_y(x,y)\Delta_y C(x,y,t)] - \Delta_y[f_y(x,y)2C(x,y,t)] + [G(x,y) - R(x,y)]\Delta t$$

By means of (4) we convert f-factors to l and then employ (2) to convert l-factors to diffusivity D. We also subtract $C(x,y,t)$ on both sides of equation and divide them by Δt. A straightforward algebraic manipulation leads to

$$\frac{C(x,y,t+\Delta t) - C(x,y,t)}{\Delta t} = \frac{\Delta_x}{\Delta x}\left[D_x(x,y)\frac{\Delta_x C(x,y,t)}{\Delta x}\right] - \frac{\Delta_x}{\Delta x}\left[D_x(x,y)\frac{F_x(x,y,t)}{kT}C(x,y,t)\right] + \frac{\Delta_y}{\Delta y}\left[D_y(x,y)\frac{\Delta_y C(x,y,t)}{\Delta y}\right] - \frac{\Delta_y}{\Delta y}\left[D_y(x,y)\frac{F_y(x,y,t)}{kT}C(x,y,t)\right] + [G(x,y) - R(x,y)]$$

In the limit when all increments go to zero the above becomes a differential equation

$$\frac{\partial C(x,y,t)}{\partial t} = \frac{\partial}{\partial x}D_x(x,y)\left[\frac{\partial C(x,y,t)}{\partial x} - \frac{F_x(x,y)C(x,y,t)}{kT}\right] \left[\frac{kT}{(t,y,x)C(x,y,t)} - \frac{F_y(x,y,t)}{(t,y,x)C\partial}\right](x,y)D_y\frac{\partial}{\partial} + G(x,y) - R(x,y)$$

While this equation may be the correct starting point for a computerized solution of the transport equation when diffusivities, forces and concentrations vary in both dimensions it certainly does not guarantee the correctness of the solution. There are countless possible ways to discretize a differential equation but very few of them yield correct answers. Therefore it is strongly advisable to start from equation (10) which is derived from fundamental physical principles and circumvents the discretization problem altogether.

4. Flux of Particles in Two Dimensions.

Equations (12) and (13) provide us with the number of particles moving in the x and y direction, respectively, across the point (x,y). The corresponding flux is that number multiplied by the particle velocity. Our initial supposition was that particles move one space interval in one time interval Δt. Consequently their velocity is $\Delta x/\Delta t$ in the x-direction and $\Delta y/\Delta t$ in the y-direction. The two respective fluxes are then

$$\Phi_x(x,y,t) = -l_x(x,y)\frac{\Delta x^2}{\Delta t}\frac{C(x_+,y,t) - C(x_-,y,t)}{\Delta x}$$

$$+ f_x(x,y)[C(x_-,y,t) + C(x_+,y,t)]$$

$$\Phi_y(x,y,t) = -l_y(x,y)\frac{\Delta y^2}{\Delta t}\frac{C(x,y_+,t) - C(x,y_-,t)}{\Delta y}$$

$$+ f_y(x,y)[C(x,y_-,t) + C(x,y_+,t)]$$

The same kind of reasoning which led to the last differential equation in the previous Section guided us from here to the particle flux equations in differential form

$$\Phi_x(x,y,t) = -D_x(x,y)\left[\frac{\partial C(x,y,t)}{\partial x} - \frac{F_x(x,y)}{kT}C(x,y,t)\right]$$

$$\Phi_y(x,y,t) = -D_y(x,y)\left[\frac{\partial C(x,y,t)}{\partial y} - \frac{F_y(x,y)}{kT}C(x,y,t)\right]$$

If we multiply the flux by electric charge q we get the electric current density. The forces in this case are qE_x and qE_y, respectively for positive charges, i.e., holes and $-qE_x$ and $-qE_y$ for electrons where E_x and E_y are the electric fields. The concentrations $C(x,y,t)$ are $n(x,y,t)$ for electrons and $p(x,y,t)$ for holes. The corresponding four current density equations are

$$j_{nx}(x,y,t) = qD_{nx}(x,y)\left[\frac{\partial n(x,y,t)}{\partial x} + \frac{q}{kT}E_x(x,y,t)n(x,y,t)\right]$$

$$j_{ny}(x,y,t) = qD_{ny}(x,y)\left[\frac{\partial n(x,y,t)}{\partial y} + \frac{q}{kT}E_y(x,y,t)n(x,y,t)\right]$$

$$j_{px}(x,y,t) = -qD_{px}(x,y)\left[\frac{\partial p(x,y,t)}{\partial x} - \frac{q}{kT}E_x(x,y,t)p(x,y,t)\right]$$

$$j_{py}(x,y,t) = -qD_{py}(x,y)\left[\frac{\partial p(x,y,t)}{\partial y} - \frac{q}{kT}E_y(x,y,t)p(x,y,t)\right]$$

The reader may note that the so called Einstein relationship between carrier diffusivity D and their mobility μ is contained in the above equations

$$\mu = D\frac{q}{kT}$$

Before we leave this section let us examine equations (12) and (13) for physical consistency. Assume that the force factor f which depends on the electric field becomes larger than l in (12) or (13). This implies that the particle flow represented by the second term of either equation reverses its direction. This could quickly deplete the particles at the adjacent point and would even make their concentration go negative. Such physical impossibilities usually present themselves as numeric instabilities. We can therefore derive a stability condition from requiring that f never exceedes the value of l. Using (4) this condition can be writen as $F\Delta x/kT < 1$. For the electric case this translates into $E\Delta x < kT/q$. Because $E\Delta x$ represents the potential difference over one grid point our stability condition requires that this potential difference never exceeds 25 mV at room temperature.

5. Forces on Particles in Two Dimensions.

The last unresolved quantity that appears in equations (6) through (9) is the force factor f. According to (4) this is directly related to the electric field for the two respective directions. The electric field, on the other hand, emanates from electric charges as suggested by the Coulomb's law. Before we apply it to our case a discussion is in order.

In real, three-dimensional space the electric field surrounding a charge possesses spherical symmetry and is as such decaying as the square of the distance from the charge. If we were dealing with a three-dimensional case we would compute the field at a given position (x,y) by adding vectorially the contributions from all surrounding charges using an inverse square law. Then we would resolve that field into its Cartesian or other components as needed by the analysis. Instead of spherical symmetry we have a cylindrical symmetry for our two-dimensional case and consequently we can allow the electric field to decay only linearly with distance away from the source charge. This may appear unconventional but is no more so than is a two-dimensional space. In a one-dimensional space the field does not decay at all and we therefore compute the field as the sum of net charges without prorating their effects for distance. One must be quite cautious when applying natural laws to unnaturals spaces. With this in mind we continue now with the evaluation of force factors to be applied to equations (6) through (9).

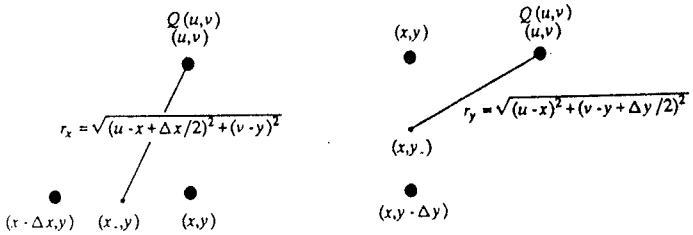

Fig. 6: Aid to calculation of the electric field components

We remember that the force factors have been defined halfway between discrete points. With the aid of Fig.6 we compute the x-component of the electric field at the point (x,y) as produced by the charge $Q(u,v)$ a distance

r_x away. A linear decay of the field with distance is employed as appropriate for our two-dimensional case.

$$E_x(x_.,y) = E(x_.,y)\frac{u-x_.}{r_x} = \frac{Q(u,v)}{2\pi r_x}\frac{u-x_.}{r_x} =$$

$$\frac{Q(u,v)}{2\pi}\frac{u-x_.}{(u-x_.)^2+(v-y)^2}$$

Similarly we get for the y-component at (x,y_) according to Fig.6

$$E_y(x,y_.) = E(x,y_.)\frac{v-y_.}{r_y} = \frac{Q(u,v)}{2\pi r_y}\frac{v-y_.}{r_y} =$$

$$\frac{Q(u,v)}{2\pi}\frac{v-y_.}{(v-y_.)^2+(u-x)^2}$$

The total field is the summation of components due to all charges Q(u,v)

$$E_x(x_.,y) = \sum_x\sum_y\sum_u\sum_v \frac{\frac{Q(u,v)}{2\pi}(u-x+\frac{\Delta x}{2})}{(u-x+\frac{\Delta x}{2})^2+(v-y)^2} \quad (16)$$

and

$$E_y(x,y_.) = \sum_x\sum_y\sum_u\sum_v \frac{\frac{Q(u,v)}{2\pi}(v-y+\frac{\Delta y}{2})}{(u-x)^2+(v-y+\frac{\Delta y}{2})^2} \quad (17)$$

We will now illustrate how (10) can be applied to impurity redistribution and to subsequent charge migration and how (16) and (17) can be used to compute the resulting forces opposing such migration. This is done in the hopes that the reader will be encouraged to experiment with multidimensional electric transport problems on his own.

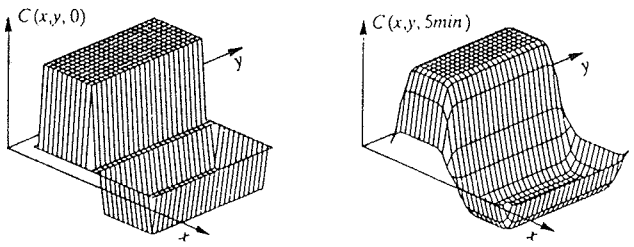

Fig. 7: Impurity profile before and after annealing.

6. An Example of a Two-Dimensional Semiconductor Structure

We start with the doping profile $C(x,y)$ shown on the LHS of Fig.7. It represents a PIN junction with a very narrow intrinsic region, with a donor density of N_D=4*10 ^{16} and an acceptor density of N_A=2*10 ^{16}. The profile shows the difference $C(x,y) = N_D(x,y) - N_A(x,y)$.

The widths of the two regions are approximately 2.5μ each. Next we anneal the sample by subjecting it to 1100°C for about 5 minutes. The diffusivity was chosen to be 5*10^{-13} at this temperature which is representative of commonly used dopants. The redistributed impurities are shown on the RHS of Fig.7.

This result was achieved by using (6) through (10) with the following parameters:

$l_x = l_y = .125$ independently of x and y, and $f_x = f_y = 0$ and $G - R = 0$ everywhere.

Equation (10) simplifies then into

$$C(x,y,t+\Delta t) = 0.5\,C(x,y,t) + 0.125\,[C(x-\Delta x,t) + C(x+\Delta x,t) + C(x,y-\Delta y,t) + C(x,y+\Delta y,t)] \quad (18)$$

Denote the number of Δx steps by X and the number of Δy steps by Y and assign a two-dimensional array $C(x,y)$ to the impurity concentration. We can then write the algorithm which represents the above equation in C-language as

```
for (y = 0; y<Y; y++)
  L[y] = 0.0; for (x = 1; x < X; x++)
  { for (y = 0; y < Y; y++)
    {temp = .5*C[x,y] + .125*(C[x - 1,y] + C[x + 1,y]
     + C[x,y - 1] + C[x,y + 1]);
     C[x - 1,y] = L[y]; L[y] = temp;
      }
} for (y=0; y < Y; y++)
{C[X -1,y] = L[y]; L[y] = 0.00;}
```

A few explanations are in order. First, the choice of l seems fairly open since it represents the diffusivity in conjuction with space and time steps according to (2). But if you consider the case where all the matter is initially concentrated in position (x,y) then we can have no more than 1/8 of the central matter moving out in each direction while preserving an equal amount of matter in the center. If we allow higher depletion, instabilities may occur. The rule of thumb is then to choose l_{max}=0.125 for highest diffusivity and proportionately less for all other values of D. In a three-dimensional case the maximum value of l should be chosen to be 1/12 and for the one- dimensional analysis l_{max}=.25.

For our example we have chosen 25 segments along the entire width and 25 along the entire lenght of the

sample corresponding to a space step of 0.2μ . According to (2) the time step Δt for our chosen diffusivity amounts to 100 seconds, meaning that every time we traverse the algorithm the time advances by this amount. Consequently, the annealed profile in Fig.7 is achieved by executing the algorithm only three times (five minutes).

In the C-algorithm above we have taken care of timing implied by (18) which demands that all RHS variables be taken at time *t*. As the execution moves from lower to higher values of x and y we must postpone the update of concentrations until we have moved to the next x-column. Therefore we save one whole x-column in L[y] and load it into C[x - 1,y] column after the old values are no longer needed. The first "for-loop" initializes the L-array to whatever boundary value we wish to assign to the *x* = 0 column. As the x-loop starts with the index 1, the column *x* = 0 receives the initial values contained in L[y]. The very last "for-loop" updates the last x-column at X-1 and resets the L-array to the initial value which will subsequently restore the boundary value at the *x* = 0 column again. Whatever we load into the X column which lies outside the reach of the x-loop becomes the upper boundary value for x. This column never gets updated so it supplies the boundary conditions which it contains to the X-1 column. In a similar fashion we can impose the boundary values to the y-column or to any point in the x-y domain, for that matter. This we do by loading such conditions into the appropriate location everytime the loops are traversed and the initial value there has been changed by the algorithm.

This lenghty discussion of the boundary value problem is intended to encourage those who have experienced difficulties with similar problems in the continuous domain. The discrete nature of computerized analysis makes the imposition of boundary conditions almost intuitive.

Next we must compute the concentrations of holes and electrons produced by the annealed profile. This is quite straightforward when we invoke the neutrality condition N_D-N_A + *p*-*n* = 0 and the mass- action law $n\,p = n_i^2$ [3,4].

The following algorithm has been used to obtain Fig.8

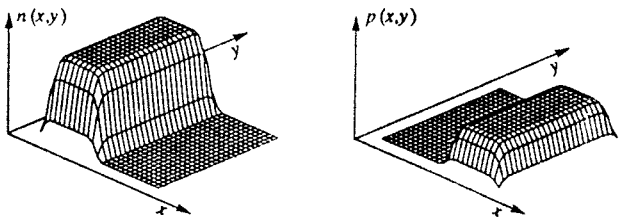

Fig. 8: Electron and hole distribution.

```
for (x = 0; x < X; x++)
  { for (y = 0; y < Y; y++)
  {if (C[x,y] > 0.0)
   {n[x,y] = C[x,y]/2.0 + sqrt(C[x,y]*C[x,y]/4.0 +
     2.0e20);
   p[x,y] = 2.0e20/n[x,y];
     }
  if (C[x,y] < 0.0)
   {p[x,y] = - C[x,y]/2.0 + sqrt(C[x,y]*C[x,y]/4.0 +
    2.0e20);
    n[x,y] = 2.0e20/p(x,y);
     }
  if (C[x,y]= = 0.0) (n[x,y] = 0.0; p[x,y] = 0.0;}
    }
}
```

where 2.0e20 stands for squared intrinsic concentration of Silicon.

Gradients of holes and electrons seen in Fig.8 give rise to their migration into the adjacent regions of low density which destroys the initially imposed neutrality. This gives rise to electric fields which tend to oppose the migration and which we will compute later. First we use (10) again to compute the movement of charges but this time we want to distinguish the migration rates of holes and

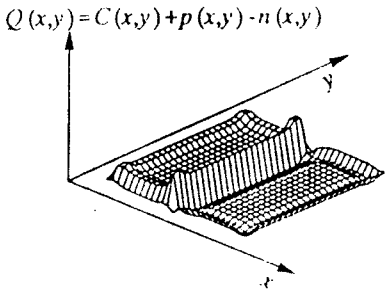

Fig. 9: Net charge distribution after 10 picoseconds

electrons. We assign a diffusivity D_n = 28 to the latter and D_p = 9 to the former. These numbers are in fair agreement with the respective impurity concentrations. We assign the value l_n = .125 to the highest diffusivity. The corresponding value of l_p must then be .125x9/28 ≈ .04. With these values and with the force factors initially set to zero we have obtained the net charge density distribution *Q(x,y)* shown in Fig.9 after approximately 10 picoseconds.

It is apparent from Fig.9 that the departure of electrons along the four boundaries of the *N*-region produces a net positive charge. Similarly the departure of holes from the *P*-region produces a net negative charge all around the boundary. A charge reversal is observed at the junction. Implemention in C-language of equations (8) through (10) which have produced Fig.9 is shown below.

```
for (i = 0; i < 5; i++)
  {for (x= 0; x < X; x++)
    { for (y= 0; y < Y; y++)
      {Fxy = .125*(fx[x,y] - fx[x+1,y] + fy[x,y]
       - fy[x,y+ 1]);
    tempN = (.5-Fxy)*n[x,y] + .125*(1.0-fx[x,y])*n[x-1,y]
+ .125*(1.0 + fx[x+1,y])*n[x+1,y]+.125*(1.0-
fy[x,y])*n[x,y-1] + .125*(1.0 + fy[x,y+1])*n[x,y+1];
    n[x-1,y] = L[y]; L[y] = tempN;
    tempP= (.84 + Fxy)*p[x,y] + .04*(1.0 + fx[x,y] *p[x-
1,y] + .04*(1.0 - fx[x+1,y]*p[x+1,y] + .04*(1.0 +
fy[x,y])*p[x,y-1] + .04*(1.0 - fy[x,y+1])*p[x,y+1];
      p[x-1,y] = L[y]; L[y] = tempP;
    }
} } for (x= 0; x < X; x++)
  for (y= 0; y < Y; y++)
    Q[x,y] = C[x,y] + p[x,y] - n[x,y];
```

In the above we have made the following substitutions

$$fx[x,y] = E_x(x_-,y)\Delta x\frac{q}{kT} \qquad fx[x+1,y] = E_x(x_+,y)\Delta x\frac{q}{kT}$$

$$fy[x,y] = E_y(x,y_-)\Delta y\frac{q}{kT} \qquad fy[x,y+1] = E_y(x,y_+)\Delta y\frac{q}{kT}$$

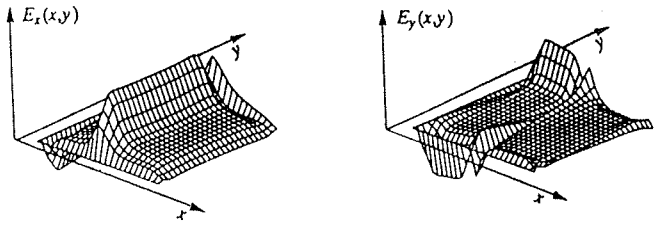

Fig. 10: Components of the electric field.

We have allowed five passes through the algorithm. Because the time interval is $\Delta t = 1/(\Delta x)^2/D$ each pass is somewhat shorter than 2 picoseconds. The net charge density $Q(x,y)$ is obtained from the resulting distributions of $p(x,y)$ and $n(x,y)$ as the last step in the algorithm. When we substitute this into (16) and (17) we obtain the two components of the electric field. These are shown in Fig.10 and a C-version of the algorithm which generated the numeric values is reproduced below the figure.

```
for (y = 0; y < Y; y++)
  for (x = 0; x < X; x++)
   {EX = EY = 0.0;
    for (u = 0; u < X; u++)
      for (v = 0; v < Y; v++)
     {delY = v-y;
    if(delY! = 0)
     {delX = u - x + .5;
     EX = EX + Q [u,v] *delX/(delY*delY + delX*delX);
      }
     delX = u - x;
     if (delX! = 0)
       {delY = v - y + .5;
      EY = EY + Q[u,v] *delY/ (*delY*delY + delX*delX),
       }
     }
   fx[x,y] = EX*1.0e-10;
   fy[x,y] = EY*1.0e-10;
    }
```

The factor 10^{-10} in the above takes account of the permittivity of Silicon, of the integration intervals of the factor q/kT and of the conversion between surface and volume density of charges involved. An interpretation of the major features of Fig. 10 tells us that the force on electrons and holes is directed inward both along the x-axis. The y field shows more intensity in the N-region than in the P which has its origin in the higher donor concentration and electron diffusivity. One major but expected feature of the electric field E_x is its high intensity along the junction region. This field eventually blocks the further migration of electrons and holes into the adjacent regions. An actual simulation of a semiconductor structure would alternate between the last two algorithms, computing the net charges from the first one and finding the resulting force factors from the latter one. These would then be substituted back into the first one. An equilibrium would eventually be established at which point te migration would cease completely.

Before we conclude let us recognize that the electronic fields could be computed in an alternate way. It is known [5], [6], [7] that the expression we have derived for the electric field from Coulomb's law is in fact the formal solution of the Poisson's differential equation

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = \frac{Q(x,y)}{\varepsilon}$$

Where Q is the charge density. Unfortunately this one equation with two unknowns and as such not very useful. But it is also known from the field vector theory [8] that field components E_x and E_y are the gradients of a function $V(x,y)$ which satisfies the Poisson equation in the following way

$$\frac{\partial^2 V(x,y)}{\partial x^2} + \frac{\partial^2 V(x,y)}{\partial y^2} = -\frac{Q(x,y)}{\varepsilon}$$

When this equation is solved for the scalar quantity $V(x,y)$ the field components can be obtained as the respective partial derivatives

$$E_x(x,y) = -\frac{\partial V(x,y)}{\partial x} \quad \text{and} \quad E_y(x,y) = -\frac{\partial V(x,y)}{\partial y}$$

It turns out that the effort in solving the Poisson equation for the potential is comparable to what we have done and our approach is more in keeping with the promise of staying close to first principles. Nevertheless it ought

to be pointed out that the Poisson equation in terms of potential $V(x,y)$ can be solved by the same algorithm we have developed for the particle transport by diffusion. The diffusion equation in two dimensions with constant diffusivity has the differential form

$$\frac{\partial V(x,y,t)}{\partial t} = \frac{\partial^2 V(x,y,t)}{\partial x^2} + \frac{\partial^2 V(x,y,t)}{\partial y^2}$$

When add a source function $\frac{Q(x,y)}{\varepsilon}$ to the RHS of the above and solve the equation for steady state condition, i.e. $\partial V(x,y,t)/\partial t$ the equation goes over into the form

$$\frac{\partial^2 V(x,y)}{\partial x^2} + \frac{\partial^2 V(x,y)}{\partial y^2} = -\frac{Q(x,y)}{\varepsilon}$$

Consequently we can solve the Poisson equation with the aid of equation (15) as implemented in the respective C-algorithm when we add the source function $Q(x,y)/\varepsilon$ to the RHS of the temp expression. When there are no more changes between two successive evaluations of temp the resulting distribution C[x,y] is the solution of the Poisson equation. The results obtained for our test case are identical with one or the other method and the time consumed is about the same.

7. Conclusions

The transport equation and the equivalent of the Poisson equation were derived from first principles without invocation of unsubstantiated abstraction. The two-dimensional case has been tested on a simple semiconductor structure and the corresponding algorithms have been presented in the C-language. The intent was to show that the availability of computers warrants a fresh look at the traditionally accepted mathematical models and that more intuitive approaches are made possible by taking advantage of present-day computer performance.

Ackknowledgement

The highly efficient and versatile algorithm which was used to plot the two-dimensional distributions was developed and given to me by Ing. Dragan Fidler of Zemun.

References

[1] SUPREME by Stanord, PISCES and DEPICT by Tehnology Modeling Associates, SIMPEL by Berkeley, MASTER by Silvacor.

[2] Z.Fazarinc, Discretization of Partial Differential Equations for Computer Evaluation, Computer Applications in Engineering Education, Vol.1, No.1, pp.73-85, John Wiley, 1992-93.

[3] C.Kittel and H.Kroemer, Thermal Physics 2-nd Ed., Freeman & Co., San Francisco 1980, p270 Derived from Thermodynamics, p268 from reaction rates, p361-3 from density of states.

[4] A.S.Grove, Physics and Technology of Semiconductor Devices, John Wiley 1967, p100-1.

[5] James C. Maxwell, A Treatise on Electricity & Magnetism, Clarendon Press, 1891, Chapter II.

[6] Oliver Heaviside, Electrician, Phil. Mag. XXVII, 1889, p.324

[7] Simeon D.Poisson, Bull.de la Soc. Philomathique iii, 1813, p.388

[8] I.Sokolnikoff and R.Redheffer, Mathematics of Physics and Modern Engineering, McGraw Hill, 1958, p386-7

prof. Zvonko Fazarinc, Ph.D
Stanford University, ret.
Hewlett-Packard Laboratories, ret.
880 La Mesa Drive
Menlo Park, CA 94025

Prispelo: 27.08.93 *Sprejeto:17.09.93*