Elaborato di Tesi ...

Titolo progetto di tesi ...

Progetto svolto da:
Andrea Bortolossi
Matr. 783023

Indice

Capitolo 1

Introduction

1.1 Brief history of VLSI devices

In 1947 John Bardeen, William Shockley and Walter Brattain (three scientists of Bell Telephone Labs) invented the bipolar transistor and since that crucial point there has been a growth of the semiconductor industry never known before, with serious impact on the way people work and live today.

Before reach the functionality and the miniaturization of modern devices, some fundamental steps has been made. In 1958 was produced the first intagrated circuits (IC) followed by the introduction of the first MOSFET(1960) and CMOS(1963). Into these inventions the first micro-processor(1971) sank his roots and since that time until present, an ever-increasing progress has continued, according to the indication of *Moore's Law* (formulated by Gordton Moore in 1965).

These events led microelectronic industry at the doors of the VLSI era (Very-Large-Scale-Integration). Indeed in the last thirty years the benefits of miniaturization have been the key in the evolutionary progress leading to today's computers, wireless units, and comunication systems that offer superior performance, dramatically reduced cost per function, and much reduced physical size.

The large worldwide investment in VLSI technology constitutes a formidable driving force that guarantee the continued progress in IC integration density and speed, for as long as physical principles will allow.

1.2 Why FEMOS

From this point we want start and remark that the aim of simulation is the full comprehension of the physical phenomenon which lies behind the func-

tion of modern device. As already underlying this situation became rapidly more important since in the last years devices became more complex and in many cases compact models are insufficent to fully describe the behavoiur of devices.

Even if there's exist many commercial software which are able to resolve different physic situations, it's really difficult satisfy the necessities of industries. A simulator should be more desireable if it could couple any kind of equations but this is a far ahievement. Modern software are often specialized on precise physic branch. Obviously this strategic choice guarantees more efficiency but it implies a lost in generality. The conseguence is that the work of the model analyst became harder when he have to afford problems located in the middle of different phenomenon.

Consider for a moment to analyse the funcionality of a new device, which its electric behavoiur is strong influeced by its mechanic response. Basically you are interested to the resolution of Maxwell's law (which is well performed by SDEVICE simulator) and the Navier-Lamè equations (which is well performed by COMSOL simulator). Now the question is: how to put in comunication the different outputs? Take into account that it's not possbile known precisely how the above programs resolves the equations, which implies a relevant risk when you decide to combine the solutions. In other words the development of an own code is at least desirable and possibly helpful. The main advantage is the total control on simulation procedure and the possibility of fully customize. Although the major drawback is that the improvement of a personal code needs time and human resources, which in many cases are not avaible.

This is the framework where FEMOS code (Finite Element Method Oriented Solver) operates.

Si introduce l'importanza e l'efficacia della gummel map. Ho bisogno di qualche fonte in più.

Una volta elencati i pregi e i difetti orientiamo l'introduzione ai reali obiettivi della tesi. Dunque chiarito il background in cui ci stiamo muovendo si è deciso di introdurre anche una libreria dedicata alla simulazione di materiali semiconduttori (i quali differiscono ampiamente rispetto agli altri materiali e questo va aggiunto nello stato dell'arte). Non solo differiscono nelle proprietà ma anche nell'impostazione del tipo

1.3 Aims of this work

The work of thesis is about the achievement of the sequent targets:

- development of a finite element based simulator for semiconductor devices which deals with multiple generation/recombination and mobility models;

- check solutions obtained with commercial software;

- definition and implementation of a new way to compute the current density inside the device;

- extension of the residual method presented in referenza for the 3D case;

- evaluation of the possibility to extend the residual method at the computation of the current density inside the device.

Capitolo 2

Semiconductor model

2.1 Basic Device Physics

2.1.1 Intrinsic semiconductor

In a silicon crystal each atom has four valence electrons to share with its four nearest neighboring atoms. The valence electrons are shared in a paired configuration called a covalent bond. The most important result of the application of quantum mechanics to the description of electrons in a solid is that the allowed energy levels of electrons are grouped into bands. The bands are separated by regions of energy that the electrons in the solid cannot possess: forbidden gaps. The highest energy band that is completely filled by electron at 0[K] is called the *valence band*. The next higher energy band, separated by a frobidden gap from the valence band, is called the *conduction band*.

What sets a semiconductor such as silicon apart from a metal or an insulator is that at absolute zero temperature, the valence band is comletely filled with electrons, while the conduction band is completely empty and the *bandgap* is on the order of 1[eV]. This implies that at room temperature a small fraction of the electrons are excited into the conduction band, leaving behind vacancies (called *holes*) in the valence band. In contrast, an insulator has a much larger forbidden gap making room-temperature conduction virtually impossible while metals have partially filled conduction bands even at absolute zero temperature, this make them good conductors at any temperature.

In every cases the energy distribution of electrons in a solid is governed by the laws of Fermi-Dirac statistics. For a system in thermal equilibrium, the principal result of these statistics is the *Fermi-Dirac distribution function*, which gives the probability that an electronic state at energy E is occupied

by an electron,

$$f_D(E) = \frac{1}{1 + exp\left(\dfrac{E - E_f}{KT}\right)} \tag{2.1}$$

here $K = 1.38 \times 10^{-23}[J/K]$ is Boltzmann's constant,T is the absolute temperature and E_f is a parameter called the *Fermi level.*

Definition 2.1. The Fermi level (E_f) is the energy at which the probability of occupation of an energy state by an electron is exactly one-half.

In most cases when the energy is at least several KT above or below the Fermi level (2.1) can be approximated with the Maxwell-Boltzmann statistics for classical particles, which read as follows:

$$f_D(E) \simeq f_{MB}(E) = \begin{cases} exp\left(-\dfrac{E - E_f}{KT}\right) & E \gg E_f \\ 1 - exp\left(-\dfrac{E_f - E}{KT}\right) & E \ll E_f \end{cases} \tag{2.2}$$

Fermi level plays an essential role in characterizing the equilibrium state of a stystem, it is important to keep in mind the sequent observation.

Observation 2.1. When two systems are in thermal equilibrium with no current flow between them, their Fermi levels must be equal, in other words for a continuous region of metals and/or semiconductors in contact, the Fermi level at thermal equilibrium is flat (spatially constant throughout the region).

What we are interested is a suitable formulation of che carrier concentration which for electrons is given by the sequent integral:

$$n = \int_{E_c}^{\infty} N(E) f_D(E) \, dE \tag{2.3}$$

With $N(E)dE$ we indicate the number of electronic states per unit volume with an energy between E and $E + dE$ in the conduction band. In general (2.3) is a Fermi integral of the order $1/2$ and must be evaluated numerically. In order to avoid this complication it's useful consider semiconductor materials such that theirs Fermi levels stay at least $3KT/q$ below the edge of the conduction band (for holes we consider the same approximation above the valence band). This kind of semiconductors are usual known as nondegenerated. The concrete consequence of this hypotesis is that we can approximate the Fermi-Dirac distribution by the Maxwell-Boltzmann distribution and resolve (2.3) in the analytically way, obtained the sequent results,

$$n = N_c exp\left(-\frac{E_c - E_f}{KT}\right) \quad (2.4)$$

$$p = N_v exp\left(-\frac{E_f - E_v}{KT}\right) \quad (2.5)$$

where N_c and N_v are the *effective density of states*. If we consider an intrisic semiconductor we can say that $n = p$ and the above equations can be rewritten in terms of the intrinsic carrier density (n_i) and the midgap energy level (E_i) :

$$n = n_i exp\left(\frac{E_f - E_i}{KT}\right) \quad (2.6)$$

$$p = n_i exp\left(\frac{E_i - E_f}{KT}\right) \quad (2.7)$$

Introdurre il discorso del midgap e dell'equilibrio np=ni2

$$np = n_i^2 \quad (2.8)$$

$$E_i = E_f = \frac{E_c + E_v}{2} - \frac{KT}{2}ln\left(\frac{N_c}{N_v}\right) \quad (2.9)$$

$$n_i = \sqrt{N_c N_v}exp\left(-\frac{E_g}{2KT}\right) \quad (2.10)$$

2.1.2 Extrinsic semiconductor

Intrinsic semiconductor at room temperature has an extremely low free-carrier concentration, therefore, its resistivity is very high. In order to make semiconductor a better conductor it's usual add impurities which introduce additional energy levels in the forbidden gap and can be easily ionized to add either electrons to the conduction band or holes to the valence band. In other words the electrical conductivity of silicon is then dominated by the type and concentration of the impurity atoms, or dopants.

There are some different materials used to construct electric devices but the most known and used is silicon. One of the most important charcateristic of silicon is that it's a column-IV element with four valence electrons per atom. This implies that there are two types of impurities which are electrically active: those from column V such as arsenic or phosphorus, and those from column III such as boron.

A column-V atom in a silicon lattice tends to have one extra electron loosely bonded after forming covalent bonds with other silicon atoms. In most cases, the thermal energy at room temperature is sufficient to ionize the impurity atom and free the exra electron to the conduction band. Such type of impurities are called *donors*; they become positively charged when ionized. Silicon material doped with column-V impurities or donors is called *n-type* silicon.

One the other hand, a column-III impurity atom in a silicon lattice tends to be deficient by one electron when forming covalent bonds with other silicon atoms. Such an impurity atom can also be ionized by accepting an electron from the valence band, which leaves a free-moving hole that contributes to electrical conduction. These impurities are called *acceptors*: they become negatively charged when ionized. Silicon material doped with column-III impurities or acceptors is called *p-type* silicon.

A p-type or an n-type is named as *extrinsic* silicon. In terms of the energy-band diagrams, donors add allowed electron states in the bandgap close to the conduction-band edge, while acceptors add allowed states just above the valence-band edge.

In contrast to intrinsic silicon, the Fermi level in an extrinsic silicon is not located at the midgap. The Fermi level in n-type silicon moves up towards the conduction band and on the other hand the Fermi level in p-type silicon moves down towards the valence band. The exact position of the Fermi level depends on both the ionization energy and the concentration of dopants. For example, for an n-type material with a donor impurity concentration N_d the charge neutrality condiction in silicon requires that

$$n = N_d^+ + p \tag{2.11}$$

where N_d^+ is the density of ionized donors. Similarly for a p-type material with acceptor impurity concentration N_a we have

$$p = N_a^- + n \tag{2.12}$$

For the sake of simplicity we consider in this work that at room temperature all impurties are ionized ($N_d = N_d^+$ and $N_a = N_a^-$). Substituting (2.4) and (2.5) in (2.11) and (2.12), solved the algebraic equation, one obtains

$$E_c - E_f = KTln\left(\frac{N_c}{N_d}\right) \tag{2.13}$$

$$E_f - E_v = KTln\left(\frac{N_v}{N_a}\right) \tag{2.14}$$

9

Equation (2.8) is independent of the dopant type and Fermi level position and in n-type and p-type respectively we obtain

$$p = \frac{n_i^2}{N_d} \qquad n = \frac{n_i^2}{N_a} \tag{2.15}$$

Instead of using N_c, N_v and referring to E_c and E_v equation (2.13) and (2.14) can be written in a more useful form in terms of n_i and E_i defined by eqautions (2.10) and (2.9):

$$E_f - E_i = KTln\left(\frac{N_d}{n_i}\right) \tag{2.16}$$

$$E_i - E_f = KTln\left(\frac{N_a}{n_i}\right) \tag{2.17}$$

Observation 2.2. The distance between the Fermi level and the intrinsic Fermi level near the midgap is a logarithmic function of doping concentration.

This situation causes two important consequences:

- non linearity relations between potential and desities,

- exponential dependence of densities from potential.

2.1.3 Densities at nonequilibrium

The above discussion applies only when both the electron end hole densisties take on their local equilibrium values and a local Fermi level can be defined. It is often in VLSI device operation to encounter nonequilibrium sistuations where the densties of one or both types of carriers depart from their equilibrium values given by (2.6) and (2.7). In particular, the minority carrier concentration can be easily overwhelmed by injection from neighboring regions. Under these circumstances, while the electrons are in local equilibrium with themselves and so are the holes, electrons and holes are not in equilibrium with each other. In order to extend the kind of relationship between Fermi level and current densities discussed above, one can introduce separate Fermi lvels for electrons and holes. They are called *quasi Fermi levels* defined as E_{fn} and E_{fp} they replace E_f in (2.6) and (2.7):

$$n = n_i exp\left(\frac{E_{fn} - E_i}{KT}\right) \tag{2.18}$$

$$p = n_i exp\left(\frac{E_i - E_{fp}}{KT}\right) \tag{2.19}$$

In this regard, quasi Fermi levels have a similar physical interpretation in terms of the state occupancy as the Fermi level.

Observation 2.3. The electron density in the conduction band can be calculated as if the Fermi level is at E_{fn}, and the hole density in the valence band ca be calculated as if the Fermi level is ag $_{fp}$.

2.1.4 Carrier transport in semiconductor

Carrier transport or current flow in silicon is driven by two different mechanisms:

- the **drift** of carriers, which is caused by the presence of an electric field;

- the **diffusion** of carriers, which is caused by and electron or hole concentration gradient.

Drif current - Ohm's law

When an electric field is applied to a conducting medium containing free carriers, the carriers are accelerated and acquire a drift velocity superimposed upon their random thermal motion.

Observation 2.4. The drift velocity of holes is in the direction of the applied field, and the drift velocity of electrons is opposite to the field.

The velocity of the carriers does not increase indefinitely under field acceleration, since they are scattered frequently and lose their acquired momentum after each collision. During their motion throughout the lattice struccture, carriers travel at an average speed definded as

$$\mathbf{v}_d^\eta = \pm \frac{q\mathbf{E}\tau_\eta}{m_\eta} \qquad \eta = \{h, e\} \tag{2.20}$$

where $q = 1.602e^{-19}[C]$ is the elementary charge, $\mathbf{E}$ is the electric field, τ_η is the average time of flight of the carrier between two consecutive interactions with the atoms of the lattice and m_η is the effective mass. The coefficient $q\tau_\eta/m_\eta$ characterizes how quickly a carrier can move through the lattice and it's well known as carrier mobility. It's usual idicate this quantity with the greek letter μ, its standard dimensions are $[m^2 V^{-1} s^{-1}]$. In general, one can use *Matthiessen's rule* to include different contributions to the mobility:

$$\frac{1}{\mu} = \frac{1}{\mu_L} + \frac{1}{\mu_I} + \cdots \tag{2.21}$$

where μ_L and μ_I correspond to the lattice and impurity scattering limited components of mobility, for a more detailed description of mobility models see [YT09]. We chose this approach during the implementation of the code.

Therefore the drift current density for a given electron (n) or hole (h) concentration, reads as follows:

$$\mathbf{J}_n = - qn\mathbf{v}_d^n = qn\mu_n\mathbf{E} \tag{2.22}$$
$$\mathbf{J}_p = + qp\mathbf{v}_d^p = qp\mu_p\mathbf{E} \tag{2.23}$$

The scalar coefficient $qn\mu_n(qp\mu_p)$ is often summerized by the electron (hole) conductivity $\sigma_n(\sigma_p)$. Now you can rewritten (2.22) and (2.23) and obtain:

$$\mathbf{J}_n = \sigma_n\mathbf{E} \tag{2.24}$$
$$\mathbf{J}_p = \sigma_p\mathbf{E} \tag{2.25}$$

Relations (2.24) and (2.25) expresses the well known *Ohm' law* stating that in the conducting material the current density is directly proportional to the applied electric field.

Diffusion current - Fick's law

In semiconductor devices it's usual have different profiles of dopant in order to allow particular behaviors, this implies a not uniform concentration of carriers which they also diffuse as a result of the concentration gradient. This leads to an additional current contribution in proportion to the concentration gradient, in mathematical terms, diffusion current densities are given accordingly by the *Fick's law* as follows:

$$\mathbf{J}_n = -D_n(-q\nabla n) \tag{2.26}$$
$$\mathbf{J}_p = -D_p(+q\nabla p) \tag{2.27}$$

The proportionally constants D_n and D_p are called the electron and hole diffusion coefficients and have units of $[cm^2 s^{-1}]$. Physically, both drift and diffusion are closely associated with the random thermal motion of carriers and their collisions with the silicon lattice in thermal equilibrium. A simple relationship between the diffusion coefficient and the mobility is the well known *Einstein relation*:

$$D_\eta = \frac{KT}{q}\mu_\eta \tag{2.28}$$

Drift-Diffusion transport equations

Finally we can write the constitutive laws for the current density of electrons and holes:

$$\mathbf{J}_n = qn\mu_n\mathbf{E} + qD_n\nabla n \tag{2.29}$$
$$\mathbf{J}_p = qp\mu_p\mathbf{E} - qD_p\nabla p \tag{2.30}$$

The total conduction current density is $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p$.

One interesting thing is that these constitutive laws can be rewritten in two other ways at least. We want remark this because it's useful find various physical explanations of the same phenomenon and moreover these reinterpreations give different start points for the discrete solver algorithm.

It should be nice connect current density with the quasi Fermi level, this is possible in fact using the relation between electric potential and electric field,

$$\mathbf{E} = -\nabla\varphi \tag{2.31}$$

one can write the current densities as:

$$\mathbf{J}_n = -qn\mu_n\left(\nabla\varphi - \frac{KT}{qn}\nabla n\right)$$
$$\mathbf{J}_p = -qp\mu_p\left(\nabla\varphi + \frac{KT}{qp}\nabla p\right)$$

Equations (2.18) and (2.19) can be substituted into the above to yeld:

$$\mathbf{J}_n = -qn\mu_n\nabla\varphi_n \tag{2.32}$$
$$\mathbf{J}_p = -qn\mu_p\nabla\varphi_p \tag{2.33}$$

With these equations we underlying an important aspect which occur in semiconductor material:

Observation 2.5. The current density is proportional to the gradient of the quasi Fermi potential instead of the electric field.

The third way to represent the current density is based on *Slotboom variables*. In 1973 Jan Slotboom proposed this change in variables for the two-dimensional numercal simulation of a bipolar transistor:

13

$$u_n = n_i exp\left(-\frac{\varphi_n}{V_{th}}\right) \tag{2.34}$$

$$u_p = n_i exp\left(\frac{\varphi_p}{V_{th}}\right) \tag{2.35}$$

Using the above equations into (2.29) and (2.30) we obtain:

$$\mathbf{J}_n = qD_n exp\left(\frac{\varphi}{V_{th}}\right) \nabla u_n \tag{2.36}$$

$$\mathbf{J}_p = -qD_p exp\left(-\frac{\varphi}{V_{th}}\right) \nabla u_p \tag{2.37}$$

This interpretation totally change the point of view on current density behavior:

Observation 2.6. It's possible explain the drift-diffusion current density in a semiconductor, with a flux totally diffusive of a new kind of carrier and diffusivity coefficient.

The vantages and drawback will be discussed better in chapter ...

2.2 Drift Diffusion Model for semiconductor

Simulations on integrated devices works on several different scale, this is the first point which a developer may consider when start to project a simulator code. Include every phenomenon from tha atomic to the macro scale is not possible. An excellent trade-off between machine time cost and physical accuracy is presented by the *Drift Diffusion model* (DD), which is by far the most widely used mathematical tool for industrial simulaiton of semiconductor devices. In this section we'll show how is obtained the DD model used for our simulations.

2.2.1 Drift Diffusion formulation

In order to describe the propagation of any electromagnetic signal in a medium, we have to start from the system of Maxwell equations, which reads as follows:

$$
\begin{cases}
\nabla \times \mathbf{H} &=& \mathbf{J} + \dfrac{\partial \mathbf{D}}{\partial t} \\[2ex]
\nabla \times \mathbf{E} &=& -\dfrac{\partial \mathbf{B}}{\partial t} \\[2ex]
\nabla \cdot \mathbf{D} &=& \rho \\[2ex]
\nabla \mathbf{B} &=& 0
\end{cases}
\tag{2.38}
$$

We are able to complete the system with the following set of constitutive laws that characterize the electromagnetic properties of the medium:

$$
\begin{array}{rcl}
\mathbf{D} &=& \epsilon \mathbf{E} \\
\mathbf{B} &=& \mu_m \mathbf{H}
\end{array}
\tag{2.39}
$$

From (2.38) we elaborate the DD model, through some interesting hypostesis which are:

- Lorentz-Gauge for the vector potential of $\mathbf{B}$.

- Quasi static approximation.

The second one is related with the IC component sizes and it is a reasonable hypotesis for our simulations. The system obtained after this suitable approximation looks as follows:

$$
\begin{cases}
\nabla \cdot (-\epsilon \nabla \varphi) &=& \rho & \quad \textit{Poisson equation} \\[2ex]
\dfrac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} &=& 0 & \quad \textit{Continuity equation}
\end{cases}
\tag{2.40}
$$

To close the above system we need to specify the mathematical form of the electric charge density (ρ) and the electric conduction current density ($\mathbf{J}$).

As we introduced in the preview section, devices are usually formed by extrinsic semiconductor and this causes the presence in the lattice of two kind of charge:

- free charge (ρ_{free}) (free electron and holes carriers),

- fixed charge (ρ_{fixed}) (inoized dopant impurities).

$$\rho = \underbrace{q(p - n)}_{\rho_{free}} + \underbrace{q(N_D^+ - N_A^-)}_{\rho_{fixed}} \qquad (2.41)$$

Notice that we assume N_D^+ and N_A^- time invariant.

Accordingly with the preview hypotesis and replacing (2.41), (2.29) and (2.23), we can split the continuity equation into the contribute of electrons and holes, the DD model formulation looks as follows:

$$\begin{cases} \nabla \cdot (-\epsilon \nabla \varphi) &= q(p - n + N_D^+ - N_A^-) \\[2ex] -q\dfrac{\partial n}{\partial t} + \nabla \cdot (-q\mu_n n \nabla \varphi + qD_n \nabla n) &= qR \\[2ex] q\dfrac{\partial p}{\partial t} + \nabla \cdot (-q\mu_p p \nabla \varphi - qD_p \nabla p) &= -qR \end{cases} \qquad (2.42)$$

The system is an incompletely parbolic initial value/boundary problem in three scalar unkown dependent variables $\varphi(\mathbf{x}, t)$, $n(\mathbf{x}, t)$ and $p(\mathbf{x}, t)$. Notice that the problem is a nonlinearly coupled system of PDE's, because of the presence of the drift terms $n\nabla \varphi$ and $p\nabla \varphi$.

From Maxwell equations we are able to guarantee only that $\mathbf{J}$ is a solenoidal field, we can't say nothing about the properties of $\mathbf{J_n}$ and $\mathbf{J_p}$. For this reason there is a new term in the right hand side. We can interpret $R(\mathbf{x}, t)$ as the net rate of generation and recombination.

We consider also the stationary form for our purpose.

$$\begin{cases} \nabla \cdot (-\epsilon \nabla \varphi) &= q(p - n + N_D^+ - N_A^-) \\[2ex] \nabla \cdot (-q\mu_n n \nabla \varphi + qD_n \nabla n) &= qR \\[2ex] \nabla \cdot (-q\mu_p p \nabla \varphi - qD_p \nabla p) &= -qR \end{cases} \qquad (2.43)$$

2.2.2 Generation and Recombination phenomenon

The modelling of $R(\mathbf{x}, t)$ is one of the most important feature which one could take into account. In this work we have implemented three distincted contribute based on the classical theory.

Shockley-Read-Hall recombination

Auger recombination

Impact ionization

Capitolo 3

Resolution of the system

3.1 Iteration algorithms

The high nonlinear nature of the problem makes an analytical treatment very difficult, if not even impossible. For this reason, numerical schemes must be used to compute an approximate solution of the system (2.43). Indisputably, the most used algorithms are *the fully coupled Newton's method* and *the decoupled Gummel map*. We spent only few words on the former as we implemented the latter, but first of all let us introduce a suitable linearization of the stationary DD equations. We can consider (2.43) in a more compact form as:

$$\mathbf{F}(\mathbf{U}) = \mathbf{0} \tag{3.1}$$

where:

$$\mathbf{U} := [\varphi, n, p]^T \qquad \mathbf{F}(\mathbf{U}) := \begin{bmatrix} F_1(\mathbf{U}) \\ F_2(\mathbf{U}) \\ F_3(\mathbf{U}) \end{bmatrix} \tag{3.2}$$

and having set:

$$F_1(\mathbf{U}) = \nabla \cdot (-\epsilon \nabla \varphi) - q(p - n + N_D^+ - N_A^-)$$
$$F_2(\mathbf{U}) = \nabla \cdot (-q\mu_n n \nabla \varphi + qD_n \nabla n) - qR$$
$$F_3(\mathbf{U}) = \nabla \cdot (-q\mu_p p \nabla \varphi - qD_p \nabla p) + qR$$

Basically the fully coupled Newton's method is an extension to differential operators of the well known Newton's method for the search of a zero for real functions ($f : \mathbb{R} \to \mathbb{R}$). In fact the vector function $\mathbf{F}$ is a nonlinear differential

operator and the associated problem which we intend to resolve is given a functional space V and the operator $\mathbf{F} : V \to V$, find $\mathbf{U} \in V$ such that (3.1) is satisfied. Just for the moment we don't care about the identity of V, which we will treat in details in the next section; now we are able to define the abstract Newton's Method for the iterative solution of problem (2.43):

Fully Coupled Newton's Method

Given $\mathbf{U}^0 \in V$, for all $k \geq 0$ until convergence, solve the following linearized problem:

$$\begin{aligned} \mathbf{F}'(\mathbf{U}^k)\delta\mathbf{U}^k &= -\mathbf{F}(\mathbf{U}^k) \\ \mathbf{U}^{k+1} &= \mathbf{U}^k + \delta\mathbf{U}^k \end{aligned} \qquad (3.3)$$

where $\mathbf{F}'$ is the Jacobian matrix of $\mathbf{F}$, whose (i,j)-th entry represents the Frechèt derivative of the i-th row with respect to the j-th variable.

The main advantage of this approach is without doubt the existence of the sequent theorem:

Theorem 3.1. *Let $\mathbf{U} \in V$ be a solution of problem (3.1). Assume that $\mathbf{F}'$ is Lipschitz continuos in the ball $\mathcal{B}(\mathbf{U}, \delta)$, i.e., that there exists $K > 0$ such that:*

$$||\mathbf{F}'(\mathbf{v}) - \mathbf{F}'(\mathbf{z})||_{L(V,V)} \leq K||\mathbf{v} - \mathbf{z}||_V \qquad \forall \mathbf{v}, \mathbf{z} \in \mathcal{B}(\mathbf{U}, \delta), \ \mathbf{v} \neq \mathbf{z} \qquad (3.4)$$

Then there exists in correspondence $\delta' > 0$, with $\delta' \leq \delta$, such that for all $\mathbf{U}^0 \in \mathcal{B}(\mathbf{U}, \delta')$ the sequence $\{\mathbf{U}^k\}$ generated by (3.3) converges quadratically to $\mathbf{U}$, i.e., there exists $C > 0$ such that, for a suitable $k_0 \geq 0$ we have:

$$||\mathbf{U} - \mathbf{U}^{k+1}||_V \leq C||\mathbf{U} - \mathbf{U}^k||_V^2 \qquad \forall k \geq k_0 \qquad (3.5)$$

Even the presence of this impressive result there are are several issues which must be evaluated before move toward this way of resolution:

- the jacobian matrix $\mathbf{F}'$ is often quite ill-conditioned and needs appropriate scaling and balancing in order to avoid problems associated with round-off error;

- to ensure convergence of the Newton iterative process (3.3), it is particularly important to ensure a very good initial guess for the unknown variables $\mathbf{U}$;

- if a direct solver is adopted (based on the LU factorization method), the number of loating point operations is of the order of N_{dofs}^3, where N_{dofs} is the number of degree of freedom used for the numerical approximation.

The fully coupled method is not the only possible approach, and many of the above points will be fixed adopting different methods, however theorem (3.1) guarantees the best error convergence estimation.

3.1.1 Gummel map algorithm

In 1964 H. K. Gummel proposed an orginal iterative algorithm in order to solve the system (2.43) in a semiconductor device in one spatial dimension. Today the Gummel decoupled iterative algorithm has become a milestone in contemporary device simulation in industrial software for the design and analysis of semiconductor devices.

The main idea of the algorithm is to move the nonlinearity to the Poisson equation only and once obtained the electric potential profile, both continuity equations are linearized. More precisely give an initial guess for φ, n and p, the functional iteration consists in the succesive solution of the nonlinear Poisson's equation (NLP) in a inner loop (Newton's method is applyed for this equation) and of the two linearized continuity equations (DD electrons and DD holes). A concisely scheme is presented in Fig.3.1.

Unfortunately there isn't any convergence result for this method like (3.1), although there are several advantages which make Gummel map algorithm preferable to the Fully Coupled Newton's Method. In fact simulations experience shows that the Gummel process is much more insesitive to the choice of the initial guess than Newton's method. This is particularly important in multidimensional problems where it is far from trivial to design a good starting point for initializing in a favorable manner.

Another attractive feature is the reduced computational and memory stage cost: at each iteration step, the Newton algorithm requires assembling a Jacobian matrix of size $3N_{dof} \times 3N_{dof}$, while the Gummel algorithm requires the successive solution of three problems, each one of size equal to $N_{dof} \times N_{dof}$.

FEMOS Gummel map

After this short presentation of the qualities and the drawbacks of the Gummel decoupled algorithm, we propose our functional iteration. For the sake of simplicity let us consider some useful hypotesis:
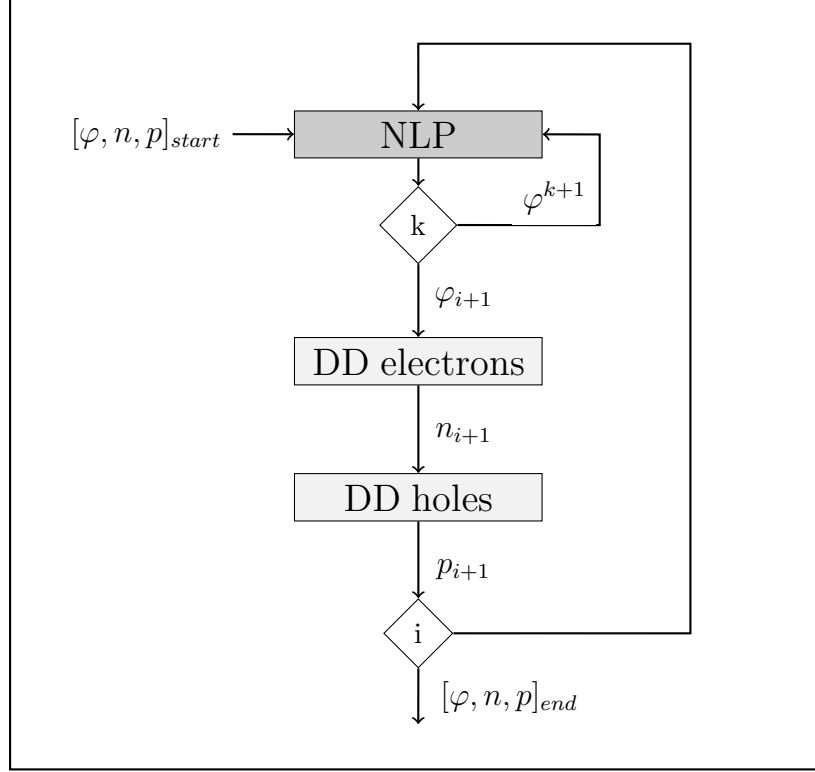
Figura 3.1: Gummel map algorithm

- we consider a polygonal simulation domain in $\mathbb{R}^3$, denoted with Ω; we indicate also with Ω_s the subset of Ω characterized by semiconductor material;

- the domain is formed only by semiconductor and/or oxide material (typically we consider silicon and SiO_2);

- contacts are assumed to be ideal, i.e. they are equipotential surfaces and no voltage drop occurs at the interface between the contact and the neighbouring material.

Take into account these considerations the appropriate Gummel map in our cases is the sequent:

> **FEMOS Gummel Map**
>
> Given n^0 and p^0, $\forall k$ until convergence:
>
> (Step 0) Compute φ_n^k and φ_p^k with (2.6) and (2.7).
>
> (Step 1) Solve the nonlinear Poisson equation over all the domain (NLP):
>
> $$\begin{cases} \nabla \cdot (-\epsilon \nabla \varphi) + n_i \left(exp \left(\dfrac{\varphi - \varphi_n^{(k)}}{V_{th}} \right) - exp \left(\dfrac{\varphi_p^{(k)} - \varphi}{V_{th}} \right) \right) = q(N_D^+ - N_A^-) & in \ \Omega_s \\ \nabla \cdot (-\epsilon \nabla \varphi) = 0 & in \ \Omega/\Omega_s \end{cases}$$
>
> Set $\varphi^k = \varphi$.
>
> (Step 2) Solve the Linear Electron Contintuity Equation (LEC):
>
> $$\nabla \cdot (-q\mu_n n \nabla \varphi^{(k)} + qD_n \nabla n) = qR(n^{k-1}, p^{k-1}) \quad in \ \Omega_s$$
>
> Set $n^k = n$.
>
> (Step 3) Solve the Linear Hole Contintuity Equation (LHC):
>
> $$\nabla \cdot (-q\mu_p p \nabla \varphi^{(k)} - qD_p \nabla p) = -qR(n^{k-1}, p^{k-1}) \quad in \ \Omega_s$$
>
> Set $p^k = p$.
>
> (Step 4) If the convergence criterion is satisfied break, otherwise restart from step 0.

Dobbiamo parlare delle slotboom variables e della possibilit di usare i quasi fermi????? It's interesting note that carrier densities are not the only useful variables. As we introduced in section (2.1.4), current density may be represented in many different ways. We underlie this because for example one can solve system (2.43) with Slotboom variables, without any change in the structure of the Gummel map.

3.2 Finite element discretization

In this section we describe the finite element discretization of the differential subproblems involved in the Gummel map previously introduced. Moreover

for each kind of PDE problem we give a briefly presentation of the well-posedness analysis.

3.2.1 Nonlinear Poisson Equation

Using the definition of the Frechèt derivative one can calculated the Jacobian matrix, which in this case has 1×1 dimensions. Therfore given φ^0, the Newton step for the linearized non linear Poisson equation looks as follows (carrier densities are computed with the Maxwell-Boltzmann approximation):

$$\begin{cases} \nabla \cdot (-\epsilon \nabla \delta \varphi^i) + \dfrac{1}{V_{th}} \sigma_s \delta \varphi^i = f_s & in \ \Omega_s \\ \nabla \cdot (-\epsilon \nabla \delta \varphi^i) = f_o & in \ \Omega/\Omega_s \\ \varphi^{i+1} = \varphi^i + \delta \varphi^i \end{cases} \tag{3.6}$$

having set,

$$\sigma_s(\varphi^i, \varphi_n^k, \varphi_p^k) = q(p(\varphi^i, \varphi_p^k) - n(\varphi^i, \varphi_n^k))$$
$$f_s(\varphi^i, \varphi_n^k, \varphi_p^k) = \nabla \cdot (-\epsilon \nabla \varphi^i) + q\left[p(\varphi^i, \varphi_p^k) - n(\varphi^i, \varphi_n^k)) + N_D^+ - N_A^-\right]$$
$$f_o(\varphi^i, \varphi_n^k, \varphi_p^k) = \nabla \cdot (-\epsilon \nabla \varphi^i)$$

The first two equations of problem (3.6) form a classical DR (Diffusion-Reaction) problem, respect the variable $\delta \varphi^i$ in Ω, considering a unique piecewise force term and reaction defined as follows:

$$f = f_s \mathcal{I}_{\Omega_s} + f_o \mathcal{I}_{\Omega/\Omega_s}$$
$$\sigma = \sigma_s \mathcal{I}_{\Omega_s}$$

The more generalized form of (3.6) reads as follows:

$$\begin{cases} \nabla \cdot (-\epsilon \nabla \varphi) + \sigma^{(k)}(\mathbf{x})\varphi &= f^{(k)}(\mathbf{x}) & in \ \Omega \\ \varphi &= \varphi_D & on \ \Gamma_D \\ \nabla \varphi \cdot \mathbf{n} &= 0 & on \ \Gamma_N \end{cases} \tag{3.7}$$

For the sake of simplicity we summerize the reaction and force term in σ and f, but we kept visible the iteration dependence. The well-posedness of such problem is ensured by several (and reasonable) hypotesis:

- $\epsilon \in L^\infty(\Omega)$ and $\exists m$ s.t. $0 < m \leq \epsilon$ (a.e.) in Ω;

- $\sigma \in L^\infty(\Omega)$ and $\exists m$ s.t. $0 < m \leq \sigma$ (a.e.) in Ω.

We proceed with the classical displacement weak formulation. Given $\varphi_D \in H^{1/2}(\Gamma_D)$ and $f \in L^2(\Omega)$ find $\varphi \in H^1(\Omega)$ such that

$$\int_\Omega \epsilon \nabla \varphi \nabla v \, d\Omega + \int_\Omega \sigma^{(k)} \varphi v \, d\Omega = \int_\Omega f^{(k)} v \, d\Omega \qquad \forall v \in H^1_{\Gamma_D}(\Omega) \qquad (3.8)$$

Numerical approximation

In view of the finite element discretization of problem, we let $\bar{\Omega} = \bigcup \bar{K}$ be a partition $\mathcal{T}_h$ of the domain into triangular elements K of area $|K|$. We suppose that $\mathcal{T}_h$ satisfied the *classical regularity condition* [Qua08]

Damping

The main problem associated with the classical Newton method is the tendency to overestimate the length of the actual correction step for the iterate. This phenomenon is frequently termed overshoot. In the case of the semiconductor equations this overshoot problem has often been treated by simply limiting the size of the correction vector $(\delta \varphi)$ determined by Newton's method. The usual established modifications to avoid overshoot are given by the seguent formulations:

$$A(\varphi_k) = \frac{1}{t_k} F'(\varphi_k) \qquad (3.9)$$

$$A(\varphi_k) = s_k I + F'(\varphi_k) \qquad (3.10)$$

t_k and s_k are properly chosen positive parameters. During the implementation of the code we chose (3.9) method. Note that for $t_k = 1$, $s_k = 0$ these modified Newton methods reduce to the classical Newton method. We have now to deal with the question how to choose t_k or s_k that the modified Newton methods exhibit superior convergence properties compared to the classical Newton method. For the case (3.9) there's a simple criterion suggested by Deuflhard referenza: t_k is taken from the interval $(0, 1]$ in such a manner that for any norm,

$$||F'(\varphi_k)^{-1} F(\varphi_k - t_k F'(\varphi_k)^{-1} F(\varphi_k))|| < ||F'(\varphi_k) F(\varphi_k)|| \qquad (3.11)$$

Condition (3.11) guarantees that the correction of the k-th iterate is an improved approximation to the final solution, in other words the residual norm can only descents. This condition can be easily evaluated only if the Jacobian matrix is factored into triangular matrices because the evaluation

of the argument of the norm on the left hand side of (3.11) is then reduced to a forward and backward substitution and the evaluation of $F(\varphi)$. Although we use an iterative methods (BCG solver) which implies serious diffuclties to the application of the criterion. Another valid possibility is to use the main diagonal of $F'(\varphi_k)$, denoted as $D(\varphi_k)$:

$$||D(\varphi_k)^{-1}F(\varphi_k - t_k D(\varphi_k)^{-1}F(\varphi_k))|| < ||F'(\varphi_k)F(\varphi_k)|| \qquad (3.12)$$

This is the criterion developed in our code. However the value to use for t_k is a question of trial and error. Frequently one chooses the following sequences:

$$t_k = \frac{1}{2^i} \qquad (3.13)$$

$$t_k = \frac{1}{\dfrac{i(i+1)}{2} \quad 2} \qquad (3.14)$$

obiuvsly i is the subiterations of damping reached when satisfied (3.12). Sufficiently close to the solution (3.11) (and so (3.12)) will be satisfied with $t_k = 1$ so that the convergence properties of the classical Newton method are recovered.

3.2.2 Continuity Equation

Weak formulation

Without loss of generality we consider only the electron continuity equation (similar reasoning could be make for the hole continuity equation). Problem referenza al problema is a classical diffusion-advection-reaction (DAR) problem written in conservative form. We will treat this PDE's equation likewise Poisson equation with the standard displacement weak formulation.

Be carefull about the right hand side: in the operation of many devices this term generates mass; this implies that a new reaction term is usually added in the left side of the equation:

$$R_n = \sigma n - f \qquad (3.15)$$

$$\begin{cases} \dfrac{\partial n}{\partial t} + \nabla \cdot (-D_n \nabla n) + \nabla \cdot (\mu_n \nabla \varphi^{(k)} n) + \sigma n &= f &\quad in \ \Omega \\ n &= n_D &\quad on \ \Gamma_D \\ \nabla n \cdot \mathbf{n} &= 0 &\quad on \ \Gamma_N \end{cases}$$
$$(3.16)$$

Given $n_D \in H^{1/2}(\Gamma_D)$ and $f \in L^2(\Omega)$ find $n \in H^1(\Omega)$ such that:

$$\begin{cases} \dfrac{\partial n}{\partial t} + \nabla \cdot (-D_n \nabla n) + \mu_n \nabla \varphi^{(k)} \nabla n + (\mu_n \Delta \varphi^{(k)} + \sigma)n &= f & \quad in \ \Omega \\ n &= n_D & \quad on \ \Gamma_D \\ \nabla n \cdot \mathbf{n} &= 0 & \quad on \ \Gamma_N \end{cases}$$

(3.17)

with $\Gamma_D \cap \Gamma_N = \varnothing$, $\Gamma_D \cup \Gamma_N = \partial\Omega$ and where $\mathbf{n}$ is the outward normal vector on Ω. We respect the standard hypotesi for the wellposdness of the problem:

- $D_n \in L^\infty(\Omega)$ and $\exists m$ s.t. $0 < m \le D_n$ a.e. in Ω;

- $\sigma \in L^\infty(\Omega)$ and $\exists m$ s.t. $0 < m \le \sigma$ a.e. in Ω;

- $\mu_n \nabla \varphi \in (W^{1,\infty}(\Omega))^d$.

formulazione con tempo?

Numerical approximation

In view of the finite element discretization of problem, we let $\bar\Omega = \bigcup \bar K$ be a partition $\mathcal{T}_h$ of the domain into triangular elements K of area $|K|$. We suppose that $\mathcal{T}_h$ satisfied the *classical regularity condition* [Qua08]

Partiamo con una semidiscretizzazione spaziale e poi trattiamo anche quella temporale?

Descrizione dettagliata (o meno?) del metodo implementato FVSG

3.2.3 Maximum discrete principle

Scriviamo qualcosa in merito?Quanto approfondito?

Capitolo 4

The current calculation problem

In many physical and engineering problems the real interesting variable of the conservation law is the flux in the domain or on specific surfaces and boundaries. The study of micro and nano electronics devices doesn't except this observation, in fact most of all models are oriented to obtain a satisfactory description of the current density. We know that the primal and not mixed formulation for the continuity equation doesn't resolved the flux density. The consequence of this fact is a binding post-processing of the quantities computed in order to reconstruct the current density of electrons and holes. It's evident which this part covers a lead role in the device simulation: as we are satisfied of the impressive results of the finite element scheme, it will be reather regrettable to lost the accuracy of our simulation during the computation of the current density. About this question many academics propose different solutions and the relative literature is boundless. Nevertheless the problem shows various aspects to take into account, among these there are some which every good method should be respect:

- reduced computational cost;

- easy extension to 3D simulation;

- detains some useful properties like orthogonal conservation across a generic surface of the domain;

- preserve consistency with the numerical scheme adopted.

It's not trivial ensure everyone of these points, thus move on toward a unique choice of a method is a delicate matter. Luckily there's some *main stone* which offers ever a good start point whence achieve new results. Probably

the most known and recognized by the inherent literature is the *Sharfetter-Gummel formula.*
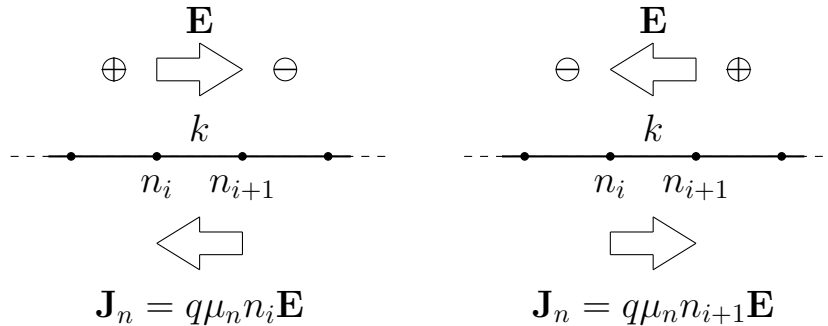
4.1 Scharfetter-Gummel formula

Consider the resolution of the continuity equation along a monodimensional domain. For the sake of simiplicity we contemplate a uniform partition (this hypotesis is not necessary for a more generic analysis). Moreover on every nodes is defined the electrostatic potential φ, and on every elements the relative electrostatic field $\mathbf{E}$. In order to avoid redundant considerations and calculuses, we proceed with our analysis considering only the current density of electrons $(\mathbf{J}_n)$.

In 1969 D. Scharfetter and H.K. Gummel (two scientists of Bell Labs), introduced a formula to compute the current density in this case, given φ and the density solution (n) on every nodes. This innovative approach led for the twenty years to follow every simulation which contemplates electric-devices.

We know that the constituve law is composed by a drift component, which depends on the electric field, and a diffusion component, which depends on the variation of the carrier density. Consider a generic element K, we define the drop in voltage $\Delta\varphi^k = \varphi_{i+1} - \varphi_i$. There are three possible situations which are well explain in the picture:

- $\Delta\varphi \gg 0$, mainly drift component from right to left

- $\Delta\varphi \ll 0$, mainly drift component from left to right

- $\Delta\varphi \simeq 0$, mainly diffusion component

$$\mathbf{J}_n = q\mu_n n_i \mathbf{E} \qquad \mathbf{J}_n = q\mu_n n_{i+1} \mathbf{E}$$

With the $Sharfetter-Gummel$ formula it's possibile taking into account every of these situations and solve boundary layer problems which occurs often in presence of strong drift component contribute.

$$J_n^k = q\frac{D_n}{h}\left[n_{i+1}\mathcal{B}\left(\frac{\Delta\varphi^k}{V_{th}}\right) - n_i\mathcal{B}\left(-\frac{\Delta\varphi^k}{V_{th}}\right)\right] \tag{4.1}$$

In the latter case $\Delta\varphi = 0$) the formula became:

$$J_n^k = qD_n\frac{n_{i+1} - n_i}{h} \tag{4.2}$$

which is the correct approximation of the current density using $\mathbb{P}_1$ basis.

Morover many mathematics discover important properties about this method questa parte la vorrei fare meglio. These changes in the direction of the electric field lead to boundary layer problems. It's not possibile afford these situations with a simply upwinding sheme ...

4.2   Extension for the 3D case

The extension of this formula for the 3D case is not trivial. We show the method for the computation of the current density of electrons (the extension for the current density of holes is quite similar). We remark the quasi fermi formula for current density:

$$\mathbf{J}_n = -q\mu_n n\nabla\varphi_n \tag{4.3}$$

where φ_n is the quasi fermi potential level. Let us write (4.3) in function of potential and in a canonic form:

$$\mathbf{J}_n\frac{exp\left(\dfrac{\varphi_n - \varphi}{V_{th}}\right)}{q\mu_n n_i} + \nabla\varphi_n = 0 \tag{4.4}$$

We consider $\mathbf{J}_n \in [L^2(\Omega)]^3$ and $\varphi_n, \varphi \in H^1(\Omega)$. We are able to multiply (4.4) with a generic function $\mathbf{q} \in [L^2(\Omega)]^3$ and then intagrate over the domain Ω:

$$\int_\Omega \frac{exp\left(\dfrac{\varphi_n - \varphi}{V_{th}}\right)}{q\mu_n n_i}\mathbf{J}_n \cdot \mathbf{q}\, d\Omega + \int_\Omega \nabla\varphi_n \cdot \mathbf{q}\, d\Omega = 0 \tag{4.5}$$

We proceed taking the usual discrete space of the constant elemenwise functions:

$$V_h = \left\{w \in L^2(\Omega) : w|_K \in \mathbb{P}_0 \forall K \in \tau_h\right\} \tag{4.6}$$

Now the discrete quantitaties are $\mathbf{J}_n^h \in [V_h]^3$ and $\nabla\varphi_n^h \in V_h$. We desire produce a system of equation on every elements for the three componente of $\mathbf{J}_n$, this is possible with a smart choice of the test function $\mathbf{q}_h \in [V_h]^3$:

$$\mathbf{q}_{1,2,3}^h = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{4.7}$$

From (4.5) we obtain the sequent system of equations defined for every element of the mesh:

$$\int_K \frac{exp\left(\dfrac{\varphi_n - \varphi}{V_{th}}\right)}{q\mu_n n_i} \mathbf{J}_n^h \cdot \mathbf{q}_i^h \, dK + \int_K \nabla\varphi_n^h \cdot \mathbf{q}_i^h \, dK = 0 \qquad \forall i = 1,2,3 \tag{4.8}$$

Operating the intagration we obtain the sequent formula for the current density components:

$$[\mathbf{J_n}]_i = -\mathbb{H}_K\left(q\mu_n n_i exp\left(\frac{\varphi - \varphi_n}{V_{th}}\right)\right) \frac{\partial \varphi_n^h}{\partial x_i} \quad i = 1...d \quad \forall K \in \tau_h \tag{4.9}$$

where $\mathbb{H}_K(f)$ is the armonic average on the elment K of the function f.

Although resolve the armonic average with a comlete 3D integration may be expensive in calculation time and propably not necessary. One approximation of this integral would be pass from a 3D integration to 1D integration along one edge of the element K.

$$\left(\frac{\int_K f^{-1} \, dK}{|K|}\right)^{-1} \simeq \left(\frac{\int_{e*} f^{-1} \, de}{|e^*|}\right)^{-1} \tag{4.10}$$

The approximation (4.10) is valid if we consider the correct edge.

Consider a quantity defined on the verteces:

$$\Phi := \varphi - \varphi_n \tag{4.11}$$

which is the difference between the electrostatic potential and the quasi fermi potential level. Now for every element consider two vertices: $\mathbf{x}_m$ s.t. $\Phi(\mathbf{x}_m) = \Phi_m := min_K(\Phi)$ and $\mathbf{x}_M$ s.t. $\Phi(\mathbf{x}_M) = \Phi_M := max_K(\Phi)$. Obviously it exists only one edge which connects these two points and on this one we perform the 1D integration (4.10). First of all as we reduce the dimension is feasible to represent $\sigma(\mathbf{x})$ in a easier mode as follows:

$$\sigma_n(s) = q\mu_n n_i exp\left(\Phi_m + (\Phi_M - \Phi_m)\frac{s - s_m}{|e^*|}\right) \qquad (4.12)$$

where $s \in [s_m, s_M]$ is the parameter refered to the edge e^* s.t. $\sigma_n(s_m) = \sigma_n(\mathbf{x}_m)$ and $\sigma_n(s_M) = \sigma_n(\mathbf{x}_M)$. We can easily resolve (4.10) with the substitution of variable:

$$\eta := \frac{s - s_m}{|e^*|}$$

this lead us to the sequent steps of integration:

$$\int_{e^*} \sigma_n^{-1} de = |e^*| \int_0^1 \frac{exp\left(-\Phi_m - (\Phi_M - \Phi_m)\eta\right)}{q\mu_n n_i} d\eta$$
$$= |e^*| \frac{exp(-\Phi_m)}{q\mu_n n_i} \frac{exp(\Phi_m - \Phi_M) - 1}{\Phi_m - \Phi_M}$$
$$= |e^*| \frac{exp(-\Phi_m)}{q\mu_n n_i} \frac{1}{\mathbf{B}(\Phi_m - \Phi_M)}$$

finally we obtain:

$$\int_K \sigma_n^{-1} dK \simeq q\mu_n n_i exp(\Phi_m)\mathbf{B}(\Phi_m - \Phi_M) \qquad (4.13)$$

Similar results may be obtained repeating the integration and considering s_M as start point:

$$\int_K \sigma_n^{-1} dK \simeq q\mu_n n_i exp(\Phi_M)\mathbf{B}(\Phi_M - \Phi_m) \qquad (4.14)$$

Numerical results (qua sarebbe carino fare un po' di test con una parte o l'altra della formula per mettere in crisi) shows that the best choice is e linea combination of these approzimations as follows:

$$\mathbf{J_n}^K = -q\mu_n \left[\frac{n_{min}\mathbf{B}(-\Delta\Phi_{max}) + n_{max}\mathbf{B}(\Delta\Phi_{max})}{2}\right] \nabla\varphi_n^h \qquad (4.15)$$

This approach is the natrual extension of the $Sharfetter - Gummel$ formula for the 1D case, indeed it's possible demonstrate the equivalence assuming a monodimensionale domain.

4.3 Residue Method

Contact method

In the following we present an accurate method for the evaluation of boundary integrals in semiconductor device based on the work . It's well known that the evaluation of boundary integrals is a difficult task occuring routinely in electron device simulations. In general, given a contact Γ_i, fluxes of current density to be calculated assume the following form:

$$\mathcal{I}_i^\nu = \int_{\Gamma_i} \mathbf{J}_\nu(\nu) \cdot \mathbf{n} \, d\Gamma_i \qquad \nu = \{n, p\} \tag{4.16}$$

where as usual $\mathbf{n}$ is the unit outward normal of the domain boundary. Difficulties in the numerical evalutaion of (4.16) arise from singularities in spatial derivatives of the approximate solution n^h or p^h near the contact edges, due to a change in the boundary condition type (from Dirichlet to Neumann) at the contact ends. In this work we extend the residue method on the 3D case and we confirm the optimal results obtained previously, matching them with SDEVICE (come bisogna scirvere il software che politica usare?). Moreover we remark that the method can be succesfully applied to a wide spread of applications, including contact charges, carrier quantum probability fluxes and heat fluxes.

Before go further with the presentation of the results, it's useful look up the anlysis made in Citazione and adapt it to our case. Before applying boundary conditions, the discretized form of referenza equazione reads:

$$\sum_{j\in\eta} A_{ij}\nu_j\psi_j(x_j) = b_i \qquad \forall i \in \eta \tag{4.17}$$

in FEMOS simulations A_{ij} is the K_{FVSG} global matrix but it's possibile use every discretization scheme. We can split the set of total nodes in contact node η_g and the complementary part η_n. The values of ν_j are known on the contacts and (4.17) can be rewritten as follows:

$$\begin{cases} \sum_{j\in\eta_n} A_{ij}\nu_j\psi_j(x_j) = b_i - \sum_{j\in\eta_g} A_{ij}\nu_j\psi_j(x_j) & \forall i \in \eta_n \\[2ex] \sum_{j\in\eta_n} A_{ij}\nu_j\psi_j(x_j) = b_i - \sum_{j\in\eta_g} A_{ij}\nu_j\psi_j(x_j) & \forall i \in \eta_g \end{cases} \tag{4.18}$$

The first set of equations is then solved for the problem unknowns ν_j (carrier density) while the second can be used for boundary flux estimation
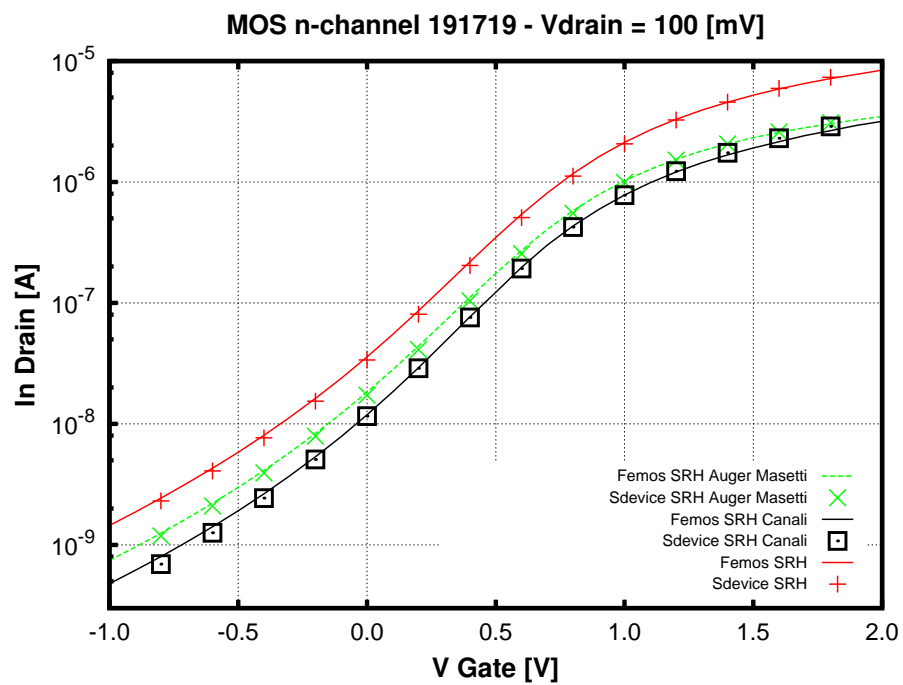
as described in the folowing. Consider a different test function v_i^h defined as:

$$v_i^h = \sum_{j \in \eta_{gi}} \psi_j \tag{4.19}$$

where η_{gi} is the set of nedes lying on contact Γ_i. We can rewrite (4.16) as follows:

$$
\begin{aligned}
\mathcal{I}_i^{\nu} &= \int_{\Gamma_i} \mathbf{J}_{\nu}(\nu) \cdot \mathbf{n} \, d\Gamma_i = \sum_{j=1}^{n_d} \int_{\Gamma_j} \mathbf{J}_{\nu}(\nu) \cdot \mathbf{n} \, v_i^h \, d\Gamma_j \\
&= \sum_{i \in \eta_{gi}} \sum_{j=1}^{n_d} \int_{\Gamma_j} \mathbf{J}_{\nu}(\nu) \cdot \mathbf{n} \, \psi_i \, d\Gamma_j = \sum_{i \in \eta_{gi}} \int_{\partial\Omega} \mathbf{J}_{\nu}(\nu) \cdot \mathbf{n} \, \psi_i \, d\partial\Omega \\
&= \sum_{i \in \eta_{gi}} \left[\int_{\Omega} \nabla \cdot \mathbf{J}_{\nu}(\nu) \psi_i \, d\Omega + \int_{\Omega} \mathbf{J}_{\nu}(\nu) \cdot \nabla \psi_i \, d\Omega \right] \\
&= \sum_{m \in \eta_{gi}} \left[\sum_{j \in \eta} A_{ij} \nu_j \psi_i - b_i \right]
\end{aligned}
\tag{4.20}
$$

Results

MOS n-channel 191719 - Vdrain = 100 [mV]

Bulk method

Now the question is: it's possible extend the residue technique at the calculation of current inside the domain? The answer at this question is not trivial. A good start point is the work of J.R. Hughes and Larson in the article The Continuous Galerikin Method Is Locally Conservative.The aim of their work is well exposed in the title and the conclusion is that this important property (locally conservation) is ensured less than the existence of an **auxiliary flux** denoted as $H(\omega)$ (with $\omega \subseteq \Omega$). To extract the statement of global and local conservation from the variational formulation, thay need to be able to set the weighting function to one. Obviously this is possible if $\Gamma_g = \varnothing$. In order to include every case an extended weak formulation is presented: consider η the set of all verteces of the discrete domain and η_g the subset of verteces on Dirichlet boundaries. Now we are able to define the discrete space $\mathcal{V}^h := span\{\psi_i\}_{i \in \eta - \eta_g}$ and $V^h := \mathcal{V}^h \oplus span\{\psi_i\}_{i \in \eta_g}$, where ψ_i is the basis function assiociated with the node i. The last space is the completion of the usual finite element space ($\mathcal{V}^h$). Note that the constant function having value 1 is contained in V^h. The modified form of Galerkin's method is given by:

Find $u^h \in \mathcal{S}^h$ and $H^h(\Omega) \in V^h - \mathcal{V}^h$ such that

$$(W^h, H^h(\Omega))_{\Gamma_g} = B(W^h, u^h) - L(W^h) \qquad \forall W^h \in V^h \qquad (4.21)$$

Note that (4.21) splits into two subproblems:

$$0 = B(w^h, u^h) - L(w^h) \qquad \forall w^h \in \mathcal{V}^h \qquad (4.22)$$

$$(W^h, H^h(\Omega))_{\Gamma_g} = B(W^h, u^h) - L(W^h) \qquad \forall W^h \in V^h - \mathcal{V}^h \qquad (4.23)$$

Equation (4.23) is a problem which determines $H^h(\Omega)$. In it we assume u^h is already determined by (4.22). The coefficient matrix for (4.23) is the mass matrix associated with Γ_g

$$\sum_{j \in \eta_g} (\psi_i, \psi_j) H_j^h(\Omega) = B(\psi_i, u^h) - L(\psi_i) \qquad \forall i \in \eta_g \qquad (4.24)$$

Il prodotto sul bordo genera la medesima matrice di massa?E cosa succede nei punti che non sono di Dirichlet?Non pu essere formulato in questo modo il problema?

In ogni caso noi abbiamo provato ad estendere il secondo problema ad ogni elemento interno che intersechi o meno i bordi di Dirichlet. Questo a cosa corrsiponde?

Abbiamo interpretato la costruzione per elemento come se in ogni elemento il problema avesse delle condizioni di dirichlet che si affacciano con

gli altri elementi adiacenti. Se assumiamo che questa procedura sia corretta allora penso che i successivi ragionamenti stiano in piedi.

Sempre nell'articolo viene testata la formulazione estesa contro la funzione test costante e si ottiene la seguente equivalenza:

$$\int_{\Gamma_g} H^h(\Omega)\, d\Gamma = \int_{\Gamma_h^+} (a_n u^h - h^+)\, d\Gamma - \int_{\Gamma_h^-} h^-\, d\Gamma - \int_{\Omega} f\, d\Omega \qquad (4.25)$$

avendo assunto ogni nodo come Dirichlet allora valgono le seguenti affermazioni:

- $\Gamma_g = \partial\Omega$

- $\Gamma_h = \varnothing$

quindi possiamo affermare che per ogni elemento vale:

$$\int_{\partial\Omega} H^h(\Omega)\, d\Gamma = -\int_{\Omega} f\, d\Omega \qquad (4.26)$$

Ora vorrei ripartire dal problema di partenza:

$$
\begin{aligned}
-\nabla \cdot \mathbf{J} &= f \\[6pt]
\int_{\Omega} -\nabla \cdot \mathbf{J}\Psi_k\, d\Omega &= \int_{\Omega} f\Psi_k\, d\Omega \qquad \forall k = 1...N_{elements} \\[6pt]
\int_{K} -\nabla \cdot \mathbf{J}\, d\Omega &= \int_{K} f\, d\Omega \qquad \forall k = 1...N_{elements} \\[6pt]
\int_{\partial K} -\mathbf{J} \cdot \mathbf{n}\, d\Omega &= \int_{K} f\, d\Omega \qquad \forall k = 1...N_{elements}
\end{aligned}
\qquad (4.27)
$$

La domanda quindi è possiamo in qualche modo dire che:

$$H^h(K) = \mathbf{J} \cdot \mathbf{n} \qquad (4.28)$$

Entrambe le quantità sono definite sul bordo dell'elemento, inoltre la ricostruzione dalle componenti normali al vettore densità di corrente non è impossibile. Tuttavia occorre passare prima a definire la grandezza sulle facce del tetraedro:

- Calcolo delle quantità nodali $H_j^h(\Omega)$

- Redistribuzione dei valori sulle facce (metodo percentuale proporzionalmente alle aree) $\bar{H}_j^h(\Omega)$

- Ortogonalizzazione dei contributi normali alle facce $\bar{H}_j^h(\Omega)\mathbf{n}_j$ tramite procedura alla Grand-Shmidt $H_j^*\mathbf{n}_j^*$

- Calcolo del vettore densità di corrente nell'elemento $\mathbf{J} = \sum_{i=1}^{4} H_j^*\mathbf{n}_j^*$

4.4 The Residue Method

4.4.1 Idea

Nel calcolo della corrente ai contatti si calcola sostanzialmente il residuo del problema DD globale, utilizzando però la matrice di sistema non ancora modificata per le condizioni al bordo di Dirichlet. Infine dato un contatto (e quindi un insieme di vertici), si sommano le componenti del residuo relative ai vertici che risiedono su quel contatto. In questo modo otteniamo la corrente di elettroni $\mathcal{I}_n^k$ (k-esimo contatto). Possiamo affermare ovviamente che:

$$\mathcal{I}_n^k = \int_{\Gamma_k} \mathbf{J}_n \cdot \mathbf{n} \, d\Gamma_k \tag{4.29}$$

Poniamo di aver risolto il problema e di conoscere le densità su ogni vertice della triangolazione.**L'idea principale è di pensare ad ogni singolo elemento come un nuovo problema con condizioni di Dirichlet sui quattro vertici.** Applico nuovamente l'idea del calcolo della corrente ai contatti utilizzata per l'intero dominio di simulazione, ma questa volta ogni faccia del tetraedro costituisce un contatto. Quindi calcolo la matrice locale e la forzante locale del problema ed infine computo il residuo con la soluzione che già possiedo. Ora non mi resta che per ogni faccia (contatto) calcolare la corrente.

Assumendo $\mathbf{J}_n$ nel discreto essa è una quantità definita sugli elementi, dunque sarà costante nel dominio che stiamo considerando in questo momento, possiamo affermare che data una faccia k dell'elemento:

$$\mathcal{I}_n^k = \mathbf{J}_n \cdot \mathbf{n}_k |\Gamma_k| \tag{4.30}$$

Così facendo è possibile recuperare quattro contributi alla corrente dell'elemento che tramite un'ortogonalizzazione alla Grand-Schmidt ci permettono di ricostruire il vettore densità di corrente.

$$
\begin{aligned}
\mathbf{J}_n \;=\; & \mathcal{I}_n^1 \mathbf{n}_1 + (\mathcal{I}_n^2 \mathbf{n}_2 - \mathcal{I}_n^2 (\mathbf{n}_2 \cdot \mathbf{n}_1)\mathbf{n}_1) \\
& + (\mathcal{I}_n^3 \mathbf{n}_3 - \mathcal{I}_n^3 (\mathbf{n}_3 \cdot \mathbf{v}_2)\mathbf{v}_2) + (\mathcal{I}_n^3 \mathbf{n}_3 - \mathcal{I}_n^3 (\mathbf{n}_3 \cdot \mathbf{n}_1)\mathbf{n}_1) \\
& + (\mathcal{I}_n^4 \mathbf{n}_4 - \mathcal{I}_n^4 (\mathbf{n}_4 \cdot \mathbf{v}_3)\mathbf{v}_3) + (\mathcal{I}_n^4 \mathbf{n}_4 - \mathcal{I}_n^4 (\mathbf{n}_4 \cdot \mathbf{v}_2)\mathbf{v}_2) + (\mathcal{I}_n^4 \mathbf{n}_4 - \mathcal{I}_n^4 (\mathbf{n}_4 \cdot \mathbf{n}_1)\mathbf{n}_1)
\end{aligned}
$$

Mi sembra una possibile estensione del metodo dei residui all'interno del dispositivo anche se probabilmente potrei avere commesso degli errori nello sviluppo dell'idea. Tuttavia anche nel caso di una corretta base teorica, per come ho presentato il metodo mi sembra che si rischi di incorrere in problemi numerici date le numerosi differenze che occorrerebbe fare.

Probabilmente aiutandoci con gli articoli di Hughes possiamo tirar fuori un'idea simile, sfruttando in qualche modo i flussi ausiliari che sono senza dubbio legati ai residui dei problemi locali.

Capitolo 5

Simulation results

5.1   Case tests presentation

In this section we dedicate some words to give informations about what kind of input file FEMOS get in input and what files it produces. Basically all .tdr format are accepted by FEMOS code. For our simulation it was possible used synopsis tools. More precisely the device design was performed with SDE and the mesh generator used was SNMESH. About output we produce .xmf files easly visible with paraview.

We focus our tests over three main structures: diode, diode inside an oxide box and a transistor mos. Here we presents the geometry tested and some mesh pattern used.

(a)   (b)   (c)
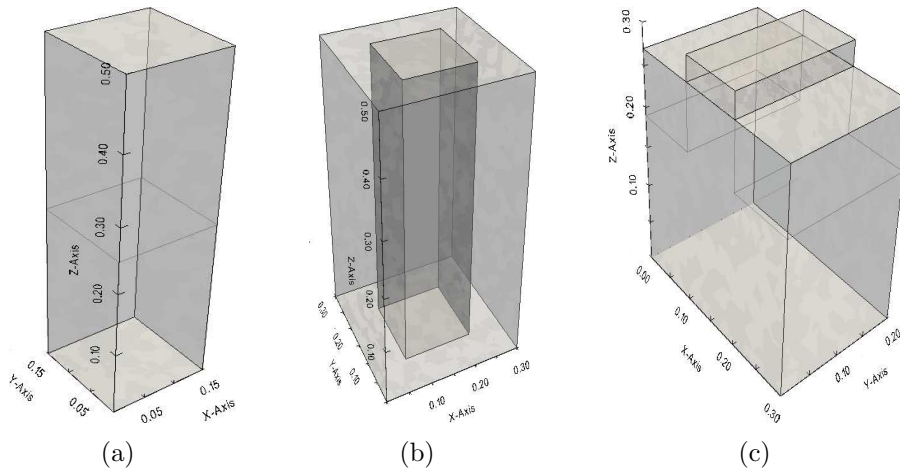
For everyone of these case test we investigated different doping concentration and distribution (like abrupt or gauss profile), multiple generation/-

recombination and mobility models applyed, ... Results are always checked with the SDEVICE simulation ...

Bibliografia

[Qua08] Alfio Quarteroni. *Modellistica Numerica per Problemi Differenziali.* Springer Italia, Milan, 2008.

[YT09] Taur Yuan and H. Ning Tak. *Fundamentals of Modern VLSI Devices.* Cambridge University Press, 2009.