THEORETICAL RESULTS IN INVERSE PROBLEMS FOR SIZE, SOLVABILITY, AND UNIQUENESS IN THE P-N JUNCTION AND DOPING PROFILE OF SEMICONDUCTORS

A Thesis by

Joseph K. Myers

Bachelor of Science, Wichita State University, 2005

Submitted to the Department of Mathematics and Statistics

and the faculty of the Graduate School of

Wichita State University

in partial fulfillment of

the requirements for the degree of

Master of Science

May 2006

THEORETICAL RESULTS IN INVERSE PROBLEMS FOR SIZE, SOLVABILITY, AND UNIQUENESS IN THE P-N JUNCTION AND DOPING PROFILE OF SEMICONDUCTORS

I have examined the final copy of this Thesis for form and content and recommend that it be accepted in partial fulfillment of the requirement for the degree of Master of Science with a major in Mathematics.

_____
Victor Isakov, Committee Chair

We have read this Thesis
and recommend its acceptance:

_____
Alexandre Boukhgueim, Committee Member

_____
Steven Skinner, Committee Member

_____
Ziqi Sun, Committee Member

ACKNOWLEDGMENTS

The author gives thanks to God, his Creator. The author also gratefully acknowledges the assistance received from Dr. Isakov, Dr. Skinner, Dr. Boukhgueim, and Dr. Sun in preparing this paper.

ABSTRACT

We present an overview of mathematical models for electrons and holes in semiconductors. We use these to pose some inverse problems for determining the doping profile of a semiconductor. We establish the solvability of the equilibrium equation $\lambda^2 \triangle u = e^u - e^{-u} - C$ in Ω. We also obtain information about the conductivity coefficient in the important case when it is piecewise constant and discontinuous.

TABLE OF CONTENTS

1 Introduction

Semiconductors are essential components of computers and many electrical devices. The junction between P-type and N-type regions of the semiconductor is important to the quality of the semiconductor device. It is very important for quality control in semiconductor manufacturing to determine how closely the location and size of this *P-N junction* adhere to specifications.

The manufacturing process of a semiconductor device begins with fabrication of a high-grade silicon wafer which has been transformed into a crystal structure by a chemical process. The wafer is further prepared by oxidation and epitaxy, a process which forms a monocrystalline film on top of the silicon surface.

In order to obtain the desired electrical properties in the semiconductor device, the concentration of carriers in the semiconductor must be regioselectively altered. Doing this involves "doping" the semiconductor by the addition of impurities with special properties.

Addition of these dopants is done through a process of diffusion [11], which forms a doped layer within the silicon substrate (Figure 1). The diffusion is carried out in a masked high-temperature process in which the conductivity of the substrate is changed over a region determined by the diffusion depth of the dopants. During diffusion the temperature range is 1173 to 1473 Kelvin, with the silicon wafer exposed to a diffusant containing the desired impurity.

There may be several doping steps included in the preparation of a semiconductor device. The doped regions are either P-type or N-type depending on their respective dopants. As the secondary set of dopants diffuses into the silicon, a P-N junction [26] is formed at the border between the +5 and +3 doped areas of the semiconductor. However, as soon as the semiconductor device is produced, it is only possible to access the bottom or the top of the semiconductor, where voltage can be applied and current measured.

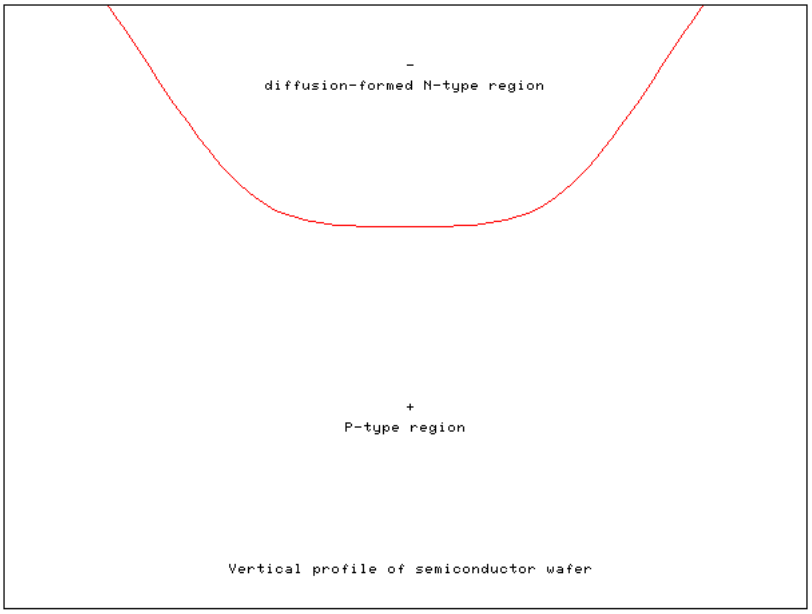

Figure 1: Drift diffusion represented within semiconductor.

The desired semiconductor devices are obtained by this method because the electrical properties of a semiconductor undergo predictable alterations upon addition of a dopant, depending on the specific impurities the dopant contains.

Following the preparation by doping, electrons and "electron holes" are distributed throughout the semiconductor according to the doping profile of the semiconductor. P-type doping results in a relative abundance of free holes, while N-type doping results in a relative abundance of free electrons. Common use, such as in the titling of *Electrons and Holes in Semiconductors* [22], justifies the use of the term holes although holes do not physically exist. The word is helpful for identifying the position vacated by an electron as a result of the electron's movement, leaving a positive charge at that position.

The goal of the inverse problem is to determine the success of the manufacturing process using electrical measurements. Using the measurements and solving the inverse problem, it is possible for one to find the extent of diffusion. For instance, one

might detect incorrectly doped computer chips and prevent hardware failures caused by distribution of such chips.

A way to accomplish this goal is by determining the concentration of electrons and holes throughout the semiconductor—which is equivalent to finding the doping profile—or at least to determine the location and size of the so-called P-N junction which separates the P-type and N-type regions of the semiconductor. Since there is restricted access to the semiconductor, it is most desirable to accomplish this by means of boundary electrical measurements.

There are mathematical models for the doping profile. However, the equations are very complex and no systematic mathematical theory has been developed to deal with the inverse problem of finding the doping profile.

Out of desperation, least-squares methods are presently being used to determine a few parameters about the doping profile. In this paper we will suggest a means of recovering information about the doping profile by using more theoretical methods.

The basis for our accomplishments is the well-established theory of elliptic partial differential equations. To our knowledge, the findings in Section 4.3 (the adjoint inverse problem) have never been published, and the results in Section 5 (evaluating the size of the N-type region) are new to this paper.

Problems related to ours have undergone extensive study. For determining the doping profile, the most advanced techniques that we are aware of are the least-squares type methods given in the papers [7], [8].

Let us summarize the major results of others which we have used, and specify our contributions toward the goal of determining the doping profile.

In this paper we rely on the drift diffusion system of equations which was developed by W. van Roosbroeck of Bell Telephone Laboratories in 1953. Then we use the semiconductor equations written in the book of P. A. Markowich, et. al. [20]. The formulation of the inverse problems we try to solve is due to V. Isakov [14], who also

referred to the work of Markowich in his own work. In particular, Dr. Isakov provided us with the adjoint inverse problem and provided the stability estimates for finding the data of the adjoint problem. Entering the adjoint inverse problem simplifies the boundary data. This makes it easier to obtain information from the elliptic partial differential equation that describes the interior of the semiconductor.

The existence of equilibrium state was established using [19] and a technical lemma from [20, Lemma 3.3.14].

We ourselves describe and explain each physical and mathematical part of the problem as we develop it in this paper. We have attempted to tie together each related aspect from physics or mathematics to form a logical structure that delineates the current knowledge and new developments. We hope that the reader who is experienced in either of the related fields of physics or mathematics will be able to understand the material which is presented.

We are responsible for the entire content of Section 5, with the exception of Section 5.2.2, for which we referred to [20]. In the main content of Section 5 we form a partial differential equation (PDE), and we obtain a small set of conditions for a category of "test functions." By inserting one of these test functions into the PDE, we are able to obtain an integral equation that allows us to evaluate characteristics of the doping profile or the P-N junction.

2 Models of semiconductors

There are physical and mathematical reasons for the way the inverse conductivity problem is formulated and solved. In trying to discover the conductivity coefficient, we are seeking a function of two variables from data measured in only one variable. This is clearly impossible, but we will find there are techniques which will obtain information about the doping profile, anyway. We are hopeful that the information will be sufficient for the semiconductor industry to incorporate new methods of quality control.

In this section we will describe an appropriate set of the physical properties that motivate the mathematical systems that were developed and which are the foundation of our research. Then we believe it is helpful for us to touch on a hydrodynamical model for semiconductors which is useful on a smaller scale than which we are dealing. This model may become more and more important as electronics are increasingly miniaturized. Finally, we will introduce the mathematical model of semiconductors based on the drift diffusion equations with which we will work in subsequent sections.

2.1 Fundamental physical models

The electromagnetic field in the semiconductor may be described appropriately by using the quantities whose behavior is described by the classical Maxwell equations and Boltzmann equation. In the case of semiconductors, these quantities are given by electrostatic potential, concentrations of charges, current density, diffusion coefficients, permittivity, recombination-generation rate, charge mobility, mass density, specific heat, thermal conductivity, and heat source. All of this information is obtained from basic knowledge about the physical properties of semiconductors.

Other appropriate models are available in addition to the one we shall describe. The hydrodynamical models for semiconductors can be derived from the Boltzmann

equation by applying the moment method. For simplicity, β, the Brillouin zone related to the crystal lattice [2], may be taken $\beta = \Re^3$ [20]. This simplification assumes that a continuous representation of the discrete lattice is valid. This assumption implicitly requires the contacts to be far away from the junction in terms of diffusion lengths. As the scale of electronics is reduced, such an assumption may be unjustified.

The motivation for our research will be clarified if we now briefly introduce some physical aspects to the structure of semiconductors. Then we will finish describing the hydrodynamical model and begin comparing it to the model for semiconductors which we will use.

2.2 Crystal lattices

The physical properties of a semiconductor arise from the molecular configuration and crystal lattice in the semiconductor material. We have already noted the influence of one of the features in the crystal lattice, so let us explain the concept of "Brillouin zones." We will see that they have an important effect, not only on the electron concentration, but also on the optical properties of materials.

Brillouin zones are defined recursively. The first Brillouin zone is defined to be the Wigner-Seitz primitive cell of the reciprocal lattice, or equivalently, the set of points in $\mathbf{k}$ space that can be reached from the origin without crossing any Bragg plane. (Here a Bragg "plane" for two points in a lattice is the plane which is perpendicular to the line between the two points and passes through the bisector of that line.) The nth Brillouin zone of a crystal lattice is the set of points that can be reached from the origin by crossing $n-1$ Bragg planes, but no fewer. [6]

Any vector between reciprocal lattice points is perpendicular to some lattice plane in the crystal lattice. Note, for instance, that the body centered cubic lattice is the reciprocal lattice to the face centered cubic lattice. This follows from the following simple considerations.

The primitive translation vectors of the bcc lattice are

$$a' = \tfrac{1}{2}a(x + y - z)$$
$$b' = \tfrac{1}{2}a(-x + y + z)$$
$$c' = \tfrac{1}{2}a(x - y + z)$$

where a is the side of the conventional unit cube and x, y, z are orthogonal unit vectors parallel to the cube edges. The volume of the primitive cell is

$$V = |a' \cdot b' \times c'| = \frac{1}{2}a^3$$

The primitive translations A, B, C of the reciprocal lattice are defined by

$$A = 2\pi \tfrac{b \times c}{a \cdot b \times c}$$
$$B = 2\pi \tfrac{c \times a}{a \cdot b \times c}$$
$$C = 2\pi \tfrac{a \times b}{a \cdot b \times c}$$

We have

$$A = \tfrac{2\pi}{a}(x + y)$$
$$B = \tfrac{2\pi}{a}(y + z)$$
$$C = \tfrac{2\pi}{a}(x + z)$$

But these are just the primitive vectors of an fcc lattice. Thus, the fcc lattice is the reciprocal lattice of the bcc lattice.

If h, k, l are integers, the general reciprocal lattice vector is

$$G = hA + kB + lC = \frac{2\pi}{a}[(h+l)x + (h+k)y + (k+l)z]$$

We show reproductions of the first three Brillouin zones of a face centered cubic lattice.

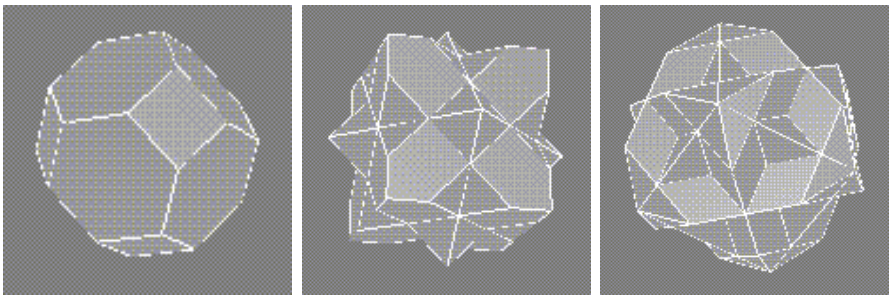

Figure 2: The first three Brillouin zones of an fcc lattice.

The electron concentration [17] $n(\rho)$ in a crystal may be expressed as a Fourier series

$$n(\rho) = \sum_K n_K e^{iK\cdot\rho} \tag{1}$$

where ρ is any point in the crystal. Following [17], let us prove an important theorem.

Theorem 2.1. *For any function that has the translational periodicity of the lattice, the only values of K that appear in the corresponding Fourier series are the reciprocal lattice vectors $G = hA + kB + lC$.*

Proof. We form $n(\rho+\rho_{mnp})$, where $\rho_{mnp} = ma+nb+pc$ is a crystal lattice translation:

$$n(\rho + \rho_{mnp}) = \sum_K n_K e^{iK\cdot\rho} e^{iK\cdot\rho_{mnp}}$$

This will have the desired translational periodicity and be equal to $n(\rho)$ only if $K \cdot (ma + nb + pc)$ is an integral multiple of 2π. But this is precisely the form of the condition $G \cdot \rho_{mnp} = 2\pi(\text{integer})$ that we encountered for points G of a reciprocal lattice. Thus, (1) reduces to

$$n(\rho) = \sum_G n_G e^{iG\cdot\rho} \tag{2}$$

where $G = hA + kB + lC$ is any reciprocal lattice vector. The quantities h, k, l are integers by definition of the G's. □

The set of quantities n_G describe completely the distribution of electrons in the entire crystal, and they also describe the diffraction of x-rays. To show this connection, we combine the fact that the scattering amplitude is proportional to the integral

$$\int dV n(\rho) e^{-i\rho\cdot\triangle k}$$

and the statement in (2).

$$\int dV n(\rho) e^{-i\rho\cdot\triangle k} = \sum_G n_G \int dV e^{iG\cdot\rho} e^{-i\triangle k\cdot\rho}$$

This integral is equal to the volume V if $\triangle k$ is equal to a reciprocal lattice vector G, and the integral may be shown to equal zero [17] if $\triangle k$ is not equal to a reciprocal lattice vector. Therefore, n_G is a measure of the amplitude of the diffracted beam.

The diffraction of light is closely related to the electrical properties of the semiconductor. It has long been known that optical pumping with circularly polarized light can generate electrons with a certain spin orientation in direct-bandgap semi-

conductors. [15]

Light is affected by the crystal structure, and so is the distribution of electrons. The discrete structure of the crystal is modeled by continuous equations that try to represent the control exercised by the crystal structure on electron flow.

2.3 Hydrodynamical model of semiconductors

Let $f = f(x, k, t)$, where $x \in \Re^3$ denotes the position variable, $k \in \beta$ denotes the wave vector, and β the Brillouin zone related to the crystal lattice.

With time $t > 0$, we obtain from the Boltzmann equation the unipolar hydrodynamical system modeling semiconductors.

$$\partial_t f + v(k)\nabla_x f + \frac{q}{h}\nabla_x V \cdot \nabla_k f = Q(f, f),\ (x, k, t) \in \Re^3 \times \Re^3 \times \Re^+, \tag{3}$$

$$f(x, k, 0) = f_0(x, k),\ (x, k) \in \Re^3 \times \Re^3. \tag{4}$$

Here $q > 0$ is the elementary charge and h the reduced Planck constant; $v(k) = (1/h)\nabla_k \epsilon(k)$ is the mean electron velocity, and $\epsilon(k)$ is the energy-wave vector function.

The collision operator $Q(f, f)$ is supposed to model the short range interactions of electrons with crystal impurities and photons. Electron interaction is neglected in this model. The form of Q is

$$Q(f, f) = \int \psi(x, v, v')(Mf' - M'f)dv',$$

where $\psi(x, v, v')$, a symmetric function in v and v', is the scattering rate, and the Maxwellian is given by

$$M(v) = \left(\frac{m}{2\pi k_\beta T}\right)^{3/2} \exp\left(\frac{-m|v|^2}{2k_\beta T}\right).$$

The system is described more fully in [18].

The hydrodynamical model is used primarily to describe high field phenomena, contrasted with the assumptions of low carrier densities and small fields. This makes the model more appropriate for shorter scale and quantum representations when crystal lattice effects are taken into consideration.

It may seem surprising that a hydrodynamical model could arise in this situation. However, when the motion of partical ensembles is described with classical principles, a kinetic transport model is formed based on Newton's second law. In particular, the model is applied to ballistic transport and scattering events of the charge currencies. Approximating the solutions of this kinetic model by performing scaling limits, one obtains fluid dynamical models.

We emphasize the unipolar nature of the hydrodynamical model as a simplifying technique. We will soon work with another unipolar model where the improved accessibility of the data is obtained by entering the adjoint problem.

We refer to [7] for more information on the hydrodynamical model, and we believe that [22] is a good source for comprehensive information about the physics of semiconductors.

2.4 Drift diffusion equations

We feel that it is appropriate to provide more background detail to the mathematical systems which we will be using.

Our starting point is a commonly accepted model of the semiconductor device, the *drift diffusion equations*. We note that this system of equations is quite similar in its properties to the equations from Combustion Theory [3, Ch. 2]. Although the most widely accepted model for semiconductor devices, the drift diffusion equations already represent a compromise that is made between the ideal of accurately describing the underlying device physics, and the feasability of computational solutions for

the chosen nonlinear system of partial differential equations.

The dependence on time that is included in this system of equations should be emphasized.

$$\text{div}(\epsilon\nabla V) = q(n - p - C), \tag{5}$$

$$\text{div } J_n = q(\partial_t n + R), \tag{6}$$

$$\text{div } J_p = q(-\partial_t p - R), \tag{7}$$

$$J_n = q(D_n(E,T)\nabla n - \mu_n(E,T)n\nabla V), \tag{8}$$

$$J_p = q(-D_p(E,T)\nabla p - \mu_p(E,T)p\nabla V), \tag{9}$$

$$\rho c(T)\partial_t T - H = \text{div } k(T)\nabla T. \tag{10}$$

Here V denotes the electrostatic potential ($-\nabla V$ is the electric field, $E = |\nabla V|$), n and p are the concentration of free carriers of negative charges (electrons) and positive charge (holes), respectively. U is the applied potential. We let J_n and J_p to be the densities of the electron and the hole current, respectively. D_n and D_p are the diffusion coefficients for electrons and holes. Coefficients μ_n and μ_p represent the mobilities of electrons and holes. The positive constants ϵ and q denote the permittivity coefficient (e.g., for silicon) and the elementary charge. The function $R = R(n, p, x)$ denotes the recombination-generation rate. The function $C = C(x)$ represents the doping concentration, which is produced by diffusion of different materials into the silicon crystal and by implantation with an ion beam. When $C > 0$, it represents the P area of the semiconductor, and $C < 0$ the N area. The constants ρ and c represent the specific mass density and specific heat of the material. In the thermodynamic description, k and H denote the thermal conductivity and the locally generated heat.

This set of equations is considered in a domain $\Omega \subset \Re^2$ representing the semiconductor device. We decompose the boundary $\partial\Omega$ into nonempty disjoint parts:

$\partial\Omega = \Gamma_N \cup \Gamma_0 \cup \Gamma_1$. The boundary data result in a mixed problem. The Dirichlet parts of the boundary, Γ_0 and Γ_1, model the Ohmic contacts, where the potential V as well as the concentrations n and p are prescribed. The Neumann part Γ_N corresponds to the insulating surfaces; thus a zero current flow and a zero electric field in the normal direction are prescribed. See the literature [26, 7, 8, 20] for more details.

Complete understanding is lacking even for this system of elliptic parabolic quasilinear equations. Depending on the type of information desired about the semiconductor device, it is necessary to isolate a certain subset of the difficult problems which solid-state physics poses to the mathematician.

Due to the partial electrical resemblance between the P-N junction and an inductor, it is doubtful that detailed information about $C(x)$, the doping profile, can be determined from inverse problems neglecting the time dependence of the equations.

It is desirable to complete this brief exposition of the semiconductor equations by describing the boundary value data and initial conditions which are typically written for the system of equations, and by showing how our specific problem is extracted from the larger system.

Let $\Gamma_D = \Gamma_0 \cup \Gamma_1$, so that Γ_D, Γ_N is a partition of the boundary $\partial\Omega$ into open parts.

The boundary data are

$$V = V_0, n = n_0, p = p_0, T = T_0 \text{ on } \Gamma_D \times (0,T) \tag{11}$$

and

$$\partial_\nu V = 0, J_n \cdot \nu = J_p \cdot \nu = 0, \partial_\nu T = 0 \text{ on } \Gamma_N \times (0,T) \tag{12}$$

We will define the notation $\partial_\nu u, u \in C^1(\Omega)$ in (41).

One also imposes the initial conditions

$$n = n_0, p = p_0, T = T_0 \text{ on } \Omega \times \{0\} \tag{13}$$

Now we make the following series of simplifications, where U_T will be the thermal voltage.

1. Consider the steady-state case when the current flow is in equilibrium. By analyzing the steady-state case in this way we can achieve something. However, an analysis which keeps track of time data would allow more electrical effects to be recorded. From the possible inductive effects at the junction the time-dependent inverse problem seems to promise more information about the semiconductor, but is very difficult to solve.

2. Neglect thermal effects: then T = const.

3. Take advantage of known relations: $D_n = U_T \mu_n, D_p = U_T \mu_p$.

4. Introduce "Slotboom" variables u, v.

$$n = n_i e^{\frac{V}{U_T}} u, p = n e^{-\frac{V}{U_T}} v$$

5. Scale by $\lambda^2 = \frac{\epsilon}{qU_T}, \delta^2 = n_i$ (intrinsic density of electrons).

Then we obtain the elliptic quasilinear system

$$\lambda^2 \triangle V = \delta^2 (e^V u - e^{-V} v) - C \tag{14}$$

$$\text{div} J_n = \delta^4 Q(V, u, x, v)(uv - 1), J_n = \mu_n n_i e^V \nabla u, \tag{15}$$

$$\text{div} J_p = -\delta^4 Q(V, u, x, v)(uv - 1), J_p = -\mu_p n_i e^{-V} \nabla v \tag{16}$$

and the boundary conditions

$$V = U + V_{bi}, u = e^{-U}, v = e^{U} \text{ on } \Gamma_D, \text{ with } V_{bi}(x) = U_T \ln\left(\frac{n_D(x)}{n_i}\right) \tag{17}$$

$$\partial_\nu V = 0, J_n \cdot \nu = J_p \cdot \nu = 0 \text{ on } \Gamma_N \tag{18}$$

where $V_{bi} \in C^0(\Omega)$ is the bipolar potential, and $\lambda \approx 10^{-3} < \delta$.

All physical parameters are known except for $C \in L_\infty(\Omega)$.

We are now almost to the start of our problem. We make a note of the materials data for silicon at room temperature (about 298° Kelvin) which give $\epsilon = 10^{-10}, q = 10^{-19}, U_T = 2.510^{-2}, n_i = 10^{16}$.

An advantage of the stationary drift diffusion system in the form (14-16) is its possession of an equilibrium solution $U = 0, V = V_0, u = 1, v = 1$.

We will deal with this case by considering the solvability of the equation which necessarily must be satisfied by V_0.

$$\lambda^2 \triangle V_0 = \delta^2(e^{V_0} - e^{-V_0}) - C \text{ in } \Omega \tag{19}$$

Before we go to the linearization V_1, u_1, v_1 at the equilibrium, we must establish the solvability of this equation. The equation is extremely nonlinear, with a right hand side that is essentially $\sinh(V_0) - C(x)$.

Refer also to [14, 2.2-2.5].

3 Existence of a steady-state solution

To deal with this problem, we need to introduce some theoretical machinery of elliptic partial differential equations. Let us begin by referring to the embedding theorems, which are implicit in the topic of existence. We will use these for showing existence of a related, time-dependent PDE after we have proved solvability in our problem.

3.1 Sobolev spaces

The $W^{k,p}$ spaces are Banach spaces where continuous differentiability of order k is replaced by weak differentiability, and Hölder continuity by p-integrability. For $p \geq 1$ and k a nonnegative integer, we let

$$W^{k,p}(\Omega) = \{u \in W^k(\Omega) : D^\alpha u \in L^p(\Omega) \ \forall \ |\alpha| \leq k\}$$

A norm is introduced by defining

$$\|u\|_{k,p;\Omega} = \left(\int_\Omega \sum_{|\alpha| \leq k} |D^\alpha u|^p \right)^{1/p}$$

Theorem 3.1. *(Embedding). Let Ω be a bounded Lipschitz domain in $\Re^n$; then*

$$W_0^{1,p}(\Omega) \subset \begin{cases} L^{np/(n-p)}(\Omega) & \text{for} \quad p < n \\ C^0(\overline{\Omega}) & \text{for} \quad p > n \end{cases}$$

and there is a constant $C(p, q, \lambda)$ such that for all functions $u \in H_{k,p}(\Omega)$ we have

$$\|u\|_q(\Omega) \leq C\|u\|_{k,p}(\Omega)$$

$$\|u\|_{m,q}(\Omega) \leq C\|u\|_{k,p}(\Omega)$$

$$|u|_\lambda(\Omega) \leq C\|u\|_{k,p}(\Omega)$$

$$when\ q \leq np/(n-kp), n > kp,$$

$$when\ m \leq k, p \leq q, n(1/p - 1/q) \leq k - m,$$

$$when\ \lambda \leq k - n/p, n < kp.$$

3.2 Geometry of the semiconductor

The geometry of the semiconductor and the approximate doping profile is shown in the diagram of the two-dimensional domain Ω, and we apply this correspondence to our own problem. One may notice that this figure is obtained by cutting the doping profile in Figure 1 along its axis of symmetry.

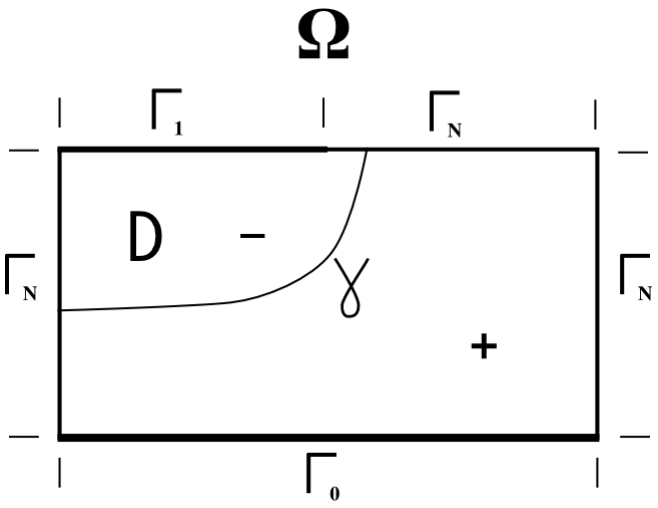

Figure 3: Vertical profile of a semiconductor.

The rectangular domain Ω is insulated except for the edge Γ_0 where input voltage may be prescribed, and Γ_1 where the voltage measurements are made. While Γ_0 extends the full length of the rectangle, Γ_1 is restricted to only a portion of the upper edge of the rectangle, beginning at the upper-left corner. This unfortunately prevents some desirable geometric simplifications from being made. Later we may assume that γ divides Ω essentially into upper and lower half-domains.

Let $\Omega, \Gamma_N, \Gamma_0, \Gamma_1, \gamma \in \Re^2$; $h, l > 0$; $\Omega = (0, l) \times (0, h)$; $\Gamma_N = \{0\} \times [0, h] \cup (\gamma_R, l) \times$

$\{h\} \cup \{l\} \times [0,h], \Gamma_0 = [0,l] \times \{0\}$, and $\Gamma_1 = [0,\gamma_R) \times \{h\}$ for some $\gamma_R \in (0,l)$. Let $\Omega_- \subset \Omega$ be star-shaped with respect to $(0,h)$, and let $\partial\Omega_- = \gamma \cup \{0\} \times (\gamma_L, h] \cup (0,\gamma_E] \times \{h\}$ for some $\gamma_L \in (0,h), \gamma_E \in (0,w)$.

With the illustration present to aid in the discussion, it is appropriate now to make some deductions about the problem. We may see later if any of these ideas are fruitful.

Observe that we cannot expect a potential field u which is differentiable throughout Ω due to the abrupt edge where the right end of Γ_1 joins with the Neumann data of 0 that are prescribed on Γ_N.

Already it is known that the small changes in data may have large effects in the solution: we are not dealing with a "well-posed" problem. The ill-posedness of the inverse problem for determining the doping profile $C(x)$ seems as if it would be insurmountable if we were trying to find details of C in the upper-right corner of the figure. The boundary conditions in the corner may create complicated swirling effects that superimpose themselves with a more stable solution.

However, we are more interested in the behavior of C in the region D which lies largely in a direct line from portions of Γ_0 to Γ_1. Relying on intuition alone, one would expect greater possibilities for attaining information about C in this region.

In fact, global uniqueness for the adjoint inverse problem for u_1^* is known under the assumption that u_1^* is decreasing toward Γ_0.

3.3 Existence of equilibrium state

Let $w = V_0, a(x) = \text{const} = \lambda^2$, and $f(x,w) = 2\sinh(w) - C(x)$. Then (19) is written in the form

$$-\text{div}(a(x)\nabla w) + f(x,w) = 0, x \in \Omega \tag{20}$$

with the mixed data

$$\partial_\nu w = 0 \text{ on } \Gamma_N, w = U + V_{bi} \text{ on } \Gamma_D. \tag{21}$$

Since $a(x) = \lambda^2$ is uniformly bounded away from 0, the equation (20) is uniformly elliptic. Moreover, $f(x, w)$ is a monotone increasing function of w, since $\partial_w f = 2\cosh(w)$, i.e., $\partial_w f = \partial_w(e^w - e^{-w}) = e^w + e^{-w}$ is positive for all w.

Let $C^* = 1 + \sup_{x\in\Omega} |C(x)|$. Then $C^* < \infty$ since $C \in L^\infty(\Omega)$.

Let $g_-(w) = 2\sinh(w) - C^*$, and $g_+(w) = 2\sinh(w) + C^*$. Then $g_-(w) \le f(x, w) \le g_+(w)\ \forall\ x \in \Omega, \forall\ w$.

Let $w_- = \sinh^{-1}(C^*/2)$, and $w_+ = \sinh^{-1}(-C^*/2)$. Then $g_-(w_-) = g_+(w_+) = 0$.

Then by [19] there exists a unique solution $w \in H^1(\Omega) \cap L^\infty(\Omega)$ to (20) and hence there exists a unique solution $V_0 \in H^1(\Omega) \cap L^\infty(\Omega)$ of (19).

This solution satisfies

$$\underline{w} \le w(x) \le \overline{w}$$

where

$$\underline{w} = \min\left\{\inf_{\partial\Omega_D} U + V_{bi}, w_-\right\}$$

$$\overline{w} = \max\left\{\sup_{\partial\Omega_D} U + V_{bi}, w_+\right\}$$

3.4 Methods of proof

One of our research goals is to explore the possibilities that more than one kind of mathematical technique gives us, as the available theory responds to the type of conditions imposed by this problem. Several options may be chosen, and we think

that we find more that is of material interest to us and to the reader, and which contributes to knowledge of the inverse problem—by trying to fit the problem into a variety of literature results that are written but often incomplete in their pertinence to such a problem—than by writing down only the answer.

Now that we have proved the desired result (existence of a solution V_0 in the equilibrium state), we now discuss some of these methods of proof.

3.4.1 Monotone iterations

We do not use this technique to provide a proof of existence. At one important point, the monotone iteration technique fails, because a lower bound is not obtained for the iterations. For numerical methods, we can make this assumption, and some results are obtained.

Let $C \in L_2(\Omega)$, $a \leq C \leq b$ for some $a, b \in \Re$.

By using the stationary drift diffusion system and the equilibrium solution $U = 0, V = V_0, u = 1, v = 1$, we will observe that V_0 is a solution to the semilinear elliptic equation

$$\lambda^2 \triangle V_0 = \delta^2 (e^{V_0} - e^{-V_0}) - C \text{ in } \Omega \tag{22}$$

and satisfies the boundary conditions

$$V_0 = V_{bi}, \text{ on } \Gamma_D, \partial_\nu V_0 = 0 \text{ on } \Gamma_N \tag{23}$$

Since the doping profile C is still unspecified, we are free to divide by δ^2 and obtain

$$\lambda^2 \triangle V_0 = e^{V_0} - e^{-V_0} - C \text{ in } \Omega \tag{24}$$

By moving terms to the left side, we obtain an operator $Qu = \lambda^2 \triangle u - (e^u - e^{-u}) + C$.

As in Theorem 10.1 [10, p. 263], we have $b(x, u, Du) = e^{-u} - e^u$ and $\partial_u b = -(e^{-u} + e^u)$, and so b is nonincreasing in u. By using Fixed Point Theorems in $H_{(1)}(\Omega) \cap L_\infty(\Omega)$ [20] one can show existence of a weak solution for sufficiently small voltages (i.e., near equilibrium).

Alternatively, we may use the important condition [7] that $0 < V_0 < 1$ in Ω, $V_0 \in H_2(\Omega)$ and establish solvability for V_0 by using monotone iterations.

In fact, given u, v, we may find V, and given V, there are corresponding u, v, all in $H_2(\Omega)$ [7].

Let u be the solution to the Poisson equation $\lambda^2 \triangle u = -C$ satisfying the boundary data, and let $V_{00} = u$.

Now we consider solutions V_{0n}, $n \in \aleph$, of the linear partial differential equations

$$\lambda^2 \triangle V_{0n} = F(V_{0n-1}) \text{ in } \Omega$$

where $F(x; V) = e^V - e^{-V} - C(x)$, and $F(V_{0n-1})$ is written for $F(x; V_{0n-1})$.

This is where a point of difficulty is reached in carrying out monotone iterations. Comparing the solutions F_{0n} and F_{0n+1} we find that the iterations are not monotone, but rather alternate. This does not indicate that the iteration scheme will be divergent, but it will be more difficult to find a proof establishing existence of a solution using this method.

3.4.2 Dependence on polar radius r

Another case is that we use polar coordinates where the origin is the upper-left corner in the diagram of Ω, and assume that D is star-shaped with respect to the origin [24, 13]. In the case where $V_0 = V_0(r)$ is independent of polar angle (a simplification

which the picture suggests), and $V_0(0) = [AV_D/(A+B)] + u_{0P}$, we can substitute into Poisson's equation using the relation

$$V_0(r) = u_{0P} + [2V_D/(A+B)][(r^2/2A) + r + (A/2)]$$

or

$$V_0(r) = u_{0P} + [AV_D/(A+B) - [2V_D/(A+B)][(r^2/2B) - r]$$

where the diffusion barrier is $V_D = V_0(B) - V_0(-A)$, and A, B are values of r from the boundaries Γ_0, Γ_1 in the linear field with respect to r.

By appropriately shifting the energy potential reference level, we solve for the doping profile $C(x) \equiv a(r)$. The case where $r < 0$ is of interest when reflecting the domain about the origin, or when choosing as the origin a point within the interior of Ω.

$$a(r) = \begin{cases} \sinh\left(u_{0P} + \frac{2V_D}{A+B}(r^2/2A + r + A/2)\right) - 2V_D/A(A+B) & \text{if} \quad -A \leq r \leq 0 \\ \sinh\left(u_{0P} + \frac{AV_D}{A+B} - \frac{2V_D}{A+B}(r^2/2B - r)\right) + 2V_D/B(A+B) & \text{if} \quad 0 \leq r \leq B \end{cases}$$

To make our information completely beneficial, it is instructive to consider also the methods from the calculus of variations.

3.4.3 Variational methods

If we write the problem from

$$\triangle u = f(u) + C, f(u) = e^u - e^{-u}$$

into

$$F = \frac{1}{2}|p|^2 + \int_0^u f(s)ds - C(x)u$$

where $p = (p_1, p_2)$, and $p_i = \partial_i u$, we obtain a variational problem for minimizing the functional $I(u) = \int_\Omega F(x, u, p)dx$.

As described in [10], for example, the solvability of the problem $\min_u I(u)$ implies the solvability of the Dirichlet problem for the equation $Qu = \text{div } D_p F(x, u, Du) - D_u F(x, u, Du) = 0$. To be able to make this assertion is a tremendous step. The original pioneers in this area were Euler and Lagrange [29, 27, 28].

Our problem is a mixed problem, however, and the references to these problems are less common in the literature of the calculus of variations.

Nevertheless, in our case it is easy to overcome this obstacle. We can't prescribe the needed Dirchlet data exactly on the entire boundary, since only Neumann data is given on Γ_N. However, we can determine a class of functions $G = \{g \in C^1(\Omega) : \partial_\nu g = 0 \text{ on } \Gamma_N\}$ which will satisfy the Neumann data. If the problem is solvable for the Dirichlet data $u = g$ on Γ_N for any $g \in G$, then the original problem is solvable.

By prescribing the normal derivative to the boundary, we reduce the problem of determining g to an $n-1$ dimensional subspace where once again Dirichlet data are prescribed. Since Γ_N has two components, it is necessary to consider both problems. However, by assumptions on V_{bi}, in each case the Dirichlet data on one boundary is 0, and on the other, 1.

Then with the complete set of Dirichlet data $V_0 = V_{bi}$ on Γ_D, $V_0 = g$ on Γ_N, with $\partial\Omega = \Gamma_D \cup \Gamma_N$, the problem is solvable for any $g \in G$, and thus the original problem is solvable.

In fact, it is a known result [20, Lemma 3.3.14] that if the function F is monotonically decreasing in u for all $x \in \Omega$, a is bounded and strictly positive, $\underline{g}(u) \leq f \leq \overline{g}(u)$

in Ω for all u for some $\underline{g}(u)$ and $\overline{g}(u)$, and there exist solutions w_1, w_2 of $\underline{g}(w_2) = 0$ and $\overline{g}(w_1) = 0$, then there exists a unique solution u of this problem in $H^1(\Omega) \cap L^\infty(\Omega)$.

To summarize, a solution V_0 will exist provided that the Neumann data are compatible with the n-dimensional Dirichlet data, when the Neumann data are interpreted as Dirichlet data for the $(n-1)$-dimensional problem (for the trace of V_0).

3.5 Existence of a solution for a similar problem

Let us make the observation that

$$\triangle u = e^u - e^{-u} - C \text{ in } \Omega$$

is similar in appearance to the time-dependent equation

$$\triangle u = f(u) + u_t \text{ in } \Omega$$

Referring to [23], the energy method may be used to bound solutions u from the initial data and thereby obtain existence.

It is necessary to use the Sobolev embedding of $W_0^{1,2}(\Omega) \subset L^6(\Omega)$ to establish this result.

4 Simplifications of the inverse problem

4.1 Linearization about the equilibrium solution

In (2) we referred to the linearization at equilibrium. This is the solution V_1, u_1, v_1 of the system of linear elliptic equations [14]

$$\lambda^2 \triangle V_1 = \delta^2 (e^{V_0} V_1 + e^{-V_0} V_1 + e^{V_0} u_1 - e^{-V_0} v_1) \tag{25}$$

$$\text{div } (e^{V_0} \nabla u_1) = C_1 Q(V_0, 1, 1, x)(u_1 + v_1) \text{ in } \Omega \tag{26}$$

$$\text{div } (e^{-V_0} \nabla v_1) = C_2 Q(V_0, 1, 1, x)(u_1 + v_1) \text{ in } \Omega \tag{27}$$

where C_1, C_2 are small constants, with the boundary data

$$V_1 = U_1, u_1 = -U_1, v_1 = U_1 \text{ on } \Gamma_D \tag{28}$$

$$\partial_\nu V_1 = \partial_\nu u_1 = \partial_\nu v_1 = 0 \text{ on } \Gamma_N \tag{29}$$

In the equations, Q is defined implicitly by

$$Q(V, u, v, x) = R(n, p, x)/(np - n_i^2)$$

where R is the recombination-generation rate [8].

Differentiation of (14) to obtain these linearized equations is justified by estimates [14] for elliptic equations.

4.2 Unipolar case

Even the linearized system looks too complicated, and so one considers practically valuable particular cases [14].

To find the concentration of electrons and holes, one is attempting to determine the doping profile $C = C(x)$. By assuming C is nonpositive (corresponding to the absence of holes), one lets $p = Q = 0$ and then the second equation in (2.7) decouples into

$$\operatorname{div}(e^{V_0}\nabla u_1) = 0 \text{ in } \Omega, u_1 = -U_1 \text{ on } \Gamma_D, \partial_\nu u_1 = 0 \text{ on } \Gamma_N \tag{30}$$

The inverse problem **U1** is to find C from the mapping

$$U_1 \to \int_{\Gamma_1} \partial_\nu u_1, U_1 = 0 \text{ on } \Gamma_1, U_1 \in H_{(\frac{1}{2})}(\Gamma_0) \tag{31}$$

One expects substantial nonuniqueness, so it is realistic to find only the P-N junction γ, or equivalently the domain D with

$$C = 1 - 2\chi_D$$

Smallness of $\lambda < 10^{-3}$ and an absence of theoretical results for the inverse problem suggest letting $\lambda = 0$ in (19) and correspondingly

$$e^{V_0} - e^{-V_0} = \delta^{-2}C, \text{ or } e^{V_0} = a^+ - a^-\chi_D$$

where a^-, a^+ are known constants. Then the inverse problem **U1** simplifies to inverse problem **S1**: Find D entering the boundary value problem

$$\operatorname{div}(a\nabla u) = 0 \text{ on } \Omega \tag{32}$$

$$u = 0 \text{ on } \Gamma_1, u = U_1 \text{ on } \Gamma_0, \partial_\nu u = 0 \text{ on } \Gamma_N \tag{33}$$

with

$$a = a^+ - a^- \chi_D \tag{34}$$

from the mapping

$$U_1 \to \int_{\Gamma_1} \partial_\nu u \tag{35}$$

However, this inverse problem is hard to handle directly since we do not have sufficient boundary or interior data. It would be very desirable to obtain a solution without this information, because then the necessary data would be gathered without inconvenient modifications to the existing quality-control methods in the manufacture of semiconductors.

For this reason, we use an "adjoint" inverse problem where instead of a functional we are given extra boundary data. This device is widely used in the inverse option pricing problem (Dupire's equation) [4] where it enables one to obtain theoretical and numerical results.

4.3 Adjoint problem

Inverse Problem S1*

Let u_1^* be the solution to the boundary value problem

$$\text{div}(a \nabla u_1^*) = 0 \text{ on } \Omega, \tag{36}$$

$$\partial_\nu u_1^* = 0 \text{ on } \Gamma_N, u_1^* = 0 \text{ on } \Gamma_0, u_1^* = 1 \text{ on } \Gamma_1 \tag{37}$$

with the given data

$$a\partial_\nu u_1^* = g_1^* \text{ on } \Gamma_0 \tag{38}$$

An important item to consider [14] is how the data for the adjoint problem are determined by the data of the original problem.

Corollary 4.1. *The data of the inverse problem* **S1** *uniquely determine* $\partial_\nu u_1^* = g_1^*$ *on* Γ_1. *Moreover, there is a constant* C *such that*

$$\|g_1^*\|_{(-\frac{1}{2})}(\Gamma_0) \leq CF \tag{39}$$

where F *is the norm of the linear functional (31) in* $H_{(\frac{1}{2})}(\Gamma_0)$.

Proof. Using the definition of a weak solution to the boundary value problem (36), (37), (38) with the test function u_1 we get

$$-\int_\Omega a\nabla u_1^* \cdot \nabla u_1 + \int_{\partial\Omega} a\partial_\nu u_1^* u_1 = 0.$$

Similarly, the definition of a weak solution to the boundary value problem (32), (33) with the test function u_1^* yields

$$-\int_\Omega a\nabla u_1 \cdot \nabla u_1^* + \int_{\partial\Omega} a\partial_\nu u_1 u_1^* = 0.$$

Therefore,

$$\int_{\partial\Omega} a(\partial_\nu u_1^* u_1 - \partial_\nu u_1 u_1^*) = 0$$

and using boundary conditions (37), (38), (18), we obtain

$$\int_{\Gamma_0} a g_1^* U_1 = \int_{\Gamma_1} a \partial_\nu u_1. \tag{40}$$

The standard elliptic estimate

$$\|\partial_\nu u_1\|_{(-\frac{1}{2})}(\Gamma_1) \leq C \|U_1\|_{(\frac{1}{2})}(\Gamma_0)$$

and (40) imply that

$$\left| \int_{\Gamma_0} a u_1^* U_1 \right| \leq C \|U_1\|_{(\frac{1}{2})}(\Gamma_0).$$

This bound implies (39). □

This lemma shows that the data of the inverse problem **S1** uniquely and in a stable way determine the data g_1^* for the adjoint inverse problem **S1***.

Let us now continue to obtain descriptions of the domains Ω_- and Ω_+.

5 Evaluation of the size of domain D

5.1 Weak solutions

The equation (36) makes sense only for $a \in C^1(\Omega)$. We need to obtain an equation which is suitable for defining solutions a which may be discontinuous. We do this through multiplication by a "test function" followed by integration by parts. The resulting equation will be the definition for a so-called "weak solution" of equation (36) in Ω. Under added regularity conditions for the boundary data, a weak solution then will also be the classical solution for the original equation.

Multiply by a test function $v \in C_0^\infty(\Omega)$ with compact support in Ω (compact support means $v = 0$ outside of some compact set in Ω), and then integrate the entire equation over Ω.

$$\operatorname{div}(a\nabla u_1^*) \cdot v = 0 \cdot v$$

$$\int_\Omega \operatorname{div}(a\nabla u_1^*)v = \int_\Omega 0 \cdot v = 0$$

We may use Integration by Parts for any $u, v \in C^1(\Omega)$.

$$\int_\Omega \partial_j uv = \int_{\partial\Omega} uv\nu_j - \int_\Omega u\partial_j v$$

where ν_j denotes the j-th component of the outward unit normal to the boundary of Ω.

Now rewrite the divergence and use the rule for integration by parts.

$$\int_\Omega \operatorname{div}(a\nabla u_1^*)v$$
$$= \int_\Omega \Sigma_j \partial_j (a\partial_j u_1^*)v$$
$$= \int_{\partial\Omega} \Sigma_j (a\partial_j u_1^*)v\nu_j - \int_\Omega \Sigma_j (a\partial_j u_1^*)\partial_j v$$
$$= 0$$

so then

$$\int_\Omega \Sigma_j (a\partial_j u_1^*)\partial_j v = \int_{\partial\Omega} \Sigma_j (a\partial_j u_1^*)v\nu_j$$

Let us now rewrite the cumbersome expression $\Sigma_j (a\partial_j u_1^*)v\nu_j$ by using the very helpful shorthand notation

$$\partial_\nu u = \nu \cdot \nabla u \tag{41}$$

This expression represents a directional derivative which is convenient when the boundary term of integration by parts contains the partial derivatives of u. By using a dot product, we can match each partial derivative from ∇u with each corresponding component of the unit outward normal ν.

$$\Sigma_j (a\partial_j u_1^*)v\nu_j$$
$$= a\Sigma_j (\partial_j u_1^*)v\nu_j$$
$$= av\Sigma_j \partial_j u_1^* \nu_j$$
$$= av < \partial_1 u_1^*, \partial_2 u_1^*, \partial_3 u_1^* > \cdot < \nu_1, \nu_2, \nu_3 >$$
$$= av\nabla u_1^* \cdot \nu$$

$$= av\nu \cdot \nabla u_1^*$$

$$= av\partial_\nu u_1^*$$

Hence, if we write $\Sigma_j(a\partial_j u_1^*)v\nu_j = av\partial_\nu u_1^*$ and $\Sigma_j(a\partial_j u_1^*)\partial_j v = a\nabla v \cdot \nabla u_1^*$ we will obtain

$$\int_\Omega a\nabla v \cdot \nabla u_1^* = \int_{\partial\Omega} av\partial_\nu u_1^* \tag{42}$$

Since this holds for any integrable a, we are free to let a be piecewise constant or discontinuous. This defines a weak solution for $u_1^*, v \in H^1(\Omega)$.

In the case that $a = \text{const}$, we may integrate further (see the Appendix in Section A) and use the boundary conditions, from which it follows that

$$\int_\Omega a\triangle v u_1^* = \int_{\Gamma_N} a\partial_\nu v u_1^* + \int_{\Gamma_1} a(\partial_\nu v - v\partial_\nu u_1^*) - \int_{\Gamma_0} v g_1^* \tag{43}$$

This is convenient because all differentiation is removed from u_1^*, except in the boundary integral where it is possible to use measurements for determining the unknown values.

It appears that the presence of a is simply useless since it can be factored from both sides of the equation (it is also a factor of the data g_1^*). However, because of the nature of the problem's data, $a\partial_\nu u_1^*$, a cannot be factored out. There is a good chance that we still can determine some facts about a.

5.2 Integrals which will yield information about the doping profile

5.2.1 Constant doping profile

Now it is up to us to pick a smart test function v.

We see from (43) that

$$a = \frac{-\int_{\Gamma_0} v g_1^*}{\int_\Omega \triangle v u_1^* - \int_{\Gamma_N} a \partial_\nu v u_1^* - \int_{\Gamma_1} a(\partial_\nu v - v \partial_\nu u_1^*)}$$

A smart choice of v will eliminate the integral $\int_{\Gamma_N} a \partial_\nu v u_1^*$ for which no knowledge is available. We can do this simply by prescribing $\partial_\nu v = 0$ on Γ_N. Since such functions are dense in $L^2(\Omega)$ we still have great freedom in making further choices about v.

$$a = \frac{-\int_{\Gamma_0} v g_1^*}{\int_\Omega \triangle v u_1^* - \int_{\Gamma_1} a(\partial_\nu v - v \partial_\nu u_1^*)}$$

This eliminates part of the problem, and we want to do more. By letting v also be harmonic, i.e., $\triangle v = 0$ in Ω, we can eliminate the integral of u_1^* over the whole domain, which we also would be unable to straightforwardly evaluate from electrical measurements.

$$a = \frac{-\int_{\Gamma_0} v g_1^*}{-\int_{\Gamma_1} a(\partial_\nu v - v \partial_\nu u_1^*)}$$

Now obtain a positive numerator and denominator.

$$a = \frac{\int_{\Gamma_0} v g_1^*}{\int_{\Gamma_1} a(\partial_\nu v - v \partial_\nu u_1^*)}$$

If we let $v = 0$ on Γ_1 and multiply by a, this formula allows us to find the value of constant a.

$$a^2 = \frac{\int_{\Gamma_0} v g_1^*}{\int_{\Gamma_1} \partial_\nu v} \tag{44}$$

Let us remember the conditions

$$\partial_\nu v = 0 \text{ on } \Gamma_N, \triangle v = 0 \text{ in } \Omega, \text{ and } v = 0 \text{ on } \Gamma_1 \tag{45}$$

$$v = v_0 \text{ on } \Gamma_0, 0 \leq v_0 \leq 1 \tag{46}$$

We also need a condition on v to ensure that the denominator term $\int_{\Gamma_1} a\partial_\nu v$ is not 0. By use of the maximum principle, it is sufficient to let $v > 0$ or $v < 0$ on Γ_0. Since v is harmonic, this ensures that $\partial_\nu v$ is strictly negative (respectively, strictly positive) on Γ_1. To generalize this condition and ensure that it is satisfied, we have written $v = v_0$ on Γ_0 and imposed sufficient conditions on v_0.

It will be helpful if we identify a function v which will give us information from this identity. Our requirements are too restrictive to allow us simply to use $v \equiv const$, since the $v = 0$ on Γ_1 forces the constant to be 0. In any case, the denominator will vanish for any constant v that is used.

Instead, we may assume that $h = 1$ for the height of the semiconductor, and take a coordinate system that chooses the origin as the lower-left corner of the semiconductor, and orients x_1 rightward and x_2 upward. Then using a test function $v(x_1, x_2) = x_2 - 1$, we find that $v(x_1, 1) = 1 - 1 = 0$ on Γ_1, $\triangle v = 0$, and $\partial_\nu v = \nu \cdot \nabla v = (\pm 1, 0) \cdot (0, 1) = 0$ on Γ_N, so it is possible for us to use (48). Inserting our choice of v, we obtain

$$a^2 = \frac{\int_{\Gamma_0} g_1^*}{\int_{\Gamma_1} \partial_\nu v} = \frac{\int_{\Gamma_0} g_1^*}{\int_{\Gamma_1} \partial_2 v} = \frac{\int_{\Gamma_0} g_1^*}{\int_{\Gamma_1} 1} = \frac{\int_{\Gamma_0} g_1^*}{\gamma_R} \tag{47}$$

where γ_R is the length of Γ_1.

5.2.2 Justification for piecewise constant doping profile

The asymptotic expansion obtained by [20] allows dropping terms with λ^2 and justifies the use of a piecewise constant doping profile C.

If the conductivity coefficient $a = e^{V_0}$ is known, then solvability of (19) allows C to be found, so determining a is essentially the same as to determine C. Therefore, the determination of doping profile may be reached through methods which obtain the conductivity coefficient $a = a(x)$ in Ω from (36) for smooth $a \in C^1(\Omega)$, or (42) for discontinuous $a \in L^\infty(\Omega)$.

5.2.3 Piecewise constant doping profile

Let us write $\Omega_- = D, \Omega_+ = \Omega - \overline{\Omega}_-$, and let us assume that $a = a^+ = \text{const}$ in Ω_+ and $a = a^- = \text{const}$ in Ω_-. Along γ, let us take ν as the outward-directed unit normal vector with respect to Ω_+.

From the definition of a weak solution in (42), we have

$$-\int_\Omega a\nabla u_1^* \cdot \nabla v + \int_{\partial\Omega} a\partial_\nu u_1^* v = 0$$

for $u_1^*, v \in H^1(\Omega)$.

From $\partial_\nu u_1^* = 0$ on Γ_N

$$-\int_\Omega a\nabla u_1^* \cdot \nabla v + \int_{\Gamma_0\cup\Gamma_1} a\partial_\nu u_1^* v = 0$$

Let $v = 0$ on Γ_1, then

$$-\int_\Omega a\nabla u_1^* \cdot \nabla v + \int_{\Gamma_0} a\partial_\nu u_1^* v = 0$$

Now

$$\int_\Omega a\nabla u_1^* \cdot \nabla v = \int_{\Omega_+} a^+\nabla u_1^* \cdot \nabla v + \int_{\Omega_-} a^-\nabla u_1^* \cdot \nabla v$$

Assume that $\triangle v = 0$ in Ω, $\partial_\nu v = 0$ on Γ_N and use the identities

$$\int_{\Omega_+} a^+ \nabla u_1^* \cdot \nabla v = \int_{\partial\Omega_+} a^+ u_1^* \partial_\nu v = \int_\gamma a^+ u_1^* \partial_{\nu+} v$$

$$\int_{\Omega_-} a^- \nabla u_1^* \cdot \nabla v = \int_{\partial\Omega_-} a^- u_1^* \partial_\nu v = \int_{\Gamma_1} a^- \partial_\nu v + \int_\gamma a^- u_1^* \partial_{\nu-} v$$

Then

$$\int_{\Gamma_0} a \partial_\nu u_1^* v = \int_\Omega a \nabla u_1^* \cdot \nabla v$$

$$= \int_{\Omega_+} a^+ \nabla u_1^* \cdot \nabla v + \int_{\Omega_-} a^- \nabla u_1^* \cdot \nabla v$$

$$= \int_\gamma a^+ u_1^* \partial_{\nu+} v + \int_{\Gamma_1} a^- \partial_\nu v + \int_\gamma a^- u_1^* \partial_{\nu-} v$$

$$= \int_{\Gamma_1} a^- \partial_\nu v + \int_\gamma (a^+ - a^-) u_1^* \partial_{\nu+} v$$

Combining the left side with the right and using $a\partial_\nu u_1^* = g_1^*$ on Γ_0, we have the final answer.

$$\int_{\Gamma_0} v g_1^* = \int_{\Gamma_1} a^- \partial_\nu v + \int_\gamma (a^+ - a^-) u_1^* \partial_{\nu+} v \tag{48}$$

We remind that the conditions imposed on the function v are

$$\partial_\nu v = 0 \text{ on } \Gamma_N, \triangle v = 0 \text{ in } \Omega, \text{ and } v = 0 \text{ on } \Gamma_1 \tag{49}$$

$$v \in H^1(\Omega) \tag{50}$$

$$v = v_0 \text{ on } \Gamma_0, 0 \le v_0 \le 1 \tag{51}$$

For assistance in writing formulas, we keep the boundary condition $v = v_0$ on Γ_0. Now we explore some promising test functions v.

5.3 Discovering the value of $a^+ - a^-$ or the slope of linear γ

In this section we undertake the most difficult and important task. So far we have established a mathematical identity (48) based on (42) which is quite safe to rely upon. However, making no further assumptions about u_1^* or v, the information that (42, 48) will yield is limited to approximations of the type in Section 5.5.

We are not justified in making very many more assumptions, however. An elliptic boundary value problem is interesting in that the boundary values characterize the solution in the interior of the domain, which is a great gift for recovering solutions to inverse problems. The downside is that one cannot carelessly assume things about the solution in the interior of the domain, at the same time that one requires certain boundary conditions to hold.

We are going to make some relatively safe assumptions and consider what happens if the assumptions are true. By doing this, we will gain insight into the problem and see how to formulate equations that extract desirable information. Then we will be able to obtain trustworthy formulas that were inspired by the information we gained.

Physical models of the electrical potential in the direct problem for the doping profile may provide justification for these assumptions. If the assumptions are perfectly true, then our results will precisely determine the area of D and the location of γ. If the assumptions are completely wrong, our safer results will still steer the electrical engineer or quality control analyst in the correct direction. It may be in some circumstances that there is simply not enough data to resolve the doping profile in a semiconductor. In those cases, the series of steps which led down to this adjoint inverse problem will have produced too much mathematical fog for numeric computations to be stable or accurate.

In most cases, the actual severity of the inverse problem for finding the doping profile will be between the two extremes. It seems reasonable that experience in

testing will provide a range of voltages for which our results are the most accurate. At the very least, we hope to provide some way of measurement for the area A_D of the domain D. Then by experimentation a manufacturer may determine and tabulate the acceptable ranges of values for A_D, and these may be used for quality control.

With both many and few assumptions, we will take the height of Ω to be 1, and choose the lower-left corner of Ω as the origin, with the rightward orientation for x_1 and the upward orientation for x_2.

Let $\gamma(t)$ be parameterized by

$$\gamma(t) = \nu_1 x_1 + \nu_2 x_2 - \nu_2 c = 0 \text{ with } \nu_1^2 + \nu_2^2 = 1. \tag{52}$$

We write the equation in this form so that the intersection of γ with the left wall $(0, x_2)$ of the semiconductor will be $x_2 = c$.

It is too restrictive to assume that γ is a level curve of the potential field or a linear Instead, the assumption says that the potential decays in a linear fashion in the direction of γ. By illustrations in [7, 8] and from inspecting solutions to harmonic boundary value problems solved by the five-point method, the curve γ appears to be close to a level curve of the harmonic potential field, except when it must intersect the right boundary of Γ_1, when the potential must be 0. We are considering the case when this intersection does not occur.

We use these assumptions as a method for developing the formula which will determine the desired area of D. We will obtain a permissible formula according to elliptical theory (not making any assumptions about u_1^* within Ω), so there is no harm from using this technique.

Now choose the test function

$$v(x_1, x_2) = 1 - x_2. \tag{53}$$

Then we may find information about ν_1, ν_2 or about the jump of a,

$$[a] = a^+ - a^-. \tag{54}$$

The assumption that $a^- = 1/a^+$ is physically meaningful value for representing the piecewise constant doping profile within Ω_- and Ω_+.

From (48) we find that

$$\int_{\Gamma_0} (1 - x_2) g_1^* = \int_{\Gamma_1} a^-(-1) + [a] \int_\gamma u_1^*(-\nu_2) dt$$

Integrating where possible and inserting $x_2 = 0$ on Γ_0, we have

$$-\int_{\Gamma_0} g_1^* = -a^- \int_{\Gamma_1} 1 - \nu_2 [a] \int_\gamma u_1^*$$

and so

$$-\int_{\Gamma_0} g_1^* = a^- \gamma_R + \nu_2 [a] \int_\gamma u_1^* \tag{55}$$

since the length of Γ_1 is γ_R.

Then we find that $A_D = w - cw + \frac{\nu_1}{\nu_2} \frac{w^2}{2}$, so rearranging to simplify with respect to ν_2 we obtain

$$A_D = w \left(\frac{w}{2} \sqrt{\frac{1}{\nu_2^2} - 1} + 1 - c \right) \tag{56}$$

Also, we note the value of ν_2.

$$\nu_2 = -\frac{\int_{\Gamma_0} g_1^* + a^- w}{[a] \int_\gamma u_1^*} \tag{57}$$

We can use the maximum principle for the harmonic function u_1^* to estimate the value of $\int_\gamma u_1^*$.

5.4 Evaluating the size of D

Using these findings, we are able to find evaluators for the size of the domain D when the doping profile is constant or piecewise constant.

5.4.1 Constant a

In the same case mentioned in Section 3.4.2, where γ displays radial symmetry with respect to the point $(x_1, x_2) = (0, h)$, we may evaluate the size of the domain D. To do so, we assume that $a = 1$, which is a physically meaningful value. For quality control purposes, it is desirable to verify that D is close enough to a certain size. It is easy to establish a table of acceptable values for the size of D using this choice $a = 1$.

Then the size A_D of D is given by

$$A_D = \pi r_D^2/4 \tag{58}$$

where

$$r_D = \int_{\Gamma_0} g_1^* \tag{59}$$

The arclength s_γ of γ will be given by

$$s_\gamma = \pi r_D/2 \tag{60}$$

5.4.2 Piecewise constant a

Once again, we may evaluate the size of the domain D. This time we assume that $[a] = 1, c = 0$, and we find that

$$A_D = w\left(\frac{w}{2}\sqrt{\frac{1}{\left(-\frac{\int_{\Gamma_0} g_1^* + a^- w}{\int_\gamma u_1^*}\right)^2} - 1} + 1\right) \tag{61}$$

5.5 Approximating the integral of a over Ω

For the moment, let us work with the more general equation which is valid even for discontinuous a. This will permit us to approximate the value of a weighted by the relative sizes of Ω_- and Ω_+.

Assume that $m_1 \leq |\nabla v| \leq M_1$ for some $0 < m_1 < M_1 < \infty$.

Then

$$\begin{aligned}\int_\Omega a\nabla v\nabla u_1^* & \\ &\leq M_1 \int_\Omega a|\nabla u_1^*| \\ &\leq M_1 \left(\int_\Omega a^2\right)^{1/2} \left(\int_\Omega |u_1^*|^2\right)^{1/2}.\end{aligned}$$

Now let

$$I = \int_{\partial\Omega} av\partial_\nu u_1^* = \int_{\partial\Gamma_0\cup\Gamma_1} av\partial_\nu u_1^*$$

Then

$$\left(\int_\Omega a^2\right)^{1/2} \geq \frac{|I|}{M_1 E(u_1^*)}$$

where

$$E(u_1^*) = \left(\int_\Omega |\nabla u_1^*|^2\right)^{1/2}. \tag{62}$$

This allows us to evaluate the size of the domain D if we make assumptions about a, or to find the value of a if we make assumptions about D.

6 Conclusion

The representation in (48) will allow us to use the methods described near equation (44) for determining the interface γ.

We obtained evaluators for the size of D in the case that the doping profile is constant (58) or piecewise constant (61). These formulas may be useful in the industry as measurements for quality control.

In two dimensions some uniqueness results for the doping profile (or equivalently, the P-N junction γ) are available [21] in the literature. These are sufficient to show uniqueness in the unipolar case which we have considered.

Since the map between the data in (31) uniquely determines data in 4.3 (38), we can obtain the desired device information by solving the easier problem. We believe that the adjoint inverse problem can be used to develop a numerical method for the reconstruction of the doping profile.

References

REFERENCES

[1] L. V. Ahlfors, *Complex Analysis*, McGraw-Hill, 1979.

[2] N. Ashcroft, N. Mermin, *Solid State Physics*, Sauners College, Philadelphia, 1976.

[3] J. Bebernes, D. Eberly, *Mathematical Problems from Combustion Theory*, Springer-Verlag, 1989.

[4] I. Bouchouev, V. Isakov, *Uniqueness, stability, and numerical methods for the inverse problem that arises in financial markets*, Inverse Problems, 15, (1999), R95-116.

[5] K. F. Brennan, *The Physics of Semiconductors*, Cambridge University Press, 1999.

[6] Brillouin Zones, Computational Physics group at the Technion, `<phycomp.technion.ac.il/~nika/brillouin_zones.html>`, 2000.

[7] M. Burger, H. W. Engl, P. A. Markowich, Paola Pietra, Identification of doping profiles in semiconductor devices, Inverse Problems, 17, (2001), 1765-1795.

[8] M. Burger, H. W. Engl, A. Leitão, P. A. Markowich, On inverse problems for semiconductor equations, Milan J. of Mathematics, 72, (2004), 273-314.

[9] W. C. Dunlap, *An Introduction to Semiconductors*, John Wiley & Sons, 1957.

[10] D. Gilbarg, N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, 1983.

[11] L. J. Herbst, *Integrated Circuit Engineering*, Oxford, 1996.

[12] V. Hutson, J. S. Pym, *Applications of Functional Analysis and Operator Theory*, Academic Press, 1980.

[13] V. Isakov, *Inverse Problems for Partial Differential Equations*, Springer-Verlag, 2005.

[14] V. Isakov, Uniqueness for inverse problems in semiconductor theory, preprint.

[15] X. Jiang, Highly efficient room-temperature tunnel spin injector using CoFe/MgO(001), IBM Journal of Research and Development, 2006.

[16] F. John, *Partial Differential Equations*, Springer, 1982.

[17] C. Kittel, *Introduction to Solid State Physics*, John Wiley & Sons, 1971.

[18] H. Li, P. Markowich, A review of Hydrodynamical Models for Semiconductors: Asymptotic Behavior, preprint.

[19] P. A. Markowich, *The Stationary Semiconductor Device Equations*, Springer, 1986.

[20] P. A. Markowich, C. A. Ringhofer, C. Schmeiser, *Semiconductor Equations*, Vienna: Springer, 1990.

[21] A. I. Nachman, Global uniqueness for a two-dimensional inverse boundary value problem, *Ann. Math.*, 143, 71-96, 1996.

[22] W. Shockley, *Electrons and Holes in Semiconductors*, D. Van Nostrand, 1963.

[23] B. Straughan, *The Energy Method, Stability, and Nonlinear Convection*

[24] Willardson, Beer. Semiconductors and Semimetals, Vol. 15, Academic Press, 1981.

[25] K. Yosida, *Functional Analysis*, Springer-Verlag, 1971.

[26] P. Y. Yu, *Fundamentals of Semiconductors: Physics and Materials Properties*, Springer, 2004.

[27] E. Zeidler, *Nonlinear Functional Analysis and its Applications II/A: Linear Monotone Operators*, Springer-Verlag, 1985.

[28] E. Zeidler, *Nonlinear Functional Analysis and its Applications II/B: Nonlinear Monotone Operators*, Springer-Verlag, 1985.

[29] E. Zeidler, *Nonlinear Functional Analysis and its Applications III: Variational Methods and Optimization*, Springer-Verlag, 1985.

Appendix

Appendix

A Integration in the case when a is constant

Let us consider

$$\int_\Omega a\nabla v \cdot \nabla u_1^* = \int_{\partial\Omega} av\partial_\nu u_1^*$$

We may leave a in its place when we integrate the left side by parts since it is a constant and integration is a linear operator.

$$\begin{aligned}\int_\Omega a\nabla v \cdot \nabla u_1^* & \\ &= \int_\Omega a\sum_j \partial_j v\partial_j u_1^* \\ &= \int_{\partial\Omega} a\sum_j \partial_j v u_1^*\nu_j \ - \int_\Omega a\sum_j \partial_j^2 v u_1^* \\ &= \int_{\partial\Omega} av\partial_\nu u_1^*\end{aligned}$$

which implies

$$\int_{\partial\Omega} a\sum_j \partial_j v u_1^*\nu_j = \int_{\partial\Omega} av\partial_\nu u_1^* + \int_\Omega a\sum_j \partial_j^2 v u_1^*$$

Therefore

$$\int_{\partial\Omega} a\partial_\nu v u_1^* = \int_{\partial\Omega} av\partial_\nu u_1^* + \int_\Omega a\triangle v u_1^*$$

so

$$\int_\Omega a\triangle v u_1^* = \int_{\partial\Omega} a(\partial_\nu v u_1^* - v\partial_\nu u_1^*)$$

which is equivalent to

$$\begin{aligned}\int_\Omega a\triangle v u_1^* & \\ &= \int_{\Gamma_N} a(\partial_\nu v u_1^* - v\partial_\nu u_1^*) \\ &+ \int_{\Gamma_0} a(\partial_\nu v u_1^* - v\partial_\nu u_1^*) \\ &+ \int_{\Gamma_1} a(\partial_\nu v u_1^* - v\partial_\nu u_1^*)\end{aligned}$$

But $\partial_\nu u_1^* = 0$ on Γ_N and $u_1^* = 0$ on Γ_0; also, $u_1^* = 1$ on Γ_1 and $a\partial_\nu u_1^* = g_1^*$ on Γ_0, implying that

$$\int_\Omega a\triangle v u_1^* = \int_{\Gamma_N} a\partial_\nu v u_1^* + \int_{\Gamma_1} a(\partial_\nu v - v\partial_\nu u_1^*) - \int_{\Gamma_0} v g_1^*$$