INVERSE DOPING PROFILE ANALYSIS FOR SEMICONDUCTOR QUALITY CONTROL

A Dissertation by
Joseph Kenneth Myers
Master of Mathematics, Wichita State University, 2006
Bachelor of Mathematics, 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
Doctor of Philosophy

December 2009
INVERSE DOPING PROFILE ANALYSIS FOR SEMICONDUCTOR QUALITY CONTROL

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy with a major in Applied Mathematics.

Victor Isakov, Committee Chair

Alexandre Boukhgueim, Committee Member

Zhiren Jin, Committee Member

Ziqi Sun, Committee Member

Melvin Zandler, Committee Member

Accepted for the College of Liberal Arts and Sciences

William D. Bischoff, Dean

Accepted for the Graduate School

J. David McDonald, Dean
DEDICATION

To the triune God, creator of the earth and heavens,
ending and endlessness, physics and mathematics,
all we can discover, and all we cannot.
For wisdom is better than rubies;  
and all the things that may be desired  
are not to be compared to it.

Proverbs 8:11
ACKNOWLEDGMENTS

Victor Isakov, my advisor, has contributed to this dissertation by his expert advice, helpful feedback, and his willingness to share his vast analytic experience and knowledge. Also, my committee members—Dr. Boukhgueim, Dr. Jin, Dr. Sun, and Dr. Zandler—have all been great teachers and taught me many inspiring things. The National Science Foundation has supported my work through grant DMS 07-07734. I am very grateful for my family—Daddy, Mommy, four brothers, and seven sisters—who kept everything in the proper perspective. I truly thank God for giving me life and health every day, and granting me grace to discover part of the beauty he has created.
ABSTRACT

Inverse doping profile problems are linked to inverse conductivity problems under the assumptions of zero space charge and low injection. Unipolar inverse conductivity problems are analyzed theoretically via three uniqueness proofs, the first of which has been published as a paper in *Inverse Problems* [34]. Also, optimized numerical methods are developed for solving the unipolar direct conductivity problem with a piecewise constant conductivity coefficient. Finally, the unipolar inverse conductivity problem is solved for inclusions defined by as many as 9 parameters, or by as many as 120 parameters when an initial guess for each parameter is known with less than 10% error. Our free boundary identification algorithm produces a sequence of improved approximations in a way that provides both regularization and accelerated convergence towards the solution.
Semiconductor devices play critical roles in our daily lives and lifestyles—in our lives through health care and medical applications\(^1\), and in our lifestyles through computers, cell phones, and the Internet. In airline flights, in hospitals, and even in calling 911, the functionality or failure of a semiconductor device can often mean the difference between life and death.

Even aside from health care and medical issues, quality control in semiconductor manufacturing is of huge importance financially. NVIDIA lost $200 million recently, essentially because of one set of defective chipsets, with a corresponding drop in stock price from $39.67 to $5.75. Clearly, our motivation to improve quality control and reliability of semiconductor devices stems from the fact that today we are dependent, more than ever before, on a wide range of semiconductor-based technology.

Quality control for semiconductor device manufacturing currently employs many destructive methods for testing semiconductor devices, such as spreading resistance profiling. Each device is destroyed when it is tested, so naturally it is impossible to use destructive methods for complete, systematic quality control.

In contrast, we hope to open up a new future for non-destructive semiconductor device testing. Our numerical methods are based on an applied mathematical analysis technique of inverting the direct relationship between inputs of doping profiles and resulting output of electrical current flow. As a result, we obtain the inverse relationship of identifying doping profiles as a result of current flow measurements. Working backwards from output to input is a classical branch of mathematics known as the study of inverse problems.

\(^1\)For instance, transistors were first used in hearing aids, and afterwards used in computing systems.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Regulating electrical conductivity by doping</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Semiconductor direct and inverse problems</td>
<td>5</td>
</tr>
<tr>
<td>1.3 One-dimensional examples</td>
<td>7</td>
</tr>
<tr>
<td>1.4 Overview of notation and content</td>
<td>10</td>
</tr>
<tr>
<td>2 Drift-diffusion and doping profile models</td>
<td>15</td>
</tr>
<tr>
<td>2.1 Hydrodynamical models of diffusion</td>
<td>16</td>
</tr>
<tr>
<td>2.2 Time-dependent drift-diffusion equations</td>
<td>18</td>
</tr>
<tr>
<td>2.3 Stationary drift-diffusion equations</td>
<td>20</td>
</tr>
<tr>
<td>2.4 Capacitance and voltage-current measurements</td>
<td>23</td>
</tr>
<tr>
<td>2.5 Linearization about the equilibrium state</td>
<td>25</td>
</tr>
<tr>
<td>2.6 Semiconductor boundary value problem</td>
<td>29</td>
</tr>
<tr>
<td>2.7 Reflected boundary value problem</td>
<td>31</td>
</tr>
<tr>
<td>3 Uniqueness</td>
<td>33</td>
</tr>
<tr>
<td>3.1 Global uniqueness of source from many measurements</td>
<td>34</td>
</tr>
<tr>
<td>3.2 Local uniqueness of $D$ from one measurement</td>
<td>42</td>
</tr>
<tr>
<td>3.3 Global uniqueness of polygonal $D$ from one measurement</td>
<td>53</td>
</tr>
<tr>
<td>4 Numerical analysis</td>
<td>60</td>
</tr>
<tr>
<td>4.1 Description of integral operator $S$</td>
<td>61</td>
</tr>
<tr>
<td>4.2 Problems for various doping levels</td>
<td>63</td>
</tr>
<tr>
<td>4.3 Algorithm for identifying the obstacle $D$</td>
<td>70</td>
</tr>
<tr>
<td>4.4 Increasing the rate of convergence</td>
<td>72</td>
</tr>
<tr>
<td>4.5 Plots of convergence and computational efficiency</td>
<td>72</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Numerical simulations</td>
<td></td>
</tr>
<tr>
<td>5.1 Discretized direct and inverse problems</td>
<td>80</td>
</tr>
<tr>
<td>5.2 Undoped semiconductor voltage profile $k = 1$</td>
<td>83</td>
</tr>
<tr>
<td>5.3 $k = \frac{1}{10}$ reverse doping versus $k = 10$ doping</td>
<td>83</td>
</tr>
<tr>
<td>5.4 Solving the inverse problem</td>
<td>85</td>
</tr>
<tr>
<td>5.5 Disk</td>
<td>86</td>
</tr>
<tr>
<td>5.6 Ellipse</td>
<td>92</td>
</tr>
<tr>
<td>5.7 Peanut, including some numerical instability</td>
<td>94</td>
</tr>
<tr>
<td>5.8 Wing domain: Positioning of detectors</td>
<td>95</td>
</tr>
<tr>
<td>5.9 Arbitrary domains</td>
<td>100</td>
</tr>
<tr>
<td>5.10 Local versus global uniqueness</td>
<td>103</td>
</tr>
<tr>
<td>5.11 Final remarks</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>109</td>
</tr>
<tr>
<td>6 Conclusion</td>
<td>117</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>History of the inverse doping problem</td>
<td>6</td>
</tr>
<tr>
<td>2.1</td>
<td>Meanings of drift diffusion symbols</td>
<td>25</td>
</tr>
<tr>
<td>5.1</td>
<td>Number of parameters able to be reconstructed</td>
<td>115</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Simulated voltage profiles of undoped and heavily doped devices</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Apparatus used for doping of semiconductor devices</td>
<td>4</td>
</tr>
<tr>
<td>4.1</td>
<td>The zeroth order reconstruction of the voltage inside of a semiconductor</td>
<td>67</td>
</tr>
<tr>
<td>4.2</td>
<td>The first order reconstruction of the voltage inside of a semiconductor</td>
<td>68</td>
</tr>
<tr>
<td>4.3</td>
<td>‘Circus-tent” behavior with $N = 25, 50, 100$ source points on $\partial D$</td>
<td>70</td>
</tr>
<tr>
<td>4.4</td>
<td>Reconstructed circle has a radius of one with ten thousand percent noise</td>
<td>70</td>
</tr>
<tr>
<td>4.5</td>
<td>First four regularization steps using Fourier method</td>
<td>74</td>
</tr>
<tr>
<td>4.6</td>
<td>Reconstructed 5-parameter domain using Fourier regularization method</td>
<td>75</td>
</tr>
<tr>
<td>4.7</td>
<td>Semilogy plot of $|g_{ik} - g_{i\infty}|$ showing convergence to 0 as $k \to \infty$</td>
<td>77</td>
</tr>
<tr>
<td>4.8</td>
<td>Semilogy plot showing convergence of $g_i^{(N+1)^2}$. The top line shows $|g_i^{(N+1)^2} - g_{i\infty}|$. The middle curve shows $|g_i^{(N+1)^2} - g_i^{1,000,000}|$. The bottom curve shows $|g_i^{(N+1)^2} - g_i^{N^2}|$</td>
<td>78</td>
</tr>
<tr>
<td>4.9</td>
<td>Plot of seconds used for computing the solution versus the value of $N$</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Voltage profile of undoped semiconductor device $u(\cdot; D(0.3, 0.3, 0, 0), 1)$</td>
<td>84</td>
</tr>
<tr>
<td>5.2</td>
<td>$u(\cdot; D(0.3, 0.3, 0, 0), 1/10)$ voltage profile versus $u(\cdot; D(0.3, 0.3, 0, 0), 10)$</td>
<td>85</td>
</tr>
<tr>
<td>5.3</td>
<td>Inverse data $g_i$ superimposed with noisy data measurements</td>
<td>86</td>
</tr>
<tr>
<td>5.4</td>
<td>$g_i$ convergence as $N$ ranges from 10 to 50</td>
<td>88</td>
</tr>
<tr>
<td>5.5</td>
<td>Optimal positioning of sources for curves of integration</td>
<td>89</td>
</tr>
<tr>
<td>5.6</td>
<td>Inverse data $g_i$ computed without appropriate source positioning algorithm</td>
<td>91</td>
</tr>
<tr>
<td>5.7</td>
<td>Reconstructions of the circular domain $D_{0.5}$ with 0, 5, 10, 50, 1,000, and 10,000 percent noise</td>
<td>92</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES (continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8</td>
<td>Reconstructions of the psuedo-elliptical domain $D$ with 5 and 10 percent noise</td>
<td>94</td>
</tr>
<tr>
<td>5.9</td>
<td>“Dinosaur” created by physically impossible domain configuration</td>
<td>96</td>
</tr>
<tr>
<td>5.10</td>
<td>Voltage profile showing classical appearances of numerical instability due to poor source placement</td>
<td>97</td>
</tr>
<tr>
<td>5.11</td>
<td>Numerical stability even with only 2 source points on each boundary</td>
<td>98</td>
</tr>
<tr>
<td>5.12</td>
<td>Good reconstruction of a peanut-shaped domain from data with 5 percent relative noise</td>
<td>99</td>
</tr>
<tr>
<td>5.13</td>
<td>Wing test for calibration of five measurement positions</td>
<td>102</td>
</tr>
<tr>
<td>5.14</td>
<td>Successful wing test for calibration of three symmetric measurement positions</td>
<td>104</td>
</tr>
<tr>
<td>5.15</td>
<td>Unsuccessful wing test for calibration of ten narrow-band measurement positions</td>
<td>105</td>
</tr>
<tr>
<td>5.16</td>
<td>Steps of multi-stage algorithm for locating a nine-parameter domain with one percent noise in measurements</td>
<td>107</td>
</tr>
<tr>
<td>5.17</td>
<td>Nearly perfect local reconstruction of 101-Fourier-parameter domain</td>
<td>108</td>
</tr>
<tr>
<td>5.18</td>
<td>Good matching in $g_i$ for a non-identical domain</td>
<td>110</td>
</tr>
<tr>
<td>5.19</td>
<td>Good matching in $g_i$ for an identical domain</td>
<td>111</td>
</tr>
<tr>
<td>5.20</td>
<td>Nearly perfect global 9-parameter (left) and local 101-parameter (right) reconstructions</td>
<td>114</td>
</tr>
</tbody>
</table>
Once upon a time, about one hundred years ago, the properties of semiconductive materials seemed almost like magical powers, and were poorly understood. For instance, beginning in the early history of radio broadcasting, people discovered that “cat’s whiskers” had the mysterious ability to drastically improve the reception of radio waves.\footnote{So-called cat’s whiskers were actually primitive versions of Schottky diodes—a special type of semiconductor diode with a very low forward-voltage drop (resulting in higher efficiency) and a very fast switching action (on the order of one-tenth of one-billionth of one second).} Just as magnetism was harnessed and compasses superceded celestial navigation, soon semiconductivity would be enlisted as the footsoldier of a technological revolution. Within the length of one human life, mankind would reach the moon, and civilization would become reliant on strange inventions unlike anything in all known world history.

Physicists and mathematicians could not have predicted the changes that would occur in a blink of an eye on history’s timescale, but they did see the great potential of semiconductor devices. The amazing promises began to be fulfilled when Philo Farnsworth (a fourteen-year-old American farm boy, 1922 high school freshman) built the world’s first working television system, and later demonstrated it in 1928. Something more ab-

\begin{quote}
\textbf{Chapter 1}
\end{quote}

\begin{quote}
\textbf{Introduction}
\end{quote}

Once upon a time, about one hundred years ago, the properties of semiconductive materials seemed almost like magical powers, and were poorly understood. For instance, beginning in the early history of radio broadcasting, people discovered that “cat’s whiskers” had the mysterious ability to drastically improve the reception of radio waves.\footnote{So-called cat’s whiskers were actually primitive versions of Schottky diodes—a special type of semiconductor diode with a very low forward-voltage drop (resulting in higher efficiency) and a very fast switching action (on the order of one-tenth of one-billionth of one second).} Just as magnetism was harnessed and compasses superceded celestial navigation, soon semiconductivity would be enlisted as the footsoldier of a technological revolution. Within the length of one human life, mankind would reach the moon, and civilization would become reliant on strange inventions unlike anything in all known world history.

Physicists and mathematicians could not have predicted the changes that would occur in a blink of an eye on history’s timescale, but they did see the great potential of semiconductor devices. The amazing promises began to be fulfilled when Philo Farnsworth (a fourteen-year-old American farm boy, 1922 high school freshman) built the world’s first working television system, and later demonstrated it in 1928. Something more ab-
stract was needed, however, that would be a basic building block for devices of arbitrary complexity.

In 1947, this call was answered by the development of the first transistor, which functions as a solid-state switch. Almost immediately, semiconductor devices took on a new and prominent role, and an unbroken stream of new semiconductor-based technology has issued forth from creative minds ever since.

We may find it difficult to imagine, but prior to the advent of transistors, the computers in the SAGE air-defense network each contained 55,000 vacuum tubes and used three megawatts of electrical power. Transistors which replaced vacuum tubes quickly cut power requirements by a factor of approximately one million, and simultaneously increased computing ability. It is no wonder that Fortune magazine designated 1953 as “The Year of the Transistor,” only six years after the transistor’s invention [23].

Less than sixty years later, computers would be able to transmit the content of all books in the entire Library of Congress within a few minutes, and store it in a two-foot stack of 6-inch diameter “DVDs.” Even more amazing, computations per unit of energy have increased by a factor of at least $2.7 \times 10^{11}$.

1.1 Regulating electrical conductivity by doping

The electrical conductivity\(^2\) of semiconductors falls between that of conductors (typical conductivities range from $10^4$ to $10^6$ siemens per centimeter) and that of insulators (on the order of $10^{-18}$ to $10^{-10}$ siemens per centimeter). This great flexibility in conductivity

\(^2\)Units of conductivity, measured in siemens, are inversely proportional to units of resistivity, which are measured in ohms.
of semiconductors, combined with the ability of doping to control the conductivity with
great freedom in precisely specified areas, is what makes semiconducting materials so
useful. Doping allows devices to be created having a built-to-order doping profile, and
thus built-to-order electrical conductivity.

![Figure 1.1: Simulated voltage profiles of undoped and heavily doped devices](image)

The field-effect transistor (FET) is commonly used in computers. An FET uses only
one intersection of n-type and p-type materials. The gate electrode is separated from the
semiconductor surface by a thin insulating layer. A current of electrons passing from a
source electrode to a drain wire through the transistor can be controlled by the amount
of charge on the gate [23].

Complex devices, like microprocessors, are formed by linking many P and N regions
together. The basic building block is a P-N type semiconductor device where the doping
profile is essentially piecewise constant. Improvements were made to the doping process
using the zone-melting manufacturing technique.

The life of a semiconductor device begins as a pure wafer of undoped silicon substrate. A semiconductor’s conductivity is sensitive to minute amounts of impurity atoms. For example, the addition of less than 0.01 percent of a particular type of impurity can increase the electrical conductivity by four or more orders of magnitude. Carefully-controlled addition of dopants\(^3\) forms a *doping profile* that regulates electrical conductivity. The result of this *doping* process is one of 28 common types of semiconductor devices, each with unique characteristics based on their doping profiles.

Figure 1.2: Apparatus used for doping of semiconductor devices

---

\(^3\)The relevant elements from the periodic table \([6]\) are boron, carbon, nitrogen, magnesium, aluminum, silicon, phosphorus, sulfur, zinc, gallium, germanium, arsenic, selenium, cadmium, indium, tin, antimony, tellurium, mercury, and lead.
1.2 Semiconductor direct and inverse problems

The movement of electrical charge through a semiconductor is mathematically described by the system of drift-diffusion equations, which have a close connection to the conductivity equation. The so-called inverse doping profile problem \[10\] is the problem of determining the source term \( C(x), \ x \in \Omega, \) in the stationary, unipolar drift-diffusion model

\[
\lambda^2 \Delta V_0 = \delta^2 (e^{V_0} - e^{-V_0}) - C \text{ in } \Omega, \tag{1.1}
\]

where \( \Omega \) represents a semiconductor device. Equation (1.1) is closely related to equation (1.2) under conditions of zero space charge and low injection \[11\]. The conductivity coefficient in (1.2) is given by \( a = e^{V_0} \) in the case of negative charge carriers (electrons), and by \( a = e^{-V_0} \) in the case of positive charge carriers (holes).

In our research we wish to determine \( C \) from exterior measurements of applied voltage and resulting current flow. The direct problem is the problem of finding \( V_0 \) when \( C \) and the boundary measurements are provided. The meanings of important drift-diffusion parameters are given in Table 2.1, and the complete descriptions and simplification of direct and inverse problems for semiconductors are given in Chapter 2.

Our approach is to analyze the properties of doping profiles and link them with the associated conductivity coefficients in a one-to-one relationship. We will simplify the inverse doping profile problem step by step into an inverse conductivity problem. The solution to the inverse conductivity problem is then expressed as the limit of a minimizing sequence to a functional depending on a parameterized class of direct problems.

We will prove uniqueness for the limit of the minimizing sequence both locally for a
Table 1.1: History of the inverse doping problem

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1947</td>
<td>The invention of an effective “transistor” takes place.</td>
</tr>
<tr>
<td>1950</td>
<td>Roosbroeck develops the drift-diffusion equations for doping profiles.</td>
</tr>
<tr>
<td>1953</td>
<td>Transistors become successful building blocks for many devices.</td>
</tr>
<tr>
<td>1980</td>
<td>Calderón brings attention to the inverse problem for conductivity.</td>
</tr>
</tbody>
</table>

wide class of domains, and globally for a restricted class of convex, polygonal domains. All
the proofs of uniqueness rely on boundary data that are as simple as possible, reflecting
the practical constraints of limited data that may be available in the real world.

Then we will perform a number of experiments for identifying domains via simulated
voltage and current measurements. In each case the limit of the minimizing sequence
is determined up to some tolerance either by Nedler-Mead simplex direct search or by
iterations of Newton’s Method.

Our solution method for the inverse doping profile problem is based on the inverse
conductivity problem. A conductive region $\Omega$ represents the rectangular two-dimensional
vertical profile of a three-dimensional semiconductor device, which has corresponding
conductivity coefficient $a(x)$, $x \in \Omega$, $a \in L^\infty(\Omega)$. The voltage potential function
$u(x)$, $x \in \Omega$, $u \in H^1(\Omega)$ then satisfies the following equation,

$$\text{div}(a(x)\nabla u(x)) = 0 \text{ in } \Omega \subset \mathbb{R}^2.$$  \hspace{1cm} (1.2)

Intuition for partial differential equations interprets the equation of conductivity as
a statement that larger magnitude of $a(x)$ corresponds to smaller magnitude of $\nabla u(x)$.
Intuition based on principles of electricity is very similar. In a more conductive region,
electrical current flows more freely, so less voltage is used for the same amount of current
flow, and the change in (or gradient of) electrical potential is smaller.

1.3 One-dimensional examples

We can visualize the conductivity equation better by finding its solution in one dimension, in which case (1.2) is equivalent to

$$ (au')' = 0. $$  \hspace{1cm} (1.3)

Integrating once, we have

$$ au' = C, $$  \hspace{1cm} (1.4)

and so $u' = \frac{C}{a}$, that is to say, the rate of change $u'$ of voltage $u$ is inversely proportional to the conductivity. This is exactly what we already guessed based on intuitive ideas of partial differential equations or physics.

It will also be very helpful to demonstrate the solution of the direct and inverse problems in a simple case when the conductivity of a three-dimensional semiconductor device is uniform in the $x$-$y$ plane, but varies depending on $z$.

We will consider the vertical component of the device to be the interval $(0, 1)$, with voltages $u(1) = 1$ and $u(0) = 0$ given uniformly over the entire upper and lower surfaces $z = 1, 0$. We will assume that the conductivity is given for some boundary $h \in (0, 1)$ by

$$ a(z) = \begin{cases} 
1, & z < h, \\
k, & z > h.
\end{cases} \hspace{1cm} (1.5)$$
Of course, we assume that the function $u$ is continuous across the plane $z = h$. Thus, there is a constant $C_h$ such that $u(x, y, z; h) = u(z; h)$ and $u(h; h) = C_h$.

Hence, given $h$, we have the following data for solving the direct problem: $(0, 0), (h, C_h), (1, 1) \in u(; h)$.

From the differential equation, and the definition of $a$, we conclude that $u(; h)$ is piecewise linear, that is,

$$u(z; h) = \begin{cases} 
m^+z, & z \leq h, \\
C_h + m^-(z - h), & z > h. 
\end{cases} \tag{1.6}$$

In the form (1.6), the boundary condition $(0, 0) \in u(; h)$ is satisfied already. In order to satisfy continuity we require that $m^+h = C_h$, and in order to satisfy the boundary condition $(1, 1) \in u(; h)$ we require that $C_h + m^- (1 - h) = 1$.

The Maximum Principle requires that $C_h \in [0, 1]$; in fact, $C_h$ is determined uniquely for each $h \in (0, 1)$ by the first slope $m^+$, and the second slope $m^-$ is determined through the conductivity relation $m^- = m^+/k$. Hence we have the solution to the direct problem:

$$u(z; h) = \begin{cases} 
m^+z, & z \leq h, \\
m^+[h + (z - h)/k], & z > h, 
\end{cases} \tag{1.7}$$

with the compatibility condition $m^+h < 1$. In fact, there is really just one free variable $h$ which determines the solution $u(; h)$ since from the second part of the piecewise definition (1.7) we can find $m^+$ by setting $u(1; h) = 1$ and using the equation $m^+ = 1/[h + (1 - h)/k] = k/(h(k - 1) + 1)$.
With respect to the parameter $k$ we find the solutions of a family of direct problems,

$$u(; D, k) = \begin{cases} 
\frac{kz}{1+h(k-1)}, & z \leq h, \\
\frac{h(k-1)+z}{1+h(k-1)}, & z > h,
\end{cases}$$

where $D$ is the slab $h < z < 1$. We have used the more general notation $u(; D, k)$ because later we will be able to use this same notation for denoting the solution to the direct problem where $a$ is much more general, and $D$ will then indicate the set $a^{-1}(\{k\})$.

**Remark 1.** We will only attempt to solve the inverse problem for $D$, not $(D, k)$. This corresponds to the ratio of conductivities being given ahead of time, perhaps determined by destructive testing, such as spreading resistance profiling. Then the components of the semiconductor device with differing conductivities are determined by solving the inverse problem for $D$ alone, by means of nondestructive testing.

Solving the inverse problem for $D$ means to determine $h$ from the form (1.8) given some additional information about the solution $u(; D, k)$. For us, this extra “inverse data” will be $g_i(; D, k) = -\partial_z u(0; D, k)$. In other words, we know a single number representing the outward normal derivative of $u(; D, k)$ on the boundary $z = 0$.

Thus we have the equality

$$-g_i(; D, k) = \frac{k}{1 + h(k-1)},$$

and the solution

$$h = \frac{1 + \frac{k}{g_i(; D, k)}}{1 - k}.$$
This one-dimensional demonstration illustrates the philosophy that will be used to solve the inverse problem in more general cases. First, the direct problem is analyzed in detail. Then, if possible, information about the direct problem will be used to solve the inverse problem explicitly. If an explicit solution cannot be found, then knowledge of how to solve the direct problem (including determination of \( g_i \)) for a given \( D = D_n \) will be used to find a sequence \( D_n \) which is a minimizing sequence for the functional 

\[
 f(D^*) = \|g_i(; D^*, k) - g_i(; D, k)\|. 
\]

This is how the inverse problem can be solved in two dimensions.

Since the data \( g_i(; D, k) \) is given as extra “inverse data” for solving the inverse problem, we can measure the functional \( f(D^*) \) for any \( D^* \) provided we have an algorithm for solving the direct problem.

### 1.4 Overview of notation and content

Traditionally, the directional derivative of a smooth function \( f \) in the unit direction \( u \) is defined in terms of the gradient by \( \partial_u f = \nabla f \cdot u \). However, by using careful mathematical formulation we can define directional derivatives in some directions even when the gradient \( \nabla f \) does not exist. If \( f \) is sufficiently smooth, we can then define the gradient in terms of directional derivatives. Therefore, we define

\[
 \partial_u \pm f = \lim_{h \to 0^\pm} \frac{f(x) - f(x - hu)}{h}. 
\]  
(1.11)
When not specified, the positive sign is always taken in the above limit, meaning that the limit is taken from the interior of a domain at a boundary point when \( u \) is an outward-facing vector. The norm \( \|u\| \) is not present in the denominators due to the assumption that \( u \) is a unit direction vector. Directional derivatives \( \partial_\nu f \) and \( \partial_\tau f \) will always be used to indicate the normal and tangential derivatives to a boundary element. Here \( \nu \) is the unit outward normal and \( \tau \) is the 90°-counterclockwise rotation of \( \nu \).

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 1. (Embedding).** Let \( \Omega \) be a bounded Lipschitz domain in \( \mathbb{R}^n \); then

\[
W^{1,p}_0(\Omega) \subset \begin{cases} L^{np/(n-p)}(\Omega) & \text{for } p < n \\ C^0(\overline{\Omega}) & \text{for } p > n \end{cases}
\]
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 $\lambda \leq k - n/p, n < kp$.

For negative $k$, the space $H^k(\Omega)$ is the dual space of $H^{-k}(\Omega)$. For fractional $k$, the space $H_{(k)}$ is defined by interpolation in terms of Fourier coefficients.

The equation of conductivity (1.2) represents a general definition of conductivity, but it will also be derived from the system of drift-diffusion equations for the specific case of doped semiconductor devices. In the second chapter, we will provide necessary models and descriptions of available measurements for identifying doping profiles. We will work through the simplifications (unipolar case, reduction to the conductivity equation, and using adjoint [20] boundary conditions) in a systematic way to arrive at two main problems.

The first main problem is for heavily-doped semiconductors [16], for which a high contrast exists between conductivity coefficients in the doped and undoped regions. The boundary value problem in this case is similar to an inverse obstacle problem for a so-called “soft” obstacle, where Dirichlet data are prescribed on the unknown boundary. The second main problem falls into the category of inverse transmission or refraction.
The proof of unique solvability of equation (1.1) for the equilibrium state $V_0 \in H^1(\Omega) \cap L^\infty(\Omega)$ for $C \in L^\infty(\Omega)$ is contained in the author’s Master’s Thesis [33], and is based on fixed point theorems given in [28]. One may consider the case of zero space charge, where $\lambda \to 0$. Passing to the limit in (1.1) is completely justified by an asymptotic analysis in [29]. Thus we obtain an algebraic relation determining $C(x)$ from $V_0(x)$. Moreover, $V_0(x)$ is determined from the conductivity equation by the conductivity coefficient $a$ via $V_0(x) = \ln a(x)$, where we are considering the unipolar case of electrons.

We shall prove global uniqueness of source for a class of semilinear elliptic equations from the Dirichlet-to-Neumann map [34]. This result is applicable to the inverse conductivity problem, but is proved for a more general setting that encompasses some problems from modeling of ion channels. For the inverse conductivity problem we shall prove local uniqueness from a single measurement for $C^{1+\lambda}$-domains, and global uniqueness from a single measurement for convex polygonal domains. Similar results for the inverse conductivity problem have been known before [2, 4], but they required more complicated boundary data and assumed the boundary of the unknown domain to be disjoint from the boundary of the device. The latter assumption, while valid for mine detection applications, is in contradiction to the manufacturing procedures of semiconductor devices, where diffusion occurs from the boundary, and the doped region extends to the boundary. Thus, we think our results make worthwhile progress towards semiconductor quality control applications.

Effective numerical methods are developed for solving the direct conductivity problem. Finally, we present some interesting and quite successful numerical simulations of the
inverse conductivity problem. As an example, we can reproduce a wing-shaped inclusion of differing conductivity from electrical measurements taken at only three points on the boundary of a simulated semiconductor device.

We would definitely like to acknowledge the work of others in the past. The foundations for our research were laid in 1947 by the invention of the solid-state transistor, the most fundamental building block of modern technology. Our analysis is based on the mathematical theory of inverse problems and mathematical models for semiconductor devices, developed and contributed to by many, many people, especially Alessandrini, Burger, Calderon, Isakov, Leitao, Lions, Markowich, Nachman, Schmeiser, Seo, and Uhlmann, just to name a few of the researchers at the top of their respective fields. Without their preparation, our new accomplishments would have been impossible.
Models of diffusion and crystal lattice structure make up the basic concepts of semiconductor physics and form the basis for all electronic properties and functional capabilities of semiconductor devices. Our scheme for model reduction of the semiconductor inverse dopant problem was motivated by the methods of Martin Burger and Peter Markowich [8]. Solvabilility for the equation (1.1) linking inverse doping to inverse conductivity was established by the author in 2006 [33] based on previous work by [29]. Finally, the jumping off point for numerical analysis will be the simplified adjoint boundary value problem developed first by [20], and also mentioned in [34].

The most up-to-date model for semiconductor devices is the system of drift diffusion equations, which may be written in time-dependent or time-independent formulations. The drift diffusion equations are the most widely-used and well-accepted model describing important features of negative and positive charge carriers in semiconducting materials.

However, in the hierarchy of semiconductor equations, the hydrodynamical model should be described first, for a variety of reasons. One reason in particular is its appli-
cability to small-scale devices. We will mention it in recognition of its importance, and provide the best references we are aware of for the reader who wishes to know more.

2.1 Hydrodynamical models of diffusion

The hydrodynamical model is used primarily to describe high field phenomena. It may seem surprising that a liquid-state hydrodynamical model could arise in the context of solid-state semiconductor devices. 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.

For shorter scale and quantum representations when crystal lattice effects are taken into consideration, the may be more appropriate than the drift-diffusion equations. Among other things, from hydrodynamical models one sees that carrier spin polarization shows drastic suppression in the P-N junction interface, rather than slow decay in a uniformly doped system [42]. We emphasize the unipolar nature of the hydrodynamical model as a simplifying technique. The same simplification will later be used in the drift diffusion equations.

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

With time \( t > 0 \), we obtain from the Boltzmann equation the unipolar hydrodynamical system modeling semiconductors.
\[ \partial_t f + v(k) \nabla_x f + \frac{q}{\hbar} \nabla_x V \cdot \nabla_k f = Q(f, f), \quad (x, k, t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^+, \quad (2.1) \]

\[ f(x, k, 0) = f_0(x, k), \quad (x, k) \in \mathbb{R}^3 \times \mathbb{R}^3. \quad (2.2) \]

Here \( q > 0 \) is the elementary charge and \( \hbar \) the reduced Planck constant; \( v(k) = (1/\hbar) \nabla_k \epsilon(k) \) 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

\[ Q(f, f) = \int \psi(x, v, v') (Mf' - Mf') dv', \]

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

\[ M(v) = \left( \frac{m}{2\pi k_B T} \right)^{3/2} \exp \left( \frac{-m|v|^2}{2k_B T} \right). \]

The system is described more fully in [27]. Also, we refer to [10] for more information on the hydrodynamical model, and we believe that [39] is a good source for comprehensive information about the physics of semiconductors, including first principles on which the hydrodynamical model is based.
2.2 Time-dependent drift-diffusion equations

Our starting point for mathematical analysis is a commonly accepted model of the semiconductor device, the drift diffusion equations [29]. We note that this system of equations is quite similar in its properties to the equations from Combustion Theory [5, 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 feasibility of computational solutions for the chosen nonlinear system of partial differential equations.

We consider the following coupled system [36] of nonlinear partial differential equations for electrostatic potential $V$, the nonnegative concentrations of free carriers of negative charge density $n$ (electrons) and positive charge density $p$ (holes), which is solved in the domain $\Omega \subset \mathbb{R}^d (d = 1, 2, 3)$ representing the semiconductor device, and in a time interval $[0, T]$. The dependence on time that is included in this system of equations should be emphasized.

\[
\begin{align*}
\text{div}(\epsilon_s \nabla V) &= q(n - p - C), \\
\text{div}(D_n(E, T^K) \nabla n - \mu_n(E, T^K) n \nabla V) &= n_t, \\
\text{div}(D_p(E, T^K) \nabla p + \mu_p(E, T^K) p \nabla V) &= p_t, \\
\text{div} k(T^K) \nabla (T^K) &= \rho c(T^K) T^K_t - H.
\end{align*}
\]

Above $\epsilon_s$ denotes the positive semiconductor permittivity coefficient (e.g., for silicon) and $q$ the positive unit of elementary charge—both $\epsilon_s$ and $q$ depend on dimension; $\mu_n$ and $\mu_p$ denote the electron and hole mobility, and $D_n$ and $D_p$ are the electron and hole
diffusion coefficients. Observe that $-\nabla V$ is the electric field with electric field strength $E = |\nabla V|$. The function $R = R(n, p, x)$ denotes the recombination-generation rate.

The constants $\rho$ 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. $T^K$ is the absolute temperature.

We assume that $R$ is of the standard form

$$R = F(n, p, x)(np - n_i^2), \quad (2.4)$$

where $F$ is a nonnegative smooth function, which holds, for example, for the frequently used Shockley-Read-Hall rate

$$R_{SRH} = \frac{np - n_i^2}{\tau_p(n + n_i) + \tau_n(p + n_i)}. \quad (2.5)$$

The function $C = C(x)$ represents the doping concentration, which is produced by diffusion of different materials into the silicon crystal (zone melting technique) and by implantation with an ion beam. When $C < 0$ it represents the P region of the semiconductor and when $C > 0$ it represents the N region. See the literature [10, 25, 7, 29, 44] for more details. [Note that there is a mistake in [12], which specifies on page 1,778 that $C > 0$ in both P and N regions.]

This system is supplemented by homogeneous Neumann boundary conditions on an insulated part $\partial \Omega_N$ (open in $\partial \Omega$ of the boundary, where zero current flow and zero electric field in the normal direction are prescribed [33]. On the remaining part $\partial \Omega_D$ (with positive $(d - 1)$-dimensional Lebesgue measure), the following Dirichlet conditions
are imposed:

\[ V(x,t) = V_D(x,t) = U(x,t) + V_b(x) = U(x) + U_T \ln \left( \frac{n_D(x)}{n_i} \right), \quad (2.6) \]

\[ n(x,t) = n_D(x) = \frac{1}{2} \left( C(x) + \sqrt{(C(x))^2 + 4n_i^2} \right), \quad (2.7) \]

\[ p(x,t) = p_D(x) = \frac{1}{2} \left( \sqrt{(C(x))^2 + 4n_i^2} - C(x) \right), \quad (2.8) \]

on \( \partial \Omega_D \times (0,T) \), where \( n_i \) is the intrinsic carrier density, \( U_T \) is the (nonnegative) thermal voltage and \( U \) is the applied potential. Moreover, the initial (time) conditions,

\[ n(x,0) = n_0(x) \geq 0, \]
\[ p(x,0) = p_0(x) \geq 0. \quad (2.9) \]

have to be supplied.

### 2.3 Stationary drift-diffusion equations

The *stationary drift-diffusion model* is obtained from the transient model by supressing time-dependence and omitting the initial (time) conditions. We use standard assumptions about the mobilities and diffusion coefficients (Einstein relations),

\[ D_n = \mu_n U_T, \]
\[ D_p = \mu_p U_T, \quad (2.10) \]
in order to transform the system using the so-called *Slotboom variables* $u$ and $v$ defined by

\begin{align}
  n &= C_0 \delta^2 e^{V/U} u, \\
  p &= C_0 \delta^2 e^{-V/U} v,
\end{align}

(2.11)

where

\begin{equation}
  \delta^2 = \frac{n_i}{C_0},
\end{equation}

(2.12)

and $C_0$ is the doping profile scale. If we rescale all quantities to dimensionless analogues (see [29] for further details), and if we define scaled current densities for electrons and holes,

\begin{align}
  J_n &= \mu_n \delta^2 e^V \nabla u, \\
  J_p &= \mu_p \delta^2 e^{-V} \nabla v,
\end{align}

(2.13)

then we obtain the stationary system

\begin{align}
  \lambda^2 \Delta V &= \delta^2 (e^V u - e^{-V} v) - C, \\
  \text{div } J_n &= \delta^4 Q(u,v,V,x)(uv-1), \\
  \text{div } J_p &= -\delta^4 Q(u,v,V,x)(uv-1).
\end{align}

(2.14)

Above $\lambda^2$ is a positive constant,

\begin{equation}
  \lambda^2 = \frac{\epsilon_s U_T}{qC_0 L^2},
\end{equation}

(2.15)

where $L$ is a typical length scale of a semiconductor device, usually about $10^{-4}$ cm. In order to obtain a completely non-dimensional system, we have used the scaled length $x/L$, and replaced the mobilities $\mu_n$ and $\mu_p$ by $(U_T/L)\mu_n$ and $(U_T/L)\mu_p$. We have also assumed $\epsilon_s$ is constant and rescaled the potential to $V/U_T$. Also, $Q$ is defined by the
relation \( F(n, p, x) = Q(u, v, V, x) \).

The Dirichlet boundary conditions can be written as

\[
V = U + V_{bi} = U + \ln \left( \frac{1}{2\delta^2 (C + \sqrt{C^2 + 4\delta^2})} \right) \quad \text{on} \quad \partial \Omega_D, \quad (2.16)
\]

and

\[
\begin{align*}
    u &= e^{-U} \quad \text{on} \quad \partial \Omega_D, \\
    v &= e^U \quad \text{on} \quad \partial \Omega_D.
\end{align*}
\]

(2.17)

On the remaining part \( \partial \Omega_N = \partial \Omega \setminus \partial \Omega_D \), the homogeneous Neumann conditions can be formulated in terms of \( J_n \) and \( J_p \), i.e.,

\[
\begin{align*}
    \partial_\nu V &= \partial_\nu J_n, \\
    \partial_\nu J_n &= \partial_\nu J_p, \\
    \partial_\nu J_p &= 0,
\end{align*}
\quad \text{on} \quad \partial \Omega_N.
\]

(2.18)

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 or capacitor, it is doubtful that extremely detailed information about \( C(x) \), the doping profile, can be determined from inverse problems neglecting the time dependence of the equations, since inductance and capacitance are both effects which involve time. However, an amazing amount of analysis can still be done using only voltage and current measurements, which assume only a steady-state equilibrium.
2.4 Capacitance and voltage-current measurements

Nondestructive experimental measurements must always be taken on the boundary of the device, more precisely on a contact $\Gamma_1 \subset \partial \Omega_D$. Following [11] we will use the notation

$$\Sigma_1 := \begin{cases} 
\Gamma_1 & \text{in the stationary case} \\
\Gamma_1 \times (0, T) & \text{in the transient case.}
\end{cases}$$

(2.19)

In the general case two different types of data can be measured from semiconductor devices.

1. Voltage-Current Data (denoted by $I_U$) are given by measurements of the normal component of the current density $J := (J_n + J_p)$ on $\Sigma_1$, i.e.,

$$I_U := (J_n + J_p)|_{\Sigma_1}$$

(2.20)

for all applied voltages $U \in \mathcal{U}$, where $\mathcal{U}$ is an appropriate class of functions.

Note that it is physically impossible to distinguish the electrical current due to $J_n$ from the electrical current due to $J_p$ [43].

2. Capacitance Data (denoted by $Q_\Phi$) around the voltage $U$ are measurements of the variation of the electric flux in the outward normal direction ($\partial_n V$ on $\Sigma_1$) with respect to the voltage $\Phi$, i.e.,

$$Q_\Phi := \lim_{s \to 0} \frac{1}{s} (\partial_n V^{s\Phi + U} - \partial_n V^U)|_{\Sigma_1}.$$
for all voltages $\Phi \in \mathcal{U}$, where $V^\Phi$ denotes the solution of the Poisson equation with $U = \Phi$. For simplicity, we assume $U \equiv 0$, i.e., we are interested in capacitance data round equilibrium.

Results about well-posedness and regularity for solutions of the transient and stationary drift-diffusion models show that for a given doping profile $C$, both current and capacitance are well-defined outputs for appropriate classes of voltages $\mathcal{U}$. In the stationary case, appropriate means smoothness, for example, $U \in H^{3/2}(\partial \Omega_D)$ and we assume (for linearized stability) smallness of $U$ as well, since hysteresis might occur for large voltages (c.f. [29] for examples of nonunique solutions).

In fact, it is a well known property of semiconductor devices, even in popular literature [23], that separate low-voltage and high-voltage states exist. For a time in both the 1960s and the 1980s, large projects were undertaken to develop computing machines from negative-resistance elements depending on precisely these bifurcations between stable states of voltage. However, the lack of a way to produce a true digital signal standard, and the difficulty in achieving required accuracy even for a few tunnel diodes, meant that these intense development efforts were eventually abandoned. In the 1970s, a great deal of attention was devoted to the possibility of computing with bistable circuits that used superconducting devices known as Josephson junctions. However, once again a standardized signal could not be achieved. Disillusionment set in, and the attempts to develop Josephson-junction-based computers were terminated.

We make note of the materials data for silicon at room temperature (about 298° Kelvin), which give $q = 10^{-19}$ and $U_T = 2.5 \cdot 10^{-2}$. Also, according to [43], $\epsilon = 1.0536 \cdot 10^{-12} \text{As/Vs}$ and $n_i = 9.65 \cdot 10^9 \text{cm}^{-3}$ in silicon at a temperature of 300 Kelvin. This data
Table 2.1: Meanings of drift diffusion symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_s$</td>
<td>semiconductor permittivity</td>
</tr>
<tr>
<td>$q$</td>
<td>elementary charge</td>
</tr>
<tr>
<td>$\mu_n$</td>
<td>electron mobility</td>
</tr>
<tr>
<td>$\mu_p$</td>
<td>hole mobility</td>
</tr>
<tr>
<td>$D_n$</td>
<td>electron diffusion coefficient</td>
</tr>
<tr>
<td>$D_p$</td>
<td>hole diffusion coefficient</td>
</tr>
<tr>
<td>$R$</td>
<td>recombination-regeneration rate (in general depends on $n, p$)</td>
</tr>
</tbody>
</table>

was verified also in [13] and [29].

Our exposition of the semiconductor equations is completed by tabulating some of the various symbols in Table 2.1. All physical parameters are known except for $C \in L^\infty(\Omega)$ (the unknown doping profile).

### 2.5 Linearization about the equilibrium state

An advantage of the stationary drift diffusion system in the form (2.14) is its possession of an equilibrium solution $U = 0, V = V_0, u = 1, v = 1$. The equation which necessarily must be satisfied by $V_0$ is

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

(2.22)

with boundary conditions

$$
V_0 = V_{bi} \text{ on } \partial \Omega_D, \quad \partial_n V_0 = 0 \text{ on } \partial \Omega_N.
$$

(2.23)
Unique solvability of (2.22), (2.23) for $V_0 \in H^1(\Omega) \cap L^\infty(\Omega)$ was established in 2006 by [33].

The linearization $(V_1, u_1, v_1)$ at equilibrium is the solution [20, 33] to the system of linear elliptic equations

\begin{align*}
&\lambda^2 \Delta V_1 = \delta^2 (a_u V_1 + a_v V_1 + a_u u_1 - a_v v_1), \\
&\text{div}(a_u \nabla u_1) = C_1 Q(\cdot, V_0, 1, 1)(u_1 + v_1), \\
&\text{div}(a_v \nabla v_1) = C_2 Q(\cdot, V_0, 1, 1)(u_1 + v_1),
\end{align*}

(2.24) \quad (2.25) \quad (2.26)

where

$$C_1 = d^2/(\mu_n q), \quad C_2 = \delta^2/(\mu_p q), \quad a_u = e^{V_0}, \quad \text{and} \quad a_v = e^{-V_0},$$

(2.27)

with the boundary data

$$V_1 = U_1, \quad u_1 = -U_1, \quad v_1 = U_1 \quad \text{on} \quad \partial \Omega_D,$$

(2.28)

$$\partial_{\nu} V_1 = \partial_{\nu} u_1 = \partial_{\nu} v_1 = 0 \quad \text{on} \quad \partial \Omega_N.$$

(2.29)

Unique solvability of the linearized problem in $H^1(\Omega)$ is known ([17][p. 215], [24] for scalar equations, and [29]).

One can assume weak interaction [20], and let $Q = 0$. In the **unipolar** case (no
holes), one lets \( v_1 = 0 \), and then (2.25) decouples into

\[
\text{div}(a_u \nabla u_1) = 0 \text{ in } \Omega, \tag{2.30}
\]

\[
u_1 = -U_1 \text{ on } \partial \Omega_D, \quad \partial_\nu u_1 = 0 \text{ on } \partial \Omega_N. \tag{2.31}
\]

In the **bipolar case**, one uses the full system with \( u_1 = v_1 = 0 \) on \( \partial \Omega_D \).

Smallness of \( \lambda \) suggests letting \( \lambda = 0 \) in (2.22). The resulting asymptotic approximation is fully justified in [29], and is given by

\[
2\delta^2 \sinh(V_0) = C. \tag{2.32}
\]

Let \( \Gamma_1 \) be a part of \( \partial \Omega_D \) and \( \Gamma_0 = \partial \Omega_D \setminus \Gamma_1 \). In applications one can prescribe various \( U_1 \) on \( \Gamma_0 \) and measure the total flux across \( \Gamma_1 \). Thus, one wishes to find \( C \) from the mapping

\[
U_1 \mapsto \int_{\Gamma_1} \left( \frac{a_u \partial_\nu u_1}{C_1} - \frac{a_u \partial_\nu v_1}{C_2} \right), \tag{2.33}
\]

for all \( U_1 = 0 \) on \( \Gamma_1, U_1 \in H_{(1/2)}(\Gamma_0) \).

This inverse problem is hard to handle directly, since we do not have sufficient boundary or interior data [20]. It would be very desirable to formulate an “adjoint” inverse problem where better data for numerical analysis are defined, in a way where the new data can be computed easily and in a stable way from the original data. This approach is widely used in the inverse option pricing problem (Dupire’s equation) [19], where it enables one to obtain theoretical and numerical results. Also in [15], an adjoint boundary value problem was utilized in the numerical solution of the inverse conductivity problem.
with laser-beam induced current (LBIC imaging).

The proper adjoint problem was developed in [20], and for the weak interaction case, (2.30) becomes

\[ \text{div}(a_u \nabla u_1^*) = 0 \text{ in } \Omega, \]
\[ u_1^* = 1 \text{ on } \Gamma_1, \quad u_1^* = 0 \text{ on } \Gamma_0, \quad \partial_n u_1^* = 0 \text{ on } \partial \Omega_N. \] (2.34) (2.35)

Then \( V_0(x) = \ln a(x) \), and therefore \( C(x) \) is determined uniquely by (2.32) since the hyperbolic sine is a one-to-one function. The adjoint inverse problem is to find \( a_u \) from the extra data

\[ g_1^* = \partial_n u_1^* \text{ on } \Gamma_0. \] (2.36)

Indeed, it has been proven that the functional in (2.33) determines in a unique and stable way the data \( g_1^* \) on \( \Gamma_0 \), assuming that \( a_u \) and \( C_1 \) are known at the boundary—which, in fact, they are, by (2.27).

The boundary data represent applied voltages of one and zero at two separate Ohmic contacts on the boundary of the semiconductor device, with the remainder of the semiconductor device being insulated. The extra data for solving the adjoint inverse problem is simply a measurement of electrical current flow (amps). It is the author’s opinion that the simplicity of these boundary measurements is a key advantage of the inverse problem for equation (2.34). Most likely these measurements can be imposed directly in quality control environments with no more difficulty than finding the mapping in (2.33). Ideally, the current flow \( g_1^* \) could be measured quickly for efficient and reliable testing of semiconductor devices, using the numerical algorithms we have developed later in this
dissertation.

We will formulate appropriate boundary value problems based on (2.34) in the next two sections, which will be used to produce theoretical and numerical results.

2.6 Semiconductor boundary value problem

We will consider the case when a vertical profile of the semiconductor device is scanned with boundary measurements of voltage and current. One must analyze these measurements and reconstruct the doping profile in the interior of the given two-dimensional domain. This scenario seems to be the standard procedure for testing of semiconductor devices.

We will assume that the physical semiconductor device is the unit rectangle,

$$\Omega_p = (-1, 1) \times (0, 1),$$

(2.37)

in which the physical voltage potential $u_p$ satisfies in a weak sense the conductivity equation,

$$\text{div}(a_p(x)\nabla(u_p(x))) = 0, \ x \in \Omega_p.$$  

(2.38)

Voltage of 1 is prescribed at a part of the top face of the semiconductor device, depending on an electrical terminal length $L$,

$$u_p = 1 \text{ on } \Gamma_{p1} := (-L, L) \times \{1\}.$$  

(2.39)
Voltage of 0 is prescribed along the entire bottom edge of \( \Omega_p \),

\[
u_p = 0 \text{ on } \Gamma_{p0} := (-1, 1) \times \{0\}.
\] (2.40)

The lateral faces (left and right components of \( \partial \Omega_p \)) of the physical device are insulated from any flow of electrical current, and the top face is also insulated except at the terminal \( \Gamma_{p1} \). Thus we are given the remaining Neumann-type boundary data,

\[
\partial \nu u_p = 0 \text{ on } \Gamma_{pN} = \{-1, 1\} \times (0, 1) \cup ((-1, 1) \times \{1\}) \setminus \Gamma_{p1}
\] (2.41)

For purposes of identifying the doped region \( D_p \subset \Omega_p \), it is assumed that the conductivity coefficient is constant in \( D_p \) and in the complement of \( D_p \); that is,

\[
a_p(x) = \begin{cases} 
  a^+, & x \in \Omega_p \setminus D_p, \\
  a^-, & x \in D_p.
\end{cases}
\] (2.42)

It is usual to assume that \( 1 \leq a^+ < a^- \). In fact, by simple scaling of electrical units, one can assume that \( a^+ = 1 \) and \( a^- = k \), in which the conductivity equation and definition of \( a_p \) can be defined simultaneously in the compact form

\[
\text{div}((1 + (k - 1)\chi_{D_p})\nabla(u_p)) = 0 \text{ in } \Omega_p,
\] (2.43)

Above \( \chi_{D_p} \) is the characteristic function of \( D_p \), equal to zero outside of \( D_p \) and equal to one inside of \( D_p \).

In the above description of the semiconductor device, it is assumed that the doping
compounds entered through the upper surface of the semiconductor device and did not reach any other parts of $\partial D_p$. It is always assumed that $L$ is small enough so that the electrical terminal at $\Gamma_1$ is a subset of $\partial D$. To make these requirements precise, we write $\Gamma_{p1} \subset \partial D$, $\partial D \cap (-1,1) \times \{1\} \neq \emptyset$, $\partial D \cap (\Gamma_0 \cup \{-1,1\} \times (0,1)) = \emptyset$.

It is necessary for diffusion to occur beyond a radius of $L$ in order for our formulation to be valid. In our experiments we will most often use $L = \frac{1}{8}$ (this is $1/8$th of the length of the semiconductor device in scaled units). Such a choice seems very adequate since masking of semiconductor wafers usually permits diffusion across much more than $1/8$th the width of a semiconductor device.

2.7 Reflected boundary value problem

To find $u_p$ numerically, one must find a way to deal with the two sharp corner points at $\partial D_p \cap \partial \Omega_p$, under the assumption that dopant diffusion is orthogonal to $\partial \Omega_p$ within a small boundary layer near $\partial \Omega_p$. This is difficult to do. However, the same orthogonality property makes it easy to reflect the domain and make these possible corners disappear completely. Since $u_p$ has zero normal derivative along $((-1,1) \times \{1\}) \setminus \Gamma_{p1}$ we can make an even reflection $u_r(x_0 + h) = u_p(x_0 - h)$, $x_0 \in (-1,1) \times \{1\}$, $x_0 - h \in \Omega_p$.

It is convenient to re-center the resulting reflected domain $\Omega_r$ so that we obtain the unit square. Almost all of our numerical simulations will use the reflected domain $\Omega_r$. Thus, we have reserved the notation $\Omega_p$ for the physical domain, and we will henceforth simply refer to the reflected domain by $\Omega := \Omega_r$, i.e., $\Omega = (-1,1)^2 \setminus \Gamma_1 := (-L,L) \times \{0\}$, and refer to the reflected solution by $u := u_r$. Likewise, we extend $D_p$, $a_p$ by reflection.
to obtain $D_r$ and $a_r$, and for convenience define $D := D_r$, $a := a_r$.

By the reflection principle [1], $u$ is piecewise harmonic in the interior of the level curves of $a$, and thus $u$ is a weak solution to

$$\text{div}((1 + (k - 1)\chi_D)\nabla(u)) = 0 \text{ in } \Omega.$$  \hfill (2.44)

The corresponding boundary conditions for $u$ are likewise simplified; in particular, we have Neumann-type data at only the symmetric lateral faces of $\partial \Omega$.

$$\partial_\nu u = 0 \text{ on } \Gamma_N := \{-1, 1\} \times (-1, 1).$$  \hfill (2.45)

The Dirichlet data are prescribed at $\Gamma_1$, and at the horizontal faces of $\partial \Omega$,

$$u = 0 \text{ on } \Gamma_0 := (-1, 1) \times \{-1, 1\}, \quad u = 1 \text{ on } \Gamma_1.$$  \hfill (2.46)

Note that if we label the open unit interval $I = (-1, 1)$, then the four faces of $\partial \Omega$ can be more conveniently defined as $\Gamma_N = \partial I \times I$ and $\Gamma_0 = I \times \partial I$.

Further decomposition of $u$ can be performed by separating $u$ into piecewise components which are harmonic in $D$ and its open complement. This procedure is based on jump relations at $\partial D$, and is introduced in the chapter about numerical methods.
Altogether, we will prove the following uniqueness results for doping profiles and conductivity coefficients. The results of the first section were recently published [34].

1. We will prove global uniqueness for any doping profile \( C \) when the Dirichlet-to-Neumann map is known. This result is new, but uses ideas similar to [19, Section 5.6 Nonlinear equations]. Our result provides a proof of uniqueness for the source term in a class of semilinear elliptic equations with additional applications to areas outside of semiconductor theory.

2. Local uniqueness for smooth \( D \). This result is based on index theory and the same methods from complex analysis used in [2, 19, 35]. However, we extend the result to boundary conditions of mixed type using the same technique as for the case of polygonal \( D \) below. Additionally, we allow \( \partial D \) to have non-empty intersection with \( \partial \Omega \). This generalization makes our results much more applicable to a typical semiconductor device, which may have doped regions extending to the boundary.
of the device.

3. We will prove that conductivity coefficients $a(x) = 1 + k\chi_D(x)$ are uniquely determined for a wide class of polygonal domains. There are three new features to our result that have not been obtained before.

(a) We consider a rectangular semiconductor device $\Omega$ and simple piecewise constant boundary data. Previous results required additional smoothness on both $\partial \Omega$ and much more complicated boundary data to prove uniqueness.

(b) Our problem is a mixed problem with partial Dirichlet data and partial Neumann data. Previous results were only available for data of a single type.

(c) $D$ is allowed to have contact with the boundary. Previous theoretical results required $D$ to be compactly contained in $\Omega$.

### 3.1 Global uniqueness of source from many measurements

Here we prove uniqueness of source for an important class of semilinear elliptic partial differential equations useful for identification of doping profiles [12] in semiconductor devices and for modeling of ion channels [26] in biological systems.

Suppose $\Omega \subset \mathbb{R}^2$ is a bounded, Lipschitz domain. We specify the following conditions on a class $\mathcal{C}$ of coefficients $c(x,u)$:

$$\mathcal{C} = \{c \in C^2(\Omega \times \mathbb{R}) : \partial_u c \geq 0, \partial_{uu}^2 c \neq 0, c(x,0) = 0\}. \quad (3.1)$$
Define
\[ \mathcal{F} = L^\infty(\Omega), \]
\[ \mathcal{G} = \{ g \in L^\infty(\partial \Omega) \cap H^{1/2}(\partial \Omega) : g \geq 0 \}, \]
and
\[ \mathcal{U} = H^1(\Omega) \cup L^\infty(\Omega). \] (3.2)

For \( f \in \mathcal{F} \) and \( g \in \mathcal{G} \), we define the operators
\[ S_f : \mathcal{G} \to \mathcal{U} \] (3.3)
by solving the following Dirichlet boundary value problem for \( u \),
\[ -\Delta u(x) + c(x, u(x)) = f(x) \text{ in } \Omega, \]
\[ u = g \text{ on } \partial \Omega, \] (3.4)
and letting
\[ S_f(g) = u. \] (3.5)

According to [19, 40, pp. 154-155, Ch. 14], the solution \( u \in \mathcal{U} \) exists and is unique. For the same \((f, g, u)\) we also define the associated Dirichlet-to-Neumann map
\[ \Lambda_f : \mathcal{G} \to H^{-1/2}(\partial \Omega) \] (3.6)
by letting
\[ \Lambda_f(g) = \partial_\nu u \text{ on } \partial \Omega. \] (3.7)

The coefficient \( c \) has a specific meaning for many problems of mathematical physics, in particular inverse problems involving semiconductor devices [12] and ion channels [9, 26].
The complicated drift diffusion equations for negative and positive charge carriers in semiconductors can be reduced [33] to a form of (3.4) where \( c(x, u) = 2 \sinh(u) \). The source term \( f \) is the semiconductor’s doping profile.

**Theorem 2.** Let \( f_1, f_2 \in \mathcal{F} \). If \( c \in \mathcal{C} \) and \( \Lambda f_1 = \Lambda f_2 \), then \( f_1 = f_2 \).

Let us give an informal outline of our proof. Essentially, we show uniqueness of \( f \) by solving for \( f(x) \) in terms of \( S_f(0)(x) \). We assume \( f \) is not known, but that its Dirichlet-to-Neumann map \( \Lambda f \) is given. In physical terms this means that although the doping profile is unknown, we can impose voltage (Dirichlet data) and measure the resulting current (Neumann data) at the boundary of the semiconductor device.

We use Dirichlet boundary data which depend on a parameter \( \tau \). Then differentiation with respect to \( \tau \) eliminates dependency on \( f \) since \( f \) does not depend on \( \tau \). In fact, the derivative \( u_\tau \) of the voltage potential \( u \) with respect to \( \tau \) satisfies the linearization of (3.4) about the solution \( S_f(0)(x) \), with an unknown coefficient \( \tilde{c}(x) = \partial_u c(x, S_f(0)(x)) \).

Next, we show that \( \Lambda f \) determines the Dirichlet-to-Neumann map for \( u_\tau \). It is a known result that the Dirichlet-to-Neumann map of the linearized equation for \( u_\tau \) determines \( \tilde{c}(x) \). Monotonicity of \( \partial_u c \) is used to recover \( u(x) = S_f(0)(x) \). From the partial differential equation for \( u \) we can recover \( f \), thus showing uniqueness.

**Remark 2.** In general, when \( c \not\in \mathcal{C} \), uniqueness of \( f \) does not hold. Consider, for example, the case when \( c(x, u) \) is linear with respect to \( u \). Let \( \Lambda f_1 = \Lambda f_2, g \in \mathcal{G}, u_1 = S_{f_1}(g), u_2 = \)
$S_{f_2}(g)$. Then for $f = f_2 - f_1$ and $v = u_2 - u_1$, we have

$$
-\Delta v + c(x,v) = f \quad \text{in } \Omega,
$$

$$
v = 0 \quad \text{on } \partial \Omega,
$$

$$
\partial_{\nu}v = 0 \quad \text{on } \partial \Omega.
$$

(3.8)

There is an infinite-dimensional set of solutions $(\Omega, f, c, v)$ for (3.8) such that $f \not\equiv 0$. To see this, let $w \in C^2_0(\Omega)$ and let $f = -\Delta w + c(x,w)$. Then $(\Omega, f, c, w)$ is a solution for (3.8), and yet $f$ need not be zero. Thus, $f_1$ need not equal $f_2$, so uniqueness of source does not hold.

**Proof of Theorem 2.** Let $f \in \mathcal{F}$ and assume that $\Lambda_f$ is given. Let $g \in \mathcal{G}$, and define

$$
u_1 = S_f(0), \nu_2(; \tau) = S_f(\tau g), \quad \tau \in (0,1).
$$

(3.9)

As a consequence of the maximum principle [40] and the fact $\tau g \geq 0$, we have $\nu_2(; \tau) \geq \nu_1$. We define a coefficient $c^*(x, \theta; \tau)$:

$$
c^*(x, \theta; \tau) = c(x, (u_1 + \theta(u_2 - u_1))(x; \tau)).
$$

(3.10)

Then by the Chain Rule,

$$
\partial_\theta c^*(x, \theta; \tau) = (u_2 - u_1)(x; \tau)\partial_\theta c(x, (u_1 + \theta(u_2 - u_1))(x; \tau)).
$$

(3.11)

Taylor’s formula in the integral form gives $c^*(x, 1; \tau) = c^*(x, 0; \tau) + \int_0^1 \partial_\theta c^*(x, s; \tau)ds$, and
thus
\[ c^*(x, 1; \tau) - c^*(x, 0; \tau) = (u_2 - u_1)(x; \tau) \int_0^1 \partial_u c(x, (u_1 + s(u_2 - u_1))(x; \tau))ds. \]  (3.12)

Now for \( u_1 \) and \( u_2 \) we have the equations
\[
\begin{align*}
-\Delta u_1(x) + c(x, u_1(x)) &= f(x), \\
-\Delta u_2(x; \tau) + c(x, u_2(x; \tau)) &= f(x).
\end{align*}
\]  (3.13)

Upon subtracting the equation for \( u_1 \) from the equation for \( u_2 \), we obtain
\[
-\Delta (u_2 - u_1)(x; \tau) + c(x, u_2(x; \tau)) - c(x, u_1(x)) = 0. \]  (3.14)

Using \( c^* \) for \( \theta = 0, 1 \), we find \(-\Delta (u_2 - u_1)(x; \tau) + c^*(x, 1; \tau) - c^*(x, 0; \tau) = 0 \), and upon using identity (3.12) for \( c^*(x, 1; \tau) - c^*(x, 0; \tau) \), we obtain
\[
\Delta (u_2 - u_1)(x; \tau) = (u_2 - u_1)(x; \tau) \int_0^1 \partial_u c(x, (u_1 + s(u_2 - u_1))(x; \tau))ds. \]  (3.15)

Let us consider the boundary value problem for the function
\[
v_\tau(x) = \frac{u_2(x; \tau) - u_1(x)}{\tau}. \]  (3.16)

For \( v_\tau \) we obtain the differential equation
\[
\Delta v_\tau(x) = v_\tau(x) \int_0^1 c_u(x, (u_1 + s(u_2 - u_1))(x; \tau))ds. \]  (3.17)
The boundary data are

\[ v_\tau = g \text{ on } \partial \Omega. \]  

(3.18)

Remember that \( \partial_u c \geq 0 \) and \( u_2(\tau) \geq u_1 \), thus making \( v_\tau \geq 0 \). Therefore, \( \Delta v_\tau \) is nonnegative since each factor of (3.17) is nonnegative, and so

\[ 0 \leq v_\tau \leq \sup_{\partial \Omega} v_\tau = \|g\|_\infty(\partial \Omega). \]  

(3.19)

Subtract the term \( v_\tau \partial_u c(,u_1) \) from both sides of (3.17). Noticing that \( -v_\tau \partial_u c(,u_1) = -v_\tau \int_0^1 \partial_u c(,u_1)ds \), we find

\[ \Delta v_\tau - v_\tau \partial_u c(,u_1) = v_\tau \int_0^1 [\partial_u c(,u_1 + s(u_2 - u_1)) - \partial_u c(,u_1)] ds. \]  

(3.20)

Let

\[ r^*(x;\tau) = v_\tau(x) \int_0^1 [\partial_u c(x,(u_1 + s(u_2 - u_1))(x;\tau)) - \partial_u c(x,u_1(x))] ds. \]  

(3.21)

Since \( \partial_u^2 c \) is continuous, there exists a constant \( C \) independent of \( x \) and \( \tau \) such that

\[ |\partial_u c(x,u_1(x)) + s(u_2(x;\tau) - u_1(x))) - \partial_u c(x,u_1(x))| \leq C\tau\|v_\tau\|_\infty(\Omega), \ s \in [0,1]. \]  

(3.22)

We have already established a bound for \( v_\tau \) using boundary data \( g \). Thus,

\[ \|r^*\|(1)(\Omega) \leq C\tau\|g\|(1/2)(\partial \Omega). \]  

(3.23)
The right side is a constant multiple of $\tau$, so clearly, $\|r^*\|(\Omega) \to 0$ as $\tau \to 0$. According to [19], the Dirichlet problem for $v_\tau$ is uniquely solvable for each $\tau > 0$.

Thus, we have proved the following lemma.

**Lemma 3.**

\[
\Delta v_\tau - v_\tau \partial_u c(, u_1) = r^* \text{ in } \Omega, \\
v_\tau = g \text{ on } \partial \Omega,
\]

where $\|r^*\|(\Omega) \to 0$ as $\tau \to 0$.

Now we will prove that the following equation for $v$ defines the limit of $v_\tau$ as $\tau \to 0$, in the sense of convergence in $H^1(\Omega)$.

Let $u_1 = S f(0)$, $\tilde{c}(x) = \partial_u c(x, u_1(x))$, and let $v$ solve

\[
-\Delta v + \tilde{c} v = 0 \quad \text{in } \Omega, \\
v = g \quad \text{on } \partial \Omega.
\]

According to [17, 19], the above linear Dirichlet problem for $v$ is uniquely solvable.

To prove that $v_\tau$ converges to $v$ in $H^1(\Omega)$ as $\tau \to 0$, we will let

\[
w = v_\tau - v.
\]

Then

\[
\Delta w - w \partial_u c(, u_1) = \Delta v_\tau - v_\tau \partial_u c(, u_1) - \Delta v + v \partial_u c(, u_1).
\]

40
The first two terms on the right are equal to $r^*$ and the last two are equal to 0. Thus,

$$\Delta w - w \partial_u c(u_1) = r^* + 0,$$

with the boundary condition

$$w = 0 \text{ on } \partial \Omega.$$

Hence,

$$\|w\|_{(1)}(\Omega) \leq C \|r^*\|_2(\Omega) \leq \|r^*\|_{(1)}(\Omega).$$

The right side goes to 0 as $\tau \to 0$.

We must now analyze the relationship between Dirichlet-to-Neumann maps for $u_1, u_2$ and $v$. Since $v_\tau$ converges to $v$, we also have

$$\partial_\nu v_\tau \to \partial_\nu v$$

in $H_{(-1/2)}(\partial \Omega)$. Since $\Lambda_f$ is given, $\partial_\nu v_\tau$ on $\partial \Omega$ are given as well, and so is their limit. In fact, equation (3.25) is the linearization of the equation for $S_f$ around the solution $u_1$.

Summing up, for any admissible Dirichlet data $g$ we are given the Neumann data for (3.25). Thus, the Dirichlet-to-Neumann map $\Lambda$ for $v$ is known. Referring to [19, Corollary 5.5.2], the coefficient $\tilde{c}(x)$ is determined by $\Lambda$ uniquely.

Furthermore, we can also uniquely determine $u_1(x) = S_f(0)(x)$. We use the fact that $c_u(x, u_1(x)) = \tilde{c}(x)$ is known as a function of $x$. By the condition $\partial_u^2 c \neq 0$, the function $\partial_u c(x, u)$ is one-to-one with respect to $u$ for each $x$. Let $x \in \Omega$. Then there is a unique $l \in \mathbb{R}$ such that $c_u(x, l) = \tilde{c}(x)$. Thus, $u_1(x) = l$. 
This proves the following lemma.

**Lemma 4.** The Dirichlet-to-Neumann map \( \Lambda_f \) determines \( S_f(0) \) uniquely.

Now we reach the conclusion of the proof of Theorem 2. Let \( \Lambda_{f_1} = \Lambda_{f_2} \) be given. According to Lemma 4, \( u = S_{f_1}(0) \) can be found uniquely, and so \( S_{f_2}(0) = u \) as well. By (3.4), \( f_1(x) = -\Delta u(x) + c(x, u(x)) = f_2(x) \) for all \( x \in \Omega \). Thus, \( f_1 = f_2 \).

\( \square \)

### 3.2 Local uniqueness of \( D \) from one measurement

Local uniqueness shall be established for an inverse domain identification problem.

Let \( I = (-1, 1) \), \( L \in (0, 1) \), \( \Gamma_1 = (-L, L) \), and define

\[
\Omega = \{x + iy : x, y \in I\} \setminus \Gamma_1.
\] (3.32)

The parts of \( \partial \Omega \) other than \( \Gamma_1 \) are denoted by \( \Gamma_0 = (I - i) \cup (I + i) \) and \( \Gamma_N = i\Gamma_0 \).

Let \( D_0 \subset \Omega \) be such that \( D_0 \cup \Gamma_1 = z_0(B) \), where \( B \) is the open unit ball in \( \mathbb{C} \). We assume that \( z_0 \) is a conformal mapping, and that \( \partial D_0 \in C^{1+\lambda} \) for some \( \lambda > 0 \). Also, we require that \( \overline{D_0} \subset \Omega \cup \Gamma_1 \), that \( D_0 \) is symmetric with respect to reflection across the real axis, and that \( z_0 \) is normalized in the canonical fashion by \( z_0(0) = 0 \), \( z_0'(0) > 0 \).

Now consider the family \( \Psi_M \) of functions \( \psi \) such that \( \psi \) is analytic in \( B \), \( \psi(0) = 0 \), \( \|\psi\|_{1+\lambda}(B) \leq M \), and \( z_\psi = z_0 + \psi \) is a conformal mapping normalized by \( z_\psi'(0) > 0 \). Let

\[
D_\psi = z_\psi(B) \setminus \Gamma_1
\] (3.33)
for all $\psi \in \Psi_M$. There exists $\epsilon_\Omega > 0$ such that if $M < \epsilon_\Omega$ then $\overline{D_\psi} \subset \Omega \cup \Gamma_1$ for all $\psi \in \Psi_M$.

We define the functions $u_\psi$, $\psi \in \Psi_M$ by

$$u_\psi(z) = \begin{cases} u^i_\psi(z), & z \in D_\psi, \\ u^e_\psi(z), & \text{otherwise,} \end{cases}$$

(3.34)

where $u^i_\psi$ and $u^e_\psi$ are harmonic functions in $\Omega \setminus \overline{D_\psi}$ and $D_\psi$, respectively. Boundary conditions on $\partial \Omega$ are given by $u^e_\psi|_{\Gamma_0} = \partial_\nu u^e_\psi|_{\Gamma_N} = 0$ and $u^i_\psi|_{\Gamma_1} = 1$. The transmission condition on $\partial D_\psi \setminus \Gamma_1$ will play an important part in what is to come, and it is given by

$$u^i_\psi = u^e_\psi,$$

(3.35)

$$\partial_\nu u^e_\psi = k \partial_\nu u^i_\psi.$$ 

(3.36)

We observe that $u_0(z) = u_0(z)$ from symmetry of $D_0$ and due to invariance of $\Gamma_0, \Gamma_1$, and $\Gamma_N$ with respect to complex conjugation.

**Remark 3 (Regularity of $u_\psi$ on $\partial D_\psi$).** When $\partial D_\psi \in C^{1+\lambda}$, the problem for $u_\psi$ has a unique solution $u_\psi \in H^1(\Omega)$ such that $u^e_\psi \in C^{1+\lambda}(\Omega \setminus \overline{D_\psi})$ and $u^i_\psi \in C^{1+\lambda}(\overline{D})$ (see [14, 31]). $\nabla u_\psi \neq 0$ on $\partial D_\psi$ is to be understood in the sense that the tangential derivative or both of the interior and exterior normal derivatives have nonzero value at $\partial D_\psi$.

Let us consider the gradient of $u_0$ in the domain $\Omega_p = \Omega \cap \mathbb{C}^+ (\mathbb{C}^+ := \{z : \Im z > 0\})$. If the sets $E_i = \{z \in \partial \Omega_p : \partial_\nu u_0(z) \leq 0\}$ and $E_o = \{z \in \partial \Omega_p : \partial_\nu u_0(z) \geq 0\}$ (representing input and output of electrical current, respectively) are connected, then $\nabla u_0(z) \neq 0$ for
all \( z \in \Omega_p \) due to results of geometric index theory [3].

We know from the maximum principle that \( \Gamma_1 \subset E_i \) and \( \Gamma_0 \cap \mathbb{C}^+ \subset E_o \). Moreover, \( \Gamma_0, \Gamma_1 \) together with \( \Gamma_N \cap \mathbb{C}^+ \) form the entire boundary of \( \Omega_p \), and \( \Gamma_N \) is a subset of both \( E_i \) and \( E_o \). Thus, \( E_i \) is connected and so is \( E_o \). Hence, \( \nabla u_0 \neq 0 \) in \( \Omega_p \).

By the previously mentioned fact, \( u_0(z) = u_0(z) \), we observe that \( \nabla u_0 \neq 0 \) in \( \Omega \) as well. Since \( \partial \Omega \cap \partial D_0 \) is empty, there is a compact set \( K \subset \Omega \) such that \( (\partial D_0 \setminus \Gamma_1) \subset (K \setminus \partial K) \). In particular, \( \nabla u_0 \neq 0 \) in \( K \). Furthermore, \( \nabla u_0 \) is continuous in \( \Omega \setminus \partial D_0 \) and the tangential, inner- and outer-normal derivatives are separately continuous on \( \partial D_0 \setminus \Gamma_1 \), and thus from compactness of \( K \) and \( \partial D_0 \setminus \Gamma_1 \), there is \( \epsilon_G > 0 \) such that if \( M < \epsilon_G \) then \( \nabla u_\psi \neq 0 \) in \( K \). (See Remark 3 for the generalized meaning of nonzero gradient on \( \partial D_\psi \).)

Let \( \epsilon_A = \min\{\epsilon_\Omega, \epsilon_C, \epsilon_G, \epsilon_L\} \), where \( \epsilon_C > 0 \) is such that if \( M < \epsilon_C \) then \( \Omega \setminus (\overline{D_0} \cup \overline{D_\psi}) \) is connected, and \( \epsilon_L > 0 \) is such that if \( M < \epsilon_L \) then \( \partial(D_0 \cap D_\psi) \) satisfies a Lipschitz condition with a constant to be specified in the proof of Theorem 7. From henceforth we shall assume \( M < \epsilon_A \). By \( C \) we shall indicate a constant which depends only on \( \Omega, D_0 \), and \( M \).

The additional data \( g_i \), to be used for the so-called inverse conductivity problem of identifying unknown \( D_i \) is given by

\[
  g_i(D_\psi) = \partial_\nu u_\psi \text{ on } \Gamma_0 \cap \mathbb{C}^+.
\]

We shall assume henceforth that \( \partial_\nu u_\psi(z) = g_i(z; \overline{D_\psi}) \) on \( \Gamma_0 \cap \mathbb{C}^- \); i.e., the data from the “physical” part of the semiconductor is reflected onto the “virtual” part.

**Lemma 5.** There exists a harmonic conjugate \( v_\psi^c \in C^{1+\lambda} \) to \( u_\psi^c \), which is unique up to a
constant.

Proof. Let \( R \) be the intersection of the real line with \( \Omega^+ := \Omega \setminus D_\psi \), and let \( A^\pm \) be the components of \( \Omega^+ \) which are separated by \( R \). Clearly, \( A^+ \) and \( A^- \) are each connected, and so there exists \([2, 37]\) a pair of harmonic conjugates \( v^e_\psi \pm \in C^{1+\lambda}(A^\pm) \) to \( u^e_\psi \), unique up to constants.

Due to the symmetry of \( \partial \Omega \) and the boundary data, as we see that \( \partial_\nu u^e_\psi = 0 \) on \( R \). Therefore \( \partial_\tau v^e_\psi \pm = 0 \) on \( R \) and so both \( v^e_\psi \pm \) are constant on \( R \). Hence, one can assume that \( v^e_\psi \pm \) are normalized by constants such that \( v^e_\psi^+ = v^e_\psi^- \) on \( R \).

The Cauchy-Riemann equations on \( R \), and the fact that \( v^e_\psi \) is harmonic throughout \( \Omega^+ \) yield

\[
\partial_\nu v^e_\psi^+ = \partial_\tau u^e_\psi^+ = \partial_\tau u^e_\psi = 0 = \partial_\tau u^e_\psi^- = \partial_\nu v^e_\psi^-.
\]

Thus, the extension of \( v^e_\psi \pm \) onto the whole of \( \Omega^+ \), given by \( v^e_\psi = v^e_\psi^+ \chi_{A^+} + v^e_\psi^- \chi_{A^-} \), is well-defined and analytic.

Let \( N = \{ \psi \in \Psi_M : g_i(\psi; D_\psi) = g_i(\psi; D_0) \} \). We define the multiply-continued function \( u^e \) by \( u^e = u^e_\psi \) on \( \Omega \setminus D_\psi \) for all \( \psi \in N \). This definition of \( u^e \) is consistent by uniqueness of continuation, and moreover \( u^e \) is harmonic in its domain of definition. It is given that \( \partial D_0 \in C^{1+\lambda} \) and \( \| \psi \|_{1+\lambda} \leq M \), which imply \( \| u^e \|_{1+\lambda}(\Omega^+) \leq C \) for some \( C > 0 \). Moreover, there exists a harmonic conjugate \( v^e \) to \( u^e \) (by applying Lemma 5 to each \( u^e_\psi \)) that is unique if we normalize by \( v^e(x_0) = 0 \) for some fixed point \( x_0 \in \Omega^+ \). Let \( U^e = u^e + iv^e \).

By previous estimates for \( u^e \), we have

\[
\| U^e \|_{1+\lambda}(\Omega^+) \leq C. \tag{3.38}
\]
Lemma 6. A domain $D_\psi$ is a solution to the inverse conductivity problem with exterior data $u^e$ if and only if

$$\phi(t) = 3U^e_\psi(z_\psi(t)) + \overline{U^i_\psi(z_\psi(t))} \text{ when } |t| = 1$$

(3.39)

for some function $\phi$ that is complex analytic in $B$ and is in $H^1(\overline{B})$.

Proof. Let $D_\psi$ be a solution to the inverse conductivity problem; that is, $D_\psi$ and $u_\psi$ solve (3.37). We see that $\partial_r v^e_\psi = -\partial_r u^e_\psi = -2\partial_r u^i_\psi = 2\partial_r v^i_\psi$ on $\partial D$. By integrating along $\partial D$ and choosing the constant of integration to be zero, we see that $v^e_\psi = 2v^i_\psi$ on $\partial D$. This relation and continuity of the function $u_\psi$ yield

$$U^e_\psi + \overline{U^i_\psi} = U^i_\psi + \overline{U^i_\psi}, \quad (U^e_\psi - \overline{U^e_\psi}) = 2(U^i_\psi - \overline{U^i_\psi}) \text{ on } \partial D.$$  

(3.40)

By substituting $\overline{U^i_\psi}$ from the first relation into the second one we obtain (3.39) with $\phi(t) = 4U^i_\psi(z_\psi(t))$.

Now let (3.39) be valid for some domain $D_\psi$ with corresponding conformal map $z_\psi$, and let $U^e_\psi$ be the exterior analytic function generated by solving the direct conductivity problem. Then we have (3.39) for $U^e_\psi$ with $4U^i_\psi(z_\psi(t))$ in place of $\phi(t)$.

We need to show that the exterior data $u^e$ is uniquely determined by $D_\psi$. Now as we saw before, a solution to the inverse problem generates $\phi(t) = 4U^i(z_\psi(t))$. Therefore, we suppose there is another $\phi(t)$ and we let $U^i(z_\psi) = \phi(t(z_\psi))/4$.

Upon subtracting the relations (3.39) for both $U$ and $U_\psi$ and letting $U_d = U - U_\psi$,
we obtain

\[ 4U_d^i = 4U_d^e + 2\overline{U}_d^e \text{ on } \partial D. \] (3.41)

Since the real parts of \( U^e \) and \( U^e_\psi \) are the same on \( \partial \Omega \) we conclude that the real part of \( U_d^e = 0 \) on \( \partial \Omega \). Subtracting and adding the boundary relations (3.41) for \( U_d \) and \( \overline{U}_d \), we arrive at

\[ U_d^e - \overline{U}_d^e = 2(U_d^i - \overline{U}_d^i), \quad U_d^e + \overline{U}_d^e = U_d^i + \overline{U}_d^i \text{ on } \partial D, \] (3.42)

which mean exactly that

\[ u_d^e = u_d^i, \quad v_d^e = 2v_d^i \text{ on } \partial D. \] (3.43)

So \( u_d \) is a solution to the direct conductivity problem with zero Dirichlet data on \( \partial \Omega \), and it follows that \( u_d^e = 0 \) outside \( D \). Therefore \( u^e = u^e_\psi \) and the proof is complete.

\[ \square \]

**Theorem 7.** There is a number \( \epsilon_0(M) \) such that if \( \| \psi \|_0(B) < \epsilon_0 \) and \( g_i(\partial D_\psi) = g_i(\partial D_0) \) then \( \psi \equiv 0 \).

**Proof.** Now if \( N \) is empty, the proof is complete. Otherwise, let \( \psi \in N \), and define \( \Omega^+ = \Omega \setminus (D_0 \cup D_\psi) \). In local coordinates near \( \partial D \) (since \( \partial D \) is a rectifiable curve), the Cauchy-Riemann equations for \( u^e \) and \( v^e \) are \( \partial_\nu u^e = \partial_\nu v^e \), \( \partial_\nu v^e = -\partial_\nu u^e \) on \( \partial D \). We conclude \( \partial_\nu v^e = 2\partial_\nu v^i \) and so \( v^e = 2v^i + C \) (\( C \) is a real constant). This and continuity of \( u_\psi \) yield

\[ U^e + \overline{U}^e = U^i + \overline{U}^i, \quad U^e - \overline{U}^e = 2(U^i - \overline{U}^i) + iC \text{ on } \partial D. \] (3.44)

Substituting \( U^i \) from the first equation above into the second, letting \( z = z(t), \phi(t) = 4U^i(z(t)) + iC \), and taking the complex conjugate, we obtain
\[ 3U^e(z(t)) + \overline{U^e(z(t))} = \phi(t) \text{ on } \partial B, \quad (3.45) \]

where, according to Lemma 6 (c.f. [19]), \( \phi \) is a uniquely determined complex analytic function in \( B \) and \( \phi \in C^{1+\lambda}(\overline{B}) \). Thus we have proved that (3.45) is equivalent to our inverse problem for \( D \). By letting \( z = z_\psi \) and by using Taylor’s formula for (3.45) we obtain

\[ A(\phi, \psi)(t) = B\psi(t) \text{ on } \partial B, \quad (3.46) \]

where \( A(\phi, \psi) = \phi - 3a\psi - \overline{a}\psi \), \( a(t) = u_\psi^e(z_0(t)) \), \( B\psi = 3B_1 + \overline{B_1}\psi \), and \( B_1\psi = U^e(z_0 + \psi) - U^e(z_0) - u_\psi^e(z_0)\psi \). Here we are considering \( A \) as an operator from \( \Phi \times \Psi \) into \( C^\lambda(\partial B) \), with \( \Phi \) being defined as the space of functions \( \phi \in C^\lambda(\overline{B}) \) analytic in \( B \) and \( \Psi \) as the space of functions \( \phi \in \Phi \) satisfying \( \psi(0) = 0, \psi'(0) > 0, \|\psi\|_{1+\lambda}(B) \leq M \). From these conditions \( A \) is a continuous operator from \( \Phi \times \Psi \) onto its range \( R(A) \subset C^\lambda(\partial B) \).

We will prove \( A \) has continuous inverse \( A^{-1} \) from \( R(A) \) onto \( \Phi \times \Psi \) by first reducing \( A \) to canonical form, secondly using index theory of one-dimensional singular integral equations [32], and thirdly referring to some known estimates for Cauchy integral operators.

Let \( t(\theta) = e^{i\theta} \) and differentiate the composition \( z_0(t(\theta)) \) with respect to \( \theta \) to find

\[ \partial_\theta u^i = u_\psi^i z_0 t e^{i\theta}, \text{ and similarly } \partial_\theta u^e = u_\psi^e z_0 t e^{i\theta}. \]

From refraction conditions we have

\[ 4U^i(z_0) = 3U^e(z_0) + \overline{U^e(z_0)} + \text{const}, \quad z_0 \in \partial D. \quad (3.47) \]

Differentiate with respect to \( \theta \), use formulas above to find derivatives \( U^i_\theta, U^e_\theta \), and multiply...
through by $\frac{|z'|}{z_0}$ to find

$$3u_\psi^e(z_0(t)) = 4u_\psi^i(z_0(t)) + \frac{u_\psi^e(z_0(t))z'_0(t)t}{z'_0(t)t}.$$  \hfill (3.48)

If we substitute (3.48) into (3.46) and use the definition of $a(t)$, we can replace $A$ by $A_*$ and prove the desired invertibility result. Here we have

$$A_*(\phi^*, \psi)(t) = \phi^*(t) - a(t)\frac{z_0(t)}{z'_0(t)t}\psi(t) - a(t)\psi(t),$$  \hfill (3.49)

and $\phi^* = \phi - 4u_\psi^i(z_0)\psi$. Since our goal will be accomplished by continuing with $A_*$ in place of $A$, we will drop $*$ from now on.

Let $c(t) = z'_0(t)t$. Then the equation $A(\varphi, \psi) = f$ is equivalent to the system of two Riemann-Hilbert boundary value problems:

$$\phi(t) = \varphi(t) - a(t)\frac{z_0(t)}{z'_0(t)t}\psi(t) - a(t)\psi(t),$$  \hfill (3.50)

and

$$\psi(t) = -\frac{c(t)}{\bar{r}(t)}(\psi(t) - f(t)) + G(t),$$  \hfill (3.51)

From the maximum principle for $u_0$ and the previous definition $a(t) = u_\psi^e(z_0(t))$, we have $|a| \neq 0$ on $\partial B$, and by the previous results of index theory we have ind$(u_\psi^e; \partial D_0) = 0$, so (3.50) and (3.51) are Riemann-Hilbert problems of index $\lambda = -2, \mu = 2$, respectively. (See [32], section 40.) Hence, if $F = 0$, the homogeneous problem (3.50) has only the
solution $\psi = 0$, which forces $G$ to be zero as well. The homogeneous problem (3.51) with index $\mu = 2$ has solutions $\psi(t) = X(t)(C_0t^2 + C_1t + C_2)$, with $C_0 = \overline{C_2}$, $C_1 \in \mathbb{R}$, where $X(t)$ is the fundamental function for the homogeneous Riemann-Hilbert problem (3.51) [32, p. 103-104].

The condition $\psi(0) = 0$ yields $C_2 = C_0 = 0$. Now a solution to the homogeneous problem (3.51) is $\psi(t) = ic(t)$, which does not satisfy $\psi'(0) > 0$. Thus, the kernel of $A$ is $\{0\}$, so $A$ has an inverse $A^{-1}$ on $R(A)$. This is the result we desired.

Let us estimate $\phi$ by [32, (40.20)]:

$$
\phi(z) = \frac{1}{\pi i}X(z)S(z), \ |z| < 1,
$$

(3.52)

where

$$
S(z) = \int_{\partial B} \frac{h(t)}{t - z} dt, \ |z| < 1.
$$

(3.53)

Here,

$$
h(t) = i\frac{\Re(-iacf)}{acX^{-}}(t),
$$

(3.54)

where $X^{-}$ is the limit of $X$ at $\partial B$ from inside $B$. Referring to [41, Theorem 4.1], we find that the norms $\|a\|_\lambda(\partial B_1)$, $\|c\|_\lambda(\partial B)$, and $\|X\|_\lambda(B)$ are bounded by $C$. In addition, $|X| > 1/C$ on $B$. So by well-known estimates [41, Theorem 1.10] for Cauchy integral operators, $\|S\|_\lambda(B) \leq C\|h\|_\lambda(\partial B)$, we have

$$
\|\phi\|_\lambda(B) \leq C\|f\|_\lambda(\partial B).
$$

(3.55)
Next we estimate $\psi$ by its representation

$$
\psi(z) = \frac{X(z)}{2\pi i} \left( \int_{\partial B} \frac{k(t)dt}{c(t)X^-(t)(t-z)} + z^2 \int_{\partial B} \frac{t^{-2}k(t)dt}{c(t)X^-(t)(t-z)} \right) - \psi_2(z),
$$

(3.56)

$$
\psi_2(z) = \frac{z^2X(z)}{2\pi i} \int_{\partial B} \frac{k(t)dt}{c(t)X^-(t)t^3} - X(z)(C_0t^2 + C_1z + C_2),
$$

(3.57)

where $2k = (\varphi - f)c/\bar{a}$ and the constants $C_0, C_1, C_2$ are chosen so that $\psi(0) = 0$, $\psi'(0) > 0$, and $\|\psi\|_{1+\lambda}(B) \leq M$. The term $-f$ in $2k$ presents no problem since clearly $f$ can be bounded in terms of $f$ itself, but we must comment further on how to estimate $\varphi$ in terms of $f$. By repeating for $\varphi$ the same argument used for $\phi$, we bound the first two terms in the representation for $\psi$ and the first term representing $\psi_2$. To bound the last term representing $\psi_2$ we recall that $C_1, C_2$ are given by integrals over the boundary

$$
C_1 = \frac{\Re X(0)}{2\pi \Im X(0)} \int_{\partial B} \frac{k(t)dt}{c(t)X^-(t)t^3},
C_2 = -\frac{1}{2\pi i} \int_{\partial B} \frac{k(t)dt}{c(t)X^-(t)t^3},
$$

(3.58)

and thus $\|\psi\|_{\lambda}(B) \leq C\|f\|_{\lambda}(\partial B)$. This establishes the estimate

$$
\|\phi\|_{\lambda}(B) + \|\psi\|_{\lambda}(B) \leq 2C\|f\|_{\lambda}(\partial B).
$$

(3.59)

This is sufficient to show the inverse $A^{-1}$ is continuous from $R(A)$ onto $\Phi \times \Psi$.

We remember that $U_d^e = 0$ in $\Omega^+$, so $U^e(z_0 + \psi) - B_1\psi$ is the first-order Taylor polynomial for $U^e(z_0 + \psi)$ centered at $\psi = 0$ (i.e., perturbations $D_\psi$ of $D$ given by $\psi$ are centered at the domain $D_0$). Using Lipschitz regularity of $\partial\Omega^+$ and the previous estimate (3.38), we find $\|B_1\psi\|_0 \leq C\|\psi\|_0^{1+\lambda}$. 51
Again by (3.38) and the assumed bounds on \( z_0 \) and \( \psi \), namely \( \partial D_0 \in C^{1+\lambda}, \| \psi \|_{1+\lambda} \leq M \), we have \( \| B_1(\psi) \|_{\lambda} \leq C \). By the general-purpose interpolation inequality

\[
\| f \|_{\mu} \leq 2 \| f \|_0^{1-\mu/\lambda} \| f \|_{\lambda}^{\mu/\lambda}, \tag{3.60}
\]

we obtain

\[
\| B \psi \|_{\mu} \leq C \| \psi \|_{\lambda}^{1+\delta} \tag{3.61}
\]

for any \( \mu, 0 < \mu < \lambda^2/(1 + \lambda) \), with \( \delta = \lambda - \mu(1 + \lambda)/\lambda \).

Now since \( g_i(; D\psi) = g_i(; D_0) \), we know that \( \psi \) satisfies (3.46), so \( B \psi \in R(A) \). Since \( A^{-1} \) is continuous from \( R(A) \) onto \( \Phi \times \Psi \) we can write (3.46) as

\[
(\phi, \psi) = A^{-1}B\psi, \tag{3.62}
\]

where we interpret \( (\phi, \psi) \) as an element of \( \Phi \times \Psi \) with the norm

\[
\|(\phi, \psi)\| = \|\phi\|_{\mu}(B) + \|\psi\|_{\mu}(B), \ 0 < \mu < \lambda. \tag{3.63}
\]

Using (3.61) we obtain

\[
\| A^{-1}B\psi \| \leq C\| \psi \|_{\mu}^{1+\delta} \leq \frac{1}{2} \|(\phi, \psi)\| \tag{3.64}
\]

for \( \| \psi \|_{\mu} < \epsilon_1 \).

By using a standard interpolation inequality [18, Theorem 1.1.1] and the given constraint \( \| \psi \|_{1+\lambda} \leq M \) we conclude that \( \| \psi \|_{\mu} < \epsilon_1 \) as soon as \( \| \psi \|_0 \leq \epsilon_0 \) for some \( \epsilon_0(M) \).
For such $\epsilon_0$ we have
\[ \|(\phi, \psi)\| \leq \frac{1}{2} \|(\phi, \psi)\|, \] (3.65)
Therefore, $\psi = 0$ and $D_\psi = D_0$.

\[ \square \]

### 3.3 Global uniqueness of polygonal $D$ from one measurement

We now consider a physically-motivated problem for a semiconductor device. The problem arises from simplification of the drift-diffusion model for solving the inverse doping profile problem, which can be linked [33] to the inverse conductivity problem in a one-to-one correspondence, under the reasonable assumptions of zero space charge and low injection [8]. For the following problem we will prove a uniqueness result which to our knowledge was not yet known. Similar results have been discovered before, but not for the type of mixed boundary conditions which correspond to real-world experimental data.

Suppose $\Omega$ is a bounded, Lipschitz, piecewise $C^2(\mathbb{R}^2)$ domain. Let $\{\Gamma_0, \Gamma_1, \Gamma_N\}$ be a partition of $\partial \Omega$ into non-empty open parts so that $\Gamma_0$ and $\Gamma_1 \cup \Gamma_N$ are connected and $\Gamma_0 \cup \Gamma_1 \cup \Gamma_N = \partial \Omega$.

We shall define a class $\mathcal{D}$ of admissible subdomains by the following rules.

1. $D \in \mathcal{D} \rightarrow D$ is an open polygon, $D$ is a proper subset of $\Omega$, and $\partial D \cap \Gamma_0$ is empty.

2. $D_1, D_2 \in \mathcal{D} \rightarrow \partial D_1 \cap \partial \Omega = \partial D_2 \cap \partial \Omega$. 

53
3. \( D \in \mathcal{D} \rightarrow D \) is convex and \( \partial D \cap \partial \Omega \) is connected.

Remark 4. Condition 3 above is sufficient, but not necessary. It can be replaced by any other condition which ensures that \( D_1, D_2 \in \mathcal{D} \rightarrow \Omega \setminus (D_1 \cup D_2) \) is connected, and the same proof of the theorem will apply.

Let \( k \neq 1 \) and let \( u = u(\cdot; D) \in H^1(\Omega) \) be the unique weak solution to the conductivity equation,

\[
\text{div}((1 + (k - 1)\chi_D(x))u(x)) = 0 \text{ in } \Omega,
\]

with boundary values

\[
u = j \text{ on } \Gamma_j, \ j \in \{0, 1\},
\]

and

\[
\partial_N u = 0 \text{ on } \Gamma_N.
\]

The inverse problem is to determine \( D \) from additional data \( g_i(\cdot; D) \),

\[
g_i(\cdot; D) = \partial_N u(\cdot; D) \text{ on } \Gamma_0.
\]

**Theorem 8.** Let \( D_1, D_2 \in \mathcal{D} \). If \( g_i(\cdot; D_1) = g_i(\cdot; D_2) \), then \( D_1 = D_2 \).

**Proof.** Suppose \( D_1, D_2 \in \mathcal{D} \) and \( g_i(\cdot; D_1) = g_i(\cdot; D_2) \). To obtain a contradiction, we will assume that \( D_1 \neq D_2 \). By relabeling \( D_1 \) and \( D_2 \) if necessary, we may assume that there is a vertex \( x_1 \) of \( D_1 \) and \( \epsilon > 0 \) such that \( B = B(x_1, \epsilon) \subset \Omega \) is disjoint from \( \overline{D_2} \). An inclusion is sufficient for our needs, and so we are still permitted to replace \( \epsilon \) by a smaller positive constant, resulting in a smaller open ball \( B \). Let \( u_1 = u(\cdot; D_1) \) and \( u_2 = u(\cdot; D_2) \).
Let \( I = D_1 \cap B \). It is to be understood that if we replace \( B = B(x_1, \epsilon) \) with \( B = B(x_1, \epsilon_2) \) \((\epsilon_2 \leq \epsilon)\), that other objects dependent on \( B \), such as \( I \), will be replaced automatically.

**Remark 5.** Let \( \Gamma_2 \) be the (1-dimensional) interior of the closure of \( \Gamma_0 \cup \Gamma_N \). Then \( \{\Gamma_1, \Gamma_2\} \) are disjoint connected arcs, and the closure of their union forms the boundary \( \partial \Omega \) of the domain \( \Omega \). By the maximum principle [19][p. 91] for \( u_2 \), we see that \( \partial_\nu u_2(\cdot) \geq 0 \) on \( \Gamma_1 \). Similarly, and by equation (3.68), we see that \( \partial_\nu u_2(\cdot) \leq 0 \) on \( \Gamma_2 \). This observation will supply the necessary contradiction after we find a critical point of \( u_2 \).

Let \( S_a, S_b \) denote the open sides of \( \partial D_1 \) adjacent to \( x_1 \) and intersected with \( B \), and let \( \tau_a, \tau_b \) be the unit tangent vectors to \( S_a, S_b \) directed towards \( x_1 \), and let \( \nu_a, \nu_b \) be the unit normal vectors to \( S_a, S_b \), directed outward from \( I \).

Let \( u_1 \) be the restriction of \( u_1 \) to \( D_1 \).

Define \( C_a = \{x_1 + h\tau_a : h \in \mathbb{R}\} \) and \( P_A = \{x \in \mathbb{R}^2 : (x-x_1) \cdot \nu_a < 0\} \). Put \( L_a = C_a \cap B \) and \( A = P_A \cap B \).

Let us consider the following boundary value problem for the function \( u_A \):

\[
\Delta u_A = 0 \text{ in } A, \tag{3.70}
\]

\[
u_A = u_2 \text{ on } L_a, \tag{3.71}
\]

\[
\partial_\nu u_2 = k \partial_\nu u_A \text{ on } L_a. \tag{3.72}
\]

The Cauchy problem for Laplace’s equation in general cannot be solved, even locally [21]. However, if the data are analytic, as they are in (3.71) and (3.72), then the Cauchy-
Kowalevski theorem guarantees existence of a unique solution in a sufficiently small neighborhood (i.e., neighborhood relative to $A$) of the line segment $L_a \subset \partial A$ where the Cauchy data are prescribed (see [21, p. 78] for an explicit reduction of the Cauchy problem to standard form and a proof that a solution exists near all of $L_a$).

In particular, the problem for $u_A$ can be solved for some $\epsilon_2 > 0$ in the set $B(x_1, \epsilon_2) \cap A$. If we replace $\epsilon$ by $\min\{\epsilon, \epsilon_2\}$—i.e., if we replace $B = B(x_1, \epsilon)$ above by $B = B(x_1, \min\{\epsilon, \epsilon_2\})$—then we retain the sufficient condition on $B$ in the first paragraph of the proof. We also replace other objects which are dependent on $B$ as described in the second paragraph of the proof.

We remind that after replacement of $\epsilon$, the set $A$ then becomes $A = \{x \in \mathbb{R}^2 : (x - x_1) \cdot \nu_a < 0\} \cap B(x_1, \epsilon_2)$, and (3.70), (3.71), (3.72) has a unique solution $u_A$ in the set $A$.

Now let $\hat{u}_A = u_A - u_2$. We see that $\hat{u}_A$ is harmonic in $A$ and zero on $L_a$, which is the intersection of $B$ with the boundary of a half-plane. Thus by the reflection principle [1][Theorem 24, p. 172], $\hat{u}_A$ can be extended harmonically across $L_a$ by an odd reflection. The resulting harmonic extension is defined on the whole of $B$ since $L_a$ is a diameter of $B$. Denote the resulting extension once again by $\hat{u}_A$. Thus we can also extend $u_A$ harmonically to all of $B$ by the formula $u_A = \hat{u}_A + u_2$ (here, as well as afterwards, we shall denote the resulting extension of $u_A$ also by $u_A$).

Moreover, since $u_1^- = u_A$ on $S_a$ and $\partial_\nu u_1^- = \partial_\nu u_A$ on $S_a$, we conclude that $u_A = u_1^-$ on $I$ by unique analytic continuation. Therefore, $\partial_\nu u_2 = k\partial_\nu u_1^- = k\partial_\nu u_A$ on $\partial D_1 \cap B = S_a \cup S_b$ by known transmission conditions for $u_2$ and $u_1^-$. 

Let us summarize our findings so far.
1. $\Delta u_2 = 0$ on $B$.

2. $\Delta u_A = 0$ on $B$.

3. $u_A = u_1^-$ on $I \cup S_a \cup S_b$.

**Remark 6.** By virtue of the fact that $x_1$ is a vertex of $D_1$, both $\{\tau_a, \tau_b\}$ and $\{\nu_a, \nu_b\}$ form a basis of $\mathbb{R}^2$.

Now by transmission conditions, we have

$$\nu_q \cdot \nabla u_2(x) = k \nu_q \cdot \nabla u_A(x), \quad x \in S_q.$$  \hfill (3.73)

We use the $q$-notation to indicate that a statement holds for both $q = a$ and $q = b$.

Therefore,

$$\nu_q \cdot \nabla u_2(x_1) = k \lim_{x \to x_1, \ x \in S_q} \nu_q \cdot \nabla u_A(x).$$  \hfill (3.74)

In light of the remark (6) and the fact that $u_A \in C^\infty(B)$ we find

$$\nabla u_2(x_1) = k \nabla u_A(x_1).$$  \hfill (3.75)

Also, we have $u_1 = u_2$ on $S_q$, which means that

$$\tau_q \cdot \nabla u_2(x) = \tau_q \cdot \nabla u_A(x), \quad x \in S_q.$$  \hfill (3.76)

Therefore,

$$\tau_q \cdot \nabla u_2(x_1) = \lim_{x \to x_1, \ x \in S_q} \tau_q \cdot \nabla u_A(x).$$  \hfill (3.77)
In light of the remark (6) and the fact that $u_A \in C^\infty(B)$ we find

$$\nabla u_2(x_1) = \nabla u_A(x_1).$$

(3.78)

By subtracting equations (3.75) and (3.78), we find that

$$0 = (k - 1)\nabla u_A(x_1),$$

(3.79)

and hence $\nabla u_A(x_1) = (0, 0)$. Substitution into (3.78) yields

$$\nabla u_2(x_1) = (0, 0).$$

(3.80)

Now according to local power series representation for $u_2$ near $x_1$, the following estimate must hold asymptotically for $x \to x_1$ and some $C > 0$.

$$|u_2(x) - u_2(x_1)| \leq C|x - x_1|^2.$$  

(3.81)

This implies that either \( \{ x \in \partial \Omega : \partial_{\nu}u_2(x) \geq 0 \} \) or \( \{ x \in \partial \Omega : \partial_{\nu}u_2(x) \leq 0 \} \) is disconnected by geometric index theory [3]. However, we know by Remark 5 that these sets are actually equal to $\Gamma_1$ and $\Gamma_2$, which are connected, so we arrive at a contradiction.

\[ \square \]

Remark 7. The parts of the boundary defined in the proof above are physically motivated by the parts of a semiconductor device, so we think it would be helpful to provide more specific definitions of $\Gamma_0, \Gamma_1, \text{and } \Gamma_N$. 

58
Let \( l \in (0, 1) \) and \( h > 0 \). Define \( \Gamma_0 = (-1, 1) \times \{0\} \) (voltage sink), \( \Gamma_1 = (-l, l) \times \{h\} \) (voltage source), and \( \Gamma_N = ((-1, -l) \cup (l, 1)) \times \{h\} \cup \{-1, 1\} \times (0, h) \) (insulated boundary). Moreover, let \( \Gamma_2 = \Gamma_0 \cup \Gamma_N \). Then \( \{\Gamma_1, \Gamma_2\} \) are disjoint arcs, and the closure of their union forms the boundary \( \partial \Omega \) of the domain \( \Omega = (-1, 1) \times (0, h) \), which represents a semiconductor device in \( x_1 \)-scaled coordinates.
Chapter 4

Numerical analysis

We will treat two special cases of equation (1.2) and one general case. When $k$ is extremely large, our analysis will be based on an analytic expansion of an integral operator $S : H_{(-1/2)}(\Gamma) \to H^1(\Omega)$ as an infinite series with respect to the parameter $\epsilon = \frac{1}{k}$ so that the solution $u$ to (1.2) is of the form

$$u = u_0 + \sum_{j=1}^{\infty} \epsilon^j u_j. \quad (4.1)$$

If we let $k \to \infty$, or equivalently $\epsilon \to 0$, then we obtain the boundary value problem corresponding to $u = u_0$. We are permitted to do this since $S$ is analytic for $\epsilon$ in a neighborhood of zero.

If we assume $k$ is infinity, then we take the boundary value problem for the term $u_0$, which is equivalent to an obstacle problem. If we assume $k$ is large, but not infinite, then it is sufficient to use the approximation $u \approx u_0 + \epsilon u_1$. This is equivalent to a coupled system of obstacle problems. In the general case, we do not use the analytic expansion at
all, but solve a full transmission problem for $u$ by using continuity across $\partial D$ and jump relations of the normal derivatives of $u$ at $\partial D$ given by

$$\partial_\nu u^+ = k \partial_\nu u^-.$$  \hfill (4.2)

In the doped region $k$ is either large or small, depending on which unipolar case is considered. We will be assuming that the conductivity coefficient is larger in the doped region, i.e., $k > 1$. When the case of electron-facilitated conductivity in silicon is considered, $k > 1$ corresponds to doping by a trivalent atom like boron or aluminum (P-type doping), and when the case of holes is considered, $k > 1$ corresponds to doping by an atom with five valence electrons, like phosphorus, arsenic, or antimony (N-type doping).

### 4.1 Description of integral operator $S$

We will decompose the function $u$ as a sum of a harmonic part $u_h$ and refracted part $u_r$. We first impose zero boundary conditions for $u_r$ on $\partial \Omega$ by letting $u_h$ solve

$$\Delta u_h = 0 \text{ in } \Omega, \quad \partial_\nu u_h = 0 \text{ on } \Gamma_N,$$  \hfill (4.3)

$$u_h = 0 \text{ on } \Gamma_0,$$  \hfill (4.4)

$$u_h = 1 \text{ on } \Gamma_1.$$  \hfill (4.5)

The existence and uniqueness of the solution $u_h \in H^1(\Omega)$ is given in [4, 33].

61
Let $G$ be the Green’s function in $\Omega$ for the mixed problem (4.3)-(4.4), but instead of (4.5), satisfying

$$G = 0 \text{ on } \Gamma_1. \quad (4.6)$$

Note that $-\Delta_x G(x, y) = \delta(-y)$ in $\Omega$.

We can represent the refracted part $u_r$ as the single layer potential of $G$ over $\partial D$ with density $\phi_0$ dependent on the jump in conductivity coefficients at $\Gamma = \partial D$,

$$u_r(x) = \int_{\Gamma} \phi_0(y)G(x, y)d\Gamma(y). \quad (4.7)$$

The single-layer operator $S : L^{\infty}(\Gamma) \to H^1(\Omega)$ is given by

$$(S\phi)(x) = \int_{\Gamma} G(x, y)\phi(y)d\Gamma(y). \quad (4.8)$$

Equation (4.2), applied to $u = u_h + u_r$, reads

$$\partial_\nu(u_h + u_r^+) = k\partial_\nu(u_h + u_r^-), \quad (4.9)$$

or, $(k - 1)\partial_\nu u_h = (\partial_\nu u_r^+ - k\partial_\nu u_r^-)$.

Letting $\epsilon = \frac{1}{k}$, we have [22]

$$\int_{\Gamma} \partial_\nu(x)G(x, y)\phi_0(y)d\Gamma(y) + \frac{1 + \epsilon}{2(1 - \epsilon)}\phi_0(x) = -\partial_\nu u_h(x), \ x \in \Gamma. \quad (4.10)$$

The left side of (4.10) can be viewed as a linear continuous operator $A(\epsilon)\phi_0$ from $H_{(-1/2)}(\Gamma)$ into itself (see [30], p. 203). Obviously $A(\epsilon)$ is an analytic function of $\epsilon$.
(with values in the Banach space of continuous linear operators from \( H_{(-1/2)}(\Gamma) \) into \( H_{(-1/2)}(\Gamma) \)). It is known [20] that this operator is invertible at \( \epsilon = 0 \) (because it corresponds to the mixed boundary value problem in \( \Omega \setminus \overline{D} \) with Neumann data on \( \Gamma \) and \( \Gamma_N \) and Dirichlet data on \( \Gamma_0 \)). According to the theory of analytic functions with values in Banach spaces, the inverse \( A^{-1}(\epsilon) \) does exist and is analytic when \( |\epsilon| < \epsilon_0 \), for some \( \epsilon_0 > 0 \). Since the right side is also analytic at \( \epsilon = 0 \), the solution \( \phi_0 \) is analytic near the origin; i.e., it is the sum of the convergent series

\[
\phi_0(x; \epsilon) = \phi_{00}(x) + \epsilon \phi_{01}(x) + \cdots + \epsilon^k \phi_{0k}(x) + \cdots
\]

(4.11)

in the space \( H_{(-1/2)}(\Gamma) \). By using properties of single-layer potentials [30], \( u \) is analytic near \( \epsilon = 0 \), i.e.,

\[
u(x; \epsilon) = u_0(x) + \epsilon u_1(x) + \cdots,
\]

(4.12)

where \( u_k \in H^1(\Omega) \) [20, 24, Theorem 3.1] and the series is convergent in this space when \( |\epsilon| < \epsilon_0 \).

### 4.2 Problems for various doping levels

We will analyze the boundary value problems for \( u_0 \) and \( u_1 \) in more detail in the following sections about superconductivity and high conductivity. Many pages of boundary value problems can be derived by inserting various partial sums of the series (4.12) into the
known jump relation for transmission problems [20], which in terms of $\epsilon$ reads

$$ \partial_{\nu} u^- = \epsilon \partial_{\nu} u^+, $$  \hspace{1cm} (4.13)

which must be true for all sufficiently small $\epsilon$. Thus, our main tool will be to form equalities between coefficients of equal powers of $\epsilon$.

Remark 8. We principally analyze the unipolar case of holes. The unipolar case of electrons may include the extreme case $k = 0$ (the dual analogue of $k = \infty$ for holes).

In order to include this case, we can write the transmission jump condition in a more general form,

$$ a^+ \partial_{\nu} u^+ = a^- \partial_{\nu} u^-, $$  \hspace{1cm} (4.14)

where $a^-$ is the conductivity coefficient inside the doped region, and $a^+$ is the conductivity coefficient outside. This equation makes sense assuming only that real constants $a^+$ and $a^-$ are not both zero.

All inverse problems use the data $g_1(\cdot; D, k) = \partial_{\nu} u(\cdot; D, k)$ on $\Gamma_0$, but the direct problems are solved with approximations of $u$, except in the case of the full transmission problem, which does not use the asymptotic expansion developed from the integral operator $S$. The superscripts - and + below will indicate the restriction of a function to either $D$ or $D$’s complement, respectively. Also, for symmetry and shortness of notation, we use $D^+$ to indicate the complement of $D$ in $\Omega$, and $\Gamma$ indicates the boundary of $D$.

- **Superconductivity**

  Conductivities of conductors range from $10^4$ to $10^6$ siemens per centimeter, but
the conductivity of a semiconductor device may be as low as $10^{-10}$ siemens per centimeter. Therefore, relative superconductivity of some regions in a semiconductor device can occur when a semiconductor device is overdoped, changing the characteristics of the doped region from a semiconductor to a conductor.

In the superconductive case we let $k \to \infty$, the conductivity problem breaks down to an obstacle problem, and we use the approximation $u \approx u_0$.

First, let’s analyze the boundary value problem for $u_0^-$, which can be solved explicitly.

\[ \Delta u_0^- = 0 \text{ in } D^-, \quad (4.15) \]
\[ \partial_{\nu} u_0^- = 0 \text{ on } \Gamma, \quad (4.16) \]
\[ u_0^- = 1 \text{ on } \Gamma_1. \quad (4.17) \]

A harmonic function with zero Neumann data has only the constant as a solution. Thus, $u_0^-$ is constant, and from the data on $g_1$ we determine that $u_0^-$ is identically equal to 1,

\[ u_0^- \equiv 1. \quad (4.18) \]

Next is the boundary value problem for $u_0^+$. We determine that

\[ \Delta u_0^+ = 0 \text{ in } D^+, \quad (4.19) \]
\[ \partial_{\nu} u_0^+ = 0 \text{ on } \Gamma_N, \quad (4.20) \]
\[ u_0^+ = u_0^- (= 1) \text{ on } \Gamma, \]  
(4.21)  

\[ u_0^+ = 0 \text{ on } \Gamma_0, \]  
(4.22)  

and finally, the extra data for the inverse problem,

\[ \partial_\nu u_0^+ = g_1 \text{ on } \Gamma_0. \]  
(4.23)  

We show the united solution \( u_0 \) in Figure 4.1. The flat region corresponds to an area of high conductivity where current easily flows and voltage change is minimal. The high conductivity is caused by the addition of doping agents to the base semiconductor substrate.

- High conductivity

When \( k \) is very large—for some semiconductor devices, the literature values for \( k \) are on the order of \( 10^3 \)—we can drop higher order terms and use the approximation \( u \approx u_0 + \epsilon u_1 \).

The boundary value problems for \( u_0 \) are the same as above. Let us write those for \( u_1 \).

\[ \triangle u_1^- = 0 \text{ in } D^-, \]  
(4.24)  

\[ u_1^- = 0 \text{ on } \Gamma_1, \]  
(4.25)  

\[ \partial_\nu u_1^- = \partial_\nu u_0^+ \text{ on } \Gamma, \]  
(4.26)
Figure 4.1: The zeroth order reconstruction of the voltage inside of a semiconductor.
\[ \Delta u_1^+ = 0 \in D^+, \quad (4.27) \]
\[ u_1^+ = u_1^- \text{ on } \Gamma, \quad (4.28) \]
\[ \partial_n u_1^+ = 0 \text{ on } \Gamma_N, \quad (4.29) \]
\[ u_1^+ = 0 \text{ on } \Gamma_0, \quad (4.30) \]

The resulting approximation \( u_0 + \epsilon u_1 \) is graphed in Figure 4.2, corresponding to \( k \approx 2 \). Note that the approximation is discontinuous at \( \Gamma \), but its normal derivatives coincide.

Figure 4.2: The first order reconstruction of the voltage inside of a semiconductor.
• Moderate conductivity

When conductivity is closer to 1, we cannot neglect any terms, so we do not use the analytic expansion at all. Instead, we solve the full transmission problem given by the jump condition (4.2) for normal derivatives at $\Gamma$ and by continuity of $u$ across $\partial D$. Moreover, the analytic expansion is only valid when $\epsilon$ is in some neighborhood of 0, so solving the full transmission problem is the only option that we have.

The full transmission model is also more accurate for high conductivity, provided that the direct problem is solved with a high-quality numerical algorithm.

Note that we do not have the choice to use the full transmission problem to model the case of superconductivity. Difficulties occur when trying to solve the transmission problem at extremely large values of $k$. We think that the observed errors are similar to Gibb’s overshoot phenomenon (when Fourier expansion of discontinuous functions results in values higher or lower at a jump than the original function’s values). To some extent this can be circumvented by increasing the concentration of source points near locations where transmission-type boundary conditions are imposed, namely $\partial D$.

Let us use this as an opportunity to show the “circus-tent” behavior exhibited by the numerical solution as the number of source points are increased. With a fine grid the solution is more or less accurate, but with fewer grid points the graphs of the numerical solutions (left of Figure 4.3) clearly contradict the Maximum Principle.

The numerical methods we have developed take all of these issues under consideration for a solution of the highest possible accuracy.
4.3 Algorithm for identifying the obstacle $D$

We propose to identify the area, shape, and location of the obstacle $D$ by using a three-step algorithm.

1. Generate initial data for inverse problem.
   
   (a) Impose potential $u = u_D$ on $\Gamma_0$.
   
   (b) Measure the initial data for the inverse problem $\partial_\nu u = u_N$ on $\Gamma_0$.

2. Compute the area and radius of $D$.
   
   (a) Create artificial obstacle $D_r$ which is circular and has radius equal to $r$.
   
   (b) Solve the direct problem with obstacle $D_r$ and the same potential $u(;r) = u_D$ on $\Gamma_0$ and find the resulting function $\partial_\nu u(;r) = u_N(;r)$ on $\Gamma_0$.
   
   (c) Find the radius $r$ that minimizes $\|u_N(;r) - u_N\|$. (Here the norm $\|\cdot\|$ is in $L_\infty$ or another norm which works well in actual numerical experiments.)
(d) Fix this value of $r$ and the area $\pi r^2$ of the obstacle $D_r$.

3. Find the exact boundary of $D$ by using perturbations which preserve the computed area of $D$.

(a) Represent $D$ as deformations of the first approximation $D_r$ ($r$ is fixed, resulting from the previous step of the algorithm).

(b) Use an orthogonal coordinate system for the boundary of $D$.

$$\partial D := \gamma := \{ <x, y> : <x, y> = (r + h(s)) <\sin(s), \cos(s)> \}$$  \hspace{1cm} (4.31)

(c) The first approximation $h(s) = 0$ corresponds to $D_r$.

(d) For better approximations take $h(s)$ as a linear combination of basis functions such as sines and cosines

$$h(s) = \sum_{k=-n}^{n} (a_k \sin(kx) + b_k \cos(kx)) := a \cdot B_{\sin} + b \cdot B_{\cos}$$  \hspace{1cm} (4.32)

where $a = (a_{-n}, a_{-n+1}, ..., a_{n-1}, a_n)$ and $b$ are multindexes. Write $\alpha = (a, b)$ as a combined multindex.

(e) Let $\partial_\nu u(\cdot; \alpha) = u_N(\cdot; \alpha)$ on $\Gamma_0$ be the data resulting from using the artificial domain $D_\alpha$ with boundary $\gamma(\alpha)$.

(f) Minimize $\|u_N(\cdot; \alpha) - u_N\|$ with respect to $\alpha$ to obtain an $n$-th order approximation $D_\alpha$ of the domain $D$. 

71
4.4 Increasing the rate of convergence

The rate of convergence for a variational algorithm as described above relies in large part on the accuracy of the initial guess. A key benefit of our algorithm given above is the ability to find a sequence of good initial guesses. Not only does our algorithm construct a series of increasingly good approximations, but it does so by varying only a few parameters, which increases the speed even more, by a multiple of as much as $10^3$.

In step one of the algorithm, a circle is reconstructed. This step has extremely high stability. Rather than solutions blowing up when the noise level increases, the reconstructed radii simply approach zero or one (depending on whether the normal derivative is measured inward or outward) as noise levels go higher than approximately $10^4$. See figure 4.4.

Then, after creating the best possible circular representation $D_r$ which matches available data of the unknown domain $D$, we add two more degrees of freedom to the variational algorithm by adding two more Fourier coefficients to the radial representation of $\partial D$. We illustrate the process with Figure 4.4 and 4.6.

4.5 Plots of convergence and computational efficiency

When $k$ is large, the solution to the full transmission can be approximated by the limiting solution $u_0 = u_\infty = \lim_{k \to \infty} u$, which solves the following boundary value problem:

$$u_\infty = 1 \text{ in } \overline{D},$$

(4.33)
Figure 4.4: Reconstructed circle has a radius of one with ten thousand percent noise added to measurements and data from outward-facing normal derivative.
Figure 4.5: First four regularization steps using Fourier method
Figure 4.6: Reconstructed 5-parameter domain using Fourier regularization method
\[ \Delta u_\infty = 0 \text{ in } \Omega^+, \]  

(4.34)

with the same boundary conditions on \( \Gamma_N \) and \( \Gamma_0 \). Let

\[ g_{i\infty} = \partial_{\nu} u_\infty \text{ on } \Gamma_0. \]  

(4.35)

Let us numerically solve both boundary value problem for the level set function,

\[ f_{D}(x) = (x_1/0.3)^2 + (x_2/0.3)^2 - 1, \]  

(4.36)

and analyze the convergence of \( g_i(\cdot; D) \) to \( g_{i\infty} \) as \( k \to \infty \). We denote by \( g_{ik} \) the data \( g_i \) corresponding to a specified constant \( k \).

Next let us analyze the convergence of the discretized approximation \( g_i^N \) with \( 8N \) source points (distributed with \( N \) points on each side of the unit square, \( 2N \) points on \( \Gamma \) and \( 2N \) points on \( \Gamma_1 \)) to the theoretical solution \( g_i \).

Finally, let us analyze the computer time used as \( N \) grows. We continued on up to \( N = 150 \), and it appeared that the solution time is \( O(N^2) \).
Figure 4.7: Semilogy plot of $\|g_{ik} - g_{i\infty}\|$ showing convergence to 0 as $k \to \infty$. 
Figure 4.8: Semilogy plot showing convergence of $g_i^{(N+1)^2}$. The top line shows $\|g_i^{(N+1)^2} - g_i^\infty\|$. The middle curve shows $\|g_i^{(N+1)^2} - g_i^{1,000,000}\|$. The bottom curve shows $\|g_i^{(N+1)^2} - g_i^{N^2}\|$.
Figure 4.9: Plot of seconds used for computing the solution versus the value of $N$
There may be slight differences between our numerical simulations and the most general types of boundary value problems introduced in previous chapters. Thus, we will describe in more exact details the direct problem we are solving for $u(; D, k)$ and the inverse problem for $D$ given extra inverse data $g_i(; D, k)$. Our numerical simulations almost always will take place in the reflected domain $\Omega$ rather than the physical domain $\Omega_p$.

When a semiconductor device is tested, it is normal to assume that two lateral faces are insulated. This is indicated by homogeneous Neumann data on a part of the boundary labeled $\Gamma_N$. We can take a square domain

$$\Omega = (-1, 1)^2, \quad (5.1)$$

and then let

$$\Gamma_N = \{-1, 1\} \times (-1, 1) \subset \partial \Omega. \quad (5.2)$$

Also, it is assumed that voltage is zero on the upper and lower edges of the device,
but that these edges are not insulated. We indicate this by homogeneous Dirichlet data on a part of the boundary called $\Gamma_0$, given by

$$\Gamma_0 = (-1,1) \times \{-1,1\} \subset \partial \Omega. \quad (5.3)$$

The voltage input is provided from an electrical contact in the center of the device. Without loss of generality, we can assume the voltage here is 1 volt, and so we call this region $\Gamma_1$,

$$\Gamma_1 = (-1/8, 1/8) \times \{0\} \subset \partial \Omega. \quad (5.4)$$

Together $\Gamma_0 \cup \Gamma_1 \cup \Gamma_N = \partial \Omega$.

All of our numerical simulations will consider only the unipolar case, and only the case where the conductivity coefficient is piecewise constant, equal to either the constant $a^+ > 0$ or the constant $a^- > 0$,

$$a(x) = \begin{cases} a^+, & x \in \Omega^+ := \Omega \setminus D, \\ a^-, & x \in \Omega^- := D. \end{cases} \quad (5.5)$$

We assume that $D$ is the doped region, with higher conductivity so that $a^- > a^+$. In an undoped semiconductor, then either $D = \phi$, or $a^- = a^+$.

Since we assume that both constants are positive, we will use the letter $k$ to indicate the ratio in conductivity $k = \frac{a^-}{a^+}$. In a doped semiconductor device then $k > 1$.

The voltage potential function $u$ satisfies the conductivity equation (1.2). We will indicate by $u = u(; D, k)$ the solution $u$ to the conductivity equation in the square
semiconductor device domain $\Omega$, with homogeneous Neumann data on $\Gamma_N$, homogeneous Dirichlet data on $\Gamma_0$, and Dirichlet data of one on $\Gamma_1$. The problem for $u(; D, k)$ is equivalent to the following transmission problem,

$$u(x; D, k) = \begin{cases} u^+(x), & x \in \Omega^+, \\ u^-(x), & x \in \Omega^-, \end{cases}$$  \hspace{1cm} (5.6)$$

with boundary conditions given by

$$\triangle u^\pm = 0 \text{ in } \Omega^\pm,$$  \hspace{1cm} (5.7)$$

$$\partial_\nu u^+ = 0 \text{ on } \Gamma_N,$$  \hspace{1cm} (5.8)$$

$$\partial_\nu u^+ = k \partial_\nu u^- \text{ on } \partial D,$$  \hspace{1cm} (5.9)$$

$$u^+ = u^- \text{ on } \partial D,$$  \hspace{1cm} (5.10)$$

$$u^+ = 0 \text{ on } \Gamma_0,$$  \hspace{1cm} (5.11)$$

$$u^- = 1 \text{ on } \Gamma_1.$$  \hspace{1cm} (5.12)$$

The trace of $u(; D, k)$ on $\partial \Omega$ is defined by the limiting values of $u^+$ near $\partial \Omega$, and $u(; D, k)$ on $\Gamma$ is defined by the limiting values of $u^+$ in $\Omega^+$ near $\partial D$ or equivalently by the limiting values of $u^-$ in $\Omega^-$ near $\partial D$. 

82
5.1 Discretized direct and inverse problems

The direct problem is to determine $u(D, k)$ when $k$ and $D$ are known. The inverse problem is to determine $D$ when the additional data $g_i(D, k)$ are given

$$g_i(D, k) = \partial_{\nu} u(D, k) \text{ on } \Gamma_0.$$

(5.13)

In practice, $g_i(D, k)$ may not be known, but only the pair $(x, g_i(x; D, k))$ for $x$ in a finite set (much smaller than the whole of $\Gamma_0$). We will definitely explore this case when dealing with a wing-shaped domain in Section 5.8 since it corresponds with what is available in a real-world quality control environment.

All of our simulations for the inverse problem will assume that $k$ is known. This does not present an insurmountable obstacle to developing quality control procedures. For instance, destructive testing, like spreading resistance profiling, could be used to determine $k$ in one semiconductor device. Then with the obtained value of $k$, non-destructive testing could be used for finding $D$ in other semiconductor devices, using the methods of our numerical simulations.

5.2 Undoped semiconductor voltage profile $k = 1$

By plotting the undoped semiconductor’s voltage profile and comparing it with the doped semiconductor’s voltage profile, at various levels of $k$ and for different doped regions $D$, we can visualize the function $u(D, k)$ and develop intuition for the direct problem that will help us solve the inverse problem.
Figure 5.1: Voltage profile of undoped semiconductor device $u(D(0.3, 0.3, 0, 0), 1)$
5.3 \( k = \frac{1}{10} \) reverse doping versus \( k = 10 \) doping

Examining the effects of reverse doping (i.e., reducing the conductivity of the doped region) on conductivity can be seen by comparing the voltage profiles of reverse-doped semiconductor devices to regularly doped device. Let’s use the same domain \( D \) and compare the voltage for 10:1 reverse doping and 1:10 reverse doping.

![Figure 5.2: \( u(D(0.3, 0.3, 0, 0), 1/10) \) voltage profile versus \( u(D(0.3, 0.3, 0, 0), 10) \)](image)

The regularly doped semiconductor device has higher conductivity in a circular region in its center, corresponding to the addition of elements with greater conductivity. Since electrical current flows more easily in this region, the gradient of the voltage profile is smaller. This results in the flatter appearance in the center of the graph of the heavily doped device. Both devices are subjected to zero voltage at the right and left, a voltage of one at an electrical contact in the center of the device, and insulated boundary conditions.
(no current flow) at the back and front.

5.4 Solving the inverse problem

So far, by determining the solutions $u(D, k)$, we have been solving the direct problem. As we have mentioned, the data for the inverse problem comes from measuring the normal derivative of the voltage potential function at part of the boundary of the semiconductor device, $\Gamma_0$. These data are called $g_i(D, k)$. There are always errors present in the measurements of $g_i(D, k)$. We will show the side-by-side graphs of $g_i((0.3, 0.3, 0, 0), 10)$ superimposed with the noisy data measurements, at one percent, two percent, and five percent noise levels.

![Figure 5.3: Inverse data $g_i$ superimposed with noisy data measurements](image)

The challenge of solving the inverse problem is in finding a solution $u(D, k)$ to the system of partial differential equations and boundary conditions, as well the domain $D$ from only the inverse data $g_i$. Even worse, the data $g_i$ may be sampled at a very small number of points. Possibly $k$ may also be unknown, but in this case at least two boundary
measurements are needed [38] to find both $k$ and $D$, according to the best mathematical theory currently available.

In our experiments we made very careful efforts to properly determine the necessary number of discretization points to accurately compute the inverse data $g_i$. We ended up using a 500-by-500 matrix (with 250,000 entries) to represent undetermined coefficients of the interior and exterior components of the voltage potential function. We made this choice based on $N = 50$ points for sampling $\Gamma_1$, and a proportional number of other points delegated to other interpolation curves used in the solution of the direct problem.

Here in figure 5.4 we show the inverse data $g_i$ computed by solving the direct problem for values of $N$ from 10 to 50. A circular domain is used for $D$. The plots of $g_i(;40)$ and $g_i(;50)$ are nearly indistinguishable.

Another important issue was positioning of source points. Ignorant rectilinear choices of source points, i.e., choosing them all a fixed distance from $\partial \Omega$, or from $\partial D$ produces disastrous results and prevented us from reconstructing any shapes more complicated than an ellipse. By using projections of the unit circle onto each axis to determine separation of sources, and by determining the distance from boundaries as a function of curvature, we were made tremendous improvements in the accuracy of the direct problems solution. Every bit of this accuracy is reflected also in the inverse problem solution, since perhaps as many as 5,000 direct problems must be solved for each iteration of the variational algorithm for finding $D$ from $g_i$.

In figure 5.5, one can see that on less complicated portions of the boundaries, source points are somewhat more widely separated (coming from projections of the unit circle), and at a further distance from the boundary (due to less curvature). The green circles
Figure 5.4: $g_i$ convergence as $N$ ranges from 10 to 50.
Figure 5.5: Optimal positioning of sources for curves of integration
represent interpolation points; the black stars represent source points for representing
the interior solution \( u^- \) (including some black points which are placed in the interior of
a small ellipse used to interpolate data placed on \( \Gamma_1 \)); the red x-marks indicate source
points for representing the exterior solution \( u^+ \). (Although it appears that the red source
points touch the corners of the domain \( \Omega \), in fact a small positive distance is always
maintained.)

For some complicated domains an appropriate positioning of source points is impos-
sible. In this case, the computer program produces a warning message that some part of
the interpolated boundary is not contained in the necessary region of harmonicity (based
on the interior bounded by inner or outer source points).

One of the most complicated parts of the computer implementation was writing a
subroutine for determining whether or not some set of points or some domain is contained
within a certain \( D \), given the Fourier coefficients for the radial representation of \( \partial D \). In
some sense, the solution to this problem was found by converting \( \partial D \) back and forth
between a level set function \( f_D \), where the interior and exterior of \( D \) are given by \( f_D < 0 \)
and \( f_D > 0 \), respectively.

Positioning of source points is probably the single most critical factor in successful
numerical analysis using integral equations. As we can prove by simple manipulation of
our source code, a single misplaced source point will destroy the entire computation or
invalidate all of the results.

We can instantly see one pay-off from improved positioning of sources. Before we
invented the automated source positioning algorithm, we noticed that the plots of inverse
data \( g_i \) looked something like figure 5.6. As you can see, there is a jump up at the left and
right sides of the graph. This jump should not be there based on practical considerations of physical laws, which say that energy should be lost the further electrical current must flow from its originating source. In fact, the data near these edges of the graph is totally worthless because of the clashing singularities of poorly placed source points which could never be unified by any choice of coefficients. Another difference in the graph is that in this case $g_i$ was computed from inward-facing normal derivatives (thus, when reaching a minimum at the boundary, it is not negative as one would usually expect from standard outward-facing normal derivatives).

Figure 5.6: Inverse data $g_i$ computed without appropriate source positioning algorithm

The locations of points used for plotting $g_i$ are shown with stars in figure 5.6.
5.5 Disk

The most hopeful case for reconstructing a doping profile is when one assumes that the doping compounds have diffused in an essentially circular region $D_r$, $r \in [0, 1]$. Then the inverse doping profile problem reduces to the inverse conductivity problem which reduces to a problem of determining one real variable $r$ on the half-open interval $(0, 1]$. [We normally exclude the degenerate case of a non-doped device which would have $r = 0$, i.e., $D$ is empty. In the scaled device geometry, $r \leq 1$ holds universally under the given hypothesis that $D$ is circular, so there is no limitation here.]

![Figure 5.7: Reconstructions of the circular domain $D_{0.5}$ with 0, 5, 10, 50, 1,000, and 10,000 percent noise](image)

Figure 5.7: Reconstructions of the circular domain $D_{0.5}$ with 0, 5, 10, 50, 1,000, and 10,000 percent noise

92
We observe quite good stability for finding the parameter $r$ even with extremely large amounts of noise. One interesting phenomenon is that when noise exceeds a relative level of 1000%, the reconstructed $D$ grows smaller and smaller as the noise level in measurements continues to increase. We feel this is due to the fact that our method for adding random noise had equal variation above and below zero. Thus, as the true data became negligible compared to the random noise, the difference between the $g_i$ for the reconstructed domain $D_r$ and the original noisy data $g^*_i$, was minimized near $r = 0$.

We use the following Matlab function to reconstruct $D_{0.5}$ from various noise-free and noisy measurements.

```matlab
function circleNoiseLevel
D = .5;
g_exact = g1(D, 50);
noise = [0 1 5 10 50 100 200 500 1000 5000 10000];
for n=1:length(noise)
    ratio = noise(n)/100;
    g_noisy = g_exact .* (1 + (rand(size(g_exact))-.5)*ratio);
    delta = @(Dg) norm(g1(Dg, 50)-g_noisy);
    Dinv = fminsearch(delta, .8);
    hold off
    simplePlotD(D);
    hold on
    simplePlotD(Dinv, 'o');
    axis([-1 1 -1 1]);
    title(sprintf('Original circle versus reconstruction with %d percent noise', 100*ratio));
    print(gcf, '-dpng', sprintf('circleNoiseLevel/Original-vs-%d.png', 100*ratio));
end
```

In figure 5.7 we see that at zero, one, five, and ten percent noise, the reconstructed circle is visually indistinguishable from the original. At larger levels, the reconstructed
shape grows larger than the original circle.

5.6 Ellipse

We find almost the same results with an ellipse or pseudo-elliptical approximation of an ellipse. However, the largest possible noise level to result in a very successful reconstruction is ten percent, shown in figure 5.8.

![Figure 5.8: Reconstructions of the psuedo-elliptical domain $D$ with 5 and 10 percent noise](image)

Figure 5.8: Reconstructions of the pseudo-elliptical domain $D$ with 5 and 10 percent noise
5.7 Peanut, including some numerical instability

For simplicity we will be considering a four-parameter family of peanut-shaped domains $D(d_1, d_2, d_3, d_4)$, where

$$\partial D = \{(x, y) : x = d_1 \cos(t) - d_3 \cos^3(t), y = d_2 \sin(t) - d_4 \sin^3(t), \ t \in [0, 2\pi]\}, \ (5.14)$$

and $D$ is defined as the bounded component(s) of $\mathbb{R}^2 \setminus \partial D$.

We will also consider peanut-shaped domains as a subset of more general domains described by their Fourier expansion around an initial peanut-shaped prototype domain. For these domains, reconstruction was extremely successful with 0, 1, 5, and 10 percent noise. Reconstruction with more than 200 percent noise failed catastrophically; in fact, the reconstructed domains often did not contain the edge $\Gamma_1$ of the semiconductor device, which is a basic necessity for any physically permissible region $D$.

Figure 5.9 shows an attempt to satisfy the boundary conditions supposedly satisfied by a self-contradictory doping profile.

Here is another slightly less extremely, but still horribly inaccurate, voltage profile

Note that the numerical stability occurred even though there were still plenty of source points.

By contrast, effective source point positioning also drastically reduces the number of source points. Figure 5.11 is essentially the same solution computed in figure 5.10. The previous difficulty with numerical instability arose because of the small size of the doped region $D$. Even though only two sources are used for each boundary, and only four around $\partial D$, the solution is still much more accurate.
Figure 5.9: “Dinosaur” created by physically impossible domain configuration
Figure 5.10: Voltage profile showing classical appearances of numerical instability due to poor source placement
Figure 5.11: Numerical stability even with only 2 source points on each boundary
Figure 5.12 shows a peanut shaped domain (blue line) reconstructed (blue circles) successfully from measurements with ±5 percent added random noise. The results shown is typical for any similar domain configuration.

Figure 5.12: Good reconstruction of a peanut-shaped domain from data with 5 percent relative noise

As we remarked earlier in this section, reconstruction of peanut-shaped domains generally failed with noise levels above 200 percent.
5.8 Wing domain: Positioning of detectors

By choosing points of electrical measurement for a wing-shaped domain we will be able to analyze the stability of the inverse problem solver. We will look at its reconstruction of a wing-shaped domain using only a few contact points for measuring electric flux—i.e., inverse data $g_i$ are sampled at only a few points.

The true domain $D$ and initial guess $D_1$ are each given by four parameters.

\[ D = (.5, .2, .1, -.03); \quad (5.15) \]

\[ D_1 = (0.8, 0.8, 0, 0). \quad (5.16) \]

Here the boundary of $D$ is given by the parameterization

\[ \begin{align*}
  * \ x(t) &= p1*\cos(t) + p3*\cos(2*t) \\
  * \ y(t) &= p2*\sin(t) + p4*\sin(2*t)
\end{align*} \]

The $p_i$’s are the four components of the parameter-vector.

The normal derivative (current flow) is measured at 5 contact points shown by blue stars at the bottom of the graph. One percent relative random noise is added to the measurements before attempting to reconstruct the domain $D$.

The normal derivatives measured at these contact points:

\[ g_{1u} = \]
Enter the amount of noise (i.e., relative error between 0 and 1): .01

glu_noisy =

0.4019
0.3982
0.3955
0.3389
0.3530

The variational minimization algorithm proceeds to find the best shape which fits the data by solving many direct problems (approximately 30 iterations are needed). To improve stability the best-fitting elliptical domain is determined first, and then used as an initial guess for the non-elliptical inverse solver.

We show the output of the algorithm below, and the reconstruction is graphed in figure 5.13, along with the five sampling points.

The best fitting ellipse is shown above with a dotted line, and the best fitting non-elliptical solution is indicated by blue diamonds. The initial guess is shown in green.
Figure 5.13: Wing test for calibration of five measurement positions
Note that the initial guess contains the true domain in its interior.

After many experiments, we determined that choosing an odd number of contact points works best so that they can be placed symmetrically about the center of the coordinate system. Also, reconstruction works best when the horizontal coordinates of the furthest contact points extend at least as far as the horizontal coordinates of the true domain D.

We will show some more of our experiments below. Each image reveals the contact points where measurements were taken, as well as the reconstruction which resulted.

One can see that only 3 points are sufficient for a good reconstruction when they are symmetric and distributed on a similar length scale as the size of the domain. However, 10 points do not result in a good reconstruction when they are restricted to a narrow band along the boundary of the semiconductor device.

5.9 Arbitrary domains

We use our automatic parameterization algorithm to locally identify arbitrary domains, which are close in some sense to an initial disk-, ellipse-, wing-, or peanut-shaped domain.

The parameterization of \( \partial D \) is given by polar coordinates \((r(\theta), \theta)\), where

\[
r(\theta) = a_0 \sum_{j=1}^{\infty} a_{2j-1} \cos(j\theta) + a_{2j} \sin(j\theta)
\]  

(5.17)

is the real Fourier expansion of \( r \). Identifying \( D \) is tantamount to identifying the sequence \((a_n) \subset L^2(\mathbb{R})\). Once again, \( D \) is defined as the bounded component(s) of \( \mathbb{R}^2 \setminus \partial D \).
Figure 5.14: Successful wing test for calibration of three symmetric measurement positions
Figure 5.15: Unsuccessful wing test for calibration of ten narrow-band measurement positions
It is necessary to form some regularization procedure for creating test domains, i.e., creating sequences \((a_n)\) which produce simply connected \(D\) containing the electrical contact curve \(\Gamma_1\). The same regularization procedure will have a two-fold purpose:

1. The regularization procedure enables us to create numerical experiments for a variety of realistic, but simulated, domains \(D\).

2. The regularization procedure enables us to improve the solution of the inverse problem for reconstructing an unknown \(D\), by limiting variations of reconstructed sequences \((a_n)\) to those which correspond to physically tenable doped regions \(D\).

The regularization procedure for creating test domains is contained in the Matlab file fourierD.m. All files of computer code will be made available to the dissertation committee members upon request.

Note that almost all of the previous experiments can be considered as subsets of the automatic parameterization algorithm, since the Fourier expansion procedure is general enough to represent all previous types of domains, since they were all star-shaped with respect to the origin (i.e., able to have their boundary represented as a single-valued function \(r(\theta)\) in polar coordinates).

Since this algorithm has a way to make a series of improved initial guesses, it represents our most reliable method for solving the inverse doping profile problem (by means of solving the inverse conductivity problem). The results are very good, and the best way to explain them is by showing two sets of numerical simulations for global reconstructions of randomly created domains depending on nine Fourier coefficients.
Figure 5.16: Steps of multi-stage algorithm for locating a nine-parameter domain with one percent noise in measurements
5.10 Local versus global uniqueness

From a numerical viewpoint, local uniqueness seems to be on solid ground, as hinted at by figure 5.17, which shows a near-perfect reconstruction of a 101-Fourier-parameter domain by using no regularization and Nedler-Mead simplex variational minimization. There was a 30 percent relative perturbation of radial Fourier coefficients between the original domain and the initial guess. The initial guess is shown with a similar shape, but clearly smaller size, and slightly differing conformation, in comparison to the original domain, which is also a blue line. The reconstruction is indicated by stars, which fall precisely on top of the original domain at all points.

Figure 5.17: Nearly perfect local reconstruction of 101-Fourier-parameter domain
The relative difference in data $g_i$ from the original domain and the initial guess is 18.16 percent; the relative difference in data $g_i$ from the original domain and the reconstructed domain is 0.148 percent. Likely with further refinement of the discretization grid (i.e., more grid points), the relative error would be even smaller.

However, global uniqueness definitely does not seem to apply for non-convex domains. Figure 5.18 shows relative matching error of less than one-tenth of one percent for a drastically different domain (labeled with stars) than the original one (solid line). Zero percent noise was added to the measurements. However, by choosing a slightly different initial guess (marked with diamonds) we are able to identify the original domain successfully (figure 5.19).

This result does not contradict Alessandrini’s theory, since we have never found close match such as this between two different domains which are subsets of one another. The parameters (70, 30, 10, 25) of the original domain are defined by equation (5.14), after dividing each of them by 100. (Decimal points were eliminated to make the graphs easier to read.)

5.11 Final remarks

Unipolarity

It is important to note that we only solve the unipolar case. What is the difference between the unipolar case and the bipolar case? In the unipolar case, one assumes that the source of current doesn’t matter, and so one assumes that all current is caused by electrons (i.e., there are no holes, so $v = 0$), or vice versa. Now in the bipolar case, the
Figure 5.18: Good matching in $g_i$ for a non-identical domain
Figure 5.19: Good matching in $g_i$ for an identical domain
additional data for the inverse problem consist of the total flux from a linear combination of flux from electron-current $\partial_\nu u$ and flux from hole-current $\partial_\nu v$. Uniqueness for both $u$ and $v$ can be proved [43] provided that $\partial_\nu u$ and $\partial_\nu v$ can be measured separately, but it is physically impossible to distinguish from electric current caused by electrons or electric current caused by holes. Some simulations have been done where a very small number of parameters are reconstructed for the bipolar case [43], sometimes by using laser-beam induced current models [15].

One of our goals is to rely on data that are as simple as possible to obtain, e.g., ones which don’t require laser beam induced current. The abstraction of assuming that all current is caused by either electrons, or all current is caused by holes, seems to be almost universally used for doping profile analysis. The assumption of unipolarity is also fundamental in related hydrodynamical models of semiconductors [10, 27, 42].

**Easier-to-obtain boundary data**

We simplified boundary data for the unipolar conductivity equation by forming the adjoint problem, where instead of functionals we are given extra boundary data. This technique is widely used in the inverse problem of option pricing (Dupire’s equation) as a simplifying device that enables one to obtain theoretical and numerical results.

The result is an inverse problem for the conductivity equation that is specified in Section 2.6. We solve this inverse problem using a special numerical algorithm. We call it a “local” reconstruction if we have have knowledge which permits us to make an initial guess quite close to the true solution. We call it a “global” reconstruction if we have absolutely no knowledge about the solution.
Regularization

In the global case, our initial guess is a circle $D_r$, and we find the radius $r$ which minimizes the difference between data $g_i(\cdot; D_r)$ and the given data for the inverse problem $g_i(\cdot; D)$. Typically we use the norm $f(D_r) = \|g_i(\cdot; D_r) - g_i(\cdot; D)\|_\infty(\Gamma_0)$.

The first stage when $r$ is varied and other Fourier coefficients are held constant serves as an effective regularization method, improving stability and the rate of convergence for subsequent minimizing sequences constructed by Newton’s Method or Nedler-Mead simplex direct search.

A similar regularization procedure is used throughout the entire global reconstruction algorithm, where two additional Fourier coefficients are allowed to vary in a neighborhood of zero, while previous lower-order Fourier coefficients have already been determined nearly exactly.

To our knowledge, no one has ever used this regularization scheme before, at least not in the inverse doping profile problem.

In the local case we use no regularization, and directly minimize the functional $f(D^*)$.

In the global case we use our own special regularization technique which works by inserting the solution $D_r$ into the next stage of reconstruction, and attempting to identify two more Fourier coefficients of $\partial D$ expressed in a special radial or star-shaped coordinate system, resulting in another domain $D_3$ which is closer to the desired solution $D$. Numerical instability is kept to a minimum since only the first free parameter, $r$, is already close to the optimal value, and primarily there are only two more parameters to be determined for $D_3$ compared to $D_r$. Then $D_3$ is used as input into the next stage of reconstruction, which attempts once more to find two additional Fourier coefficients,
resulting in five parameters being identified, and a solution $D_5$ which is even closer to the desired solution $D$.

We can also combine this with a Tikhonov regularization scheme which penalizes high frequency components of the solution, for instance the $n$-th Fourier coefficient is penalized by combining it with $n^2$ as a factor of the regularization parameter $\alpha$. We find good results in almost all cases, however, even when $\alpha = 0$.

Although all of our input data and inverse problem solutions come from simulations, we believe that it would be straightforward to apply our algorithm to the identification of areas of higher conductivity within actual semiconductor devices. In principle this is sufficient to identify the doping profile for PN-type semiconductor devices, which are the building blocks of more complex devices.

Figure 5.20: Nearly perfect global 9-parameter (left) and local 101-parameter (right) reconstructions
Table 5.1: Number of parameters able to be reconstructed

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Local</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>More than 50</td>
<td>Up to 5</td>
</tr>
<tr>
<td>1%</td>
<td>More than 50</td>
<td>Up to 4</td>
</tr>
<tr>
<td>5%</td>
<td>More than 50</td>
<td>Not more than 3</td>
</tr>
<tr>
<td>25%</td>
<td>Around 10</td>
<td>1 or 2</td>
</tr>
</tbody>
</table>

“Local” means that an initial guess is provided with a 5% relative perturbation from the true solution. Global means that no initial guess is provided (thus an initial guess with a circle of radius 1/2 is used). Newton’s Method works well only with about $10/df$ percent relative error in the initial guess.

**Benefits**

We think our method has three primary benefits over existing methods for identifying conductivity or doping profiles:

1. The necessary boundary measurements are very simple. For a square device with four sides in its vertical profile, our input measurement requires constant voltage (e.g., 1) at a small contact on one side $\Gamma_1$, and different constant voltage (e.g., 0) on the opposite side $\Gamma_0$. The last requirement is for the remainder of the device to be insulated.

2. We think our method would provide higher resolution for mapping the conductivity or doping profile of a device. This is based on results we have seen in the literature for identifying a circle (one parameter), off-center circle (two parameters), or a linear transformation of the Heaviside function (two parameters). The published results we have seen are quite inaccurate, even though they are based on simulated
data with no more than 10% noise. Our global method can easily reconstruct 5 or more parameters with no a priori information, except that the desired domain be simply connected. Our local method can reconstruct well over 50 parameters. The relative error in reconstructed Fourier coefficients appears to be no more than double the relative error in the data given for solving the inverse problem.

We have been careful to avoid inverse crimes by testing our reconstruction algorithm with input data produced by different discretizations as well as different placement of sources, compared to the discretizations and placement of sources used to produce the output data.

3. Finally, the type of measurements necessary are non-destructive, so applying our method would be more efficient than destructive quality-control methods.

In cases where it is necessary to identify either conductivity or the doping profile of a PN-type semiconductor device, we believe that our algorithm, precisely defined in Section 4.3, is superior to previous methods.
We have reduced the stationary drift-diffusion equations into the conductivity equation, under reasonable and widely-used assumptions of zero space charge and low injection. The result is a one-to-one relationship between solutions to the inverse problem for the conductivity equation and solutions of the original inverse problem for doping profiles. Then we have proved important global and local uniqueness results for the inverse conductivity problem, as well as a more general class of semilinear elliptic equations. Moreover, we have developed a very efficient and accurate algorithm for solving the direct conductivity problem given in Section 2.7. Finally, we have solved the resulting inverse conductivity problem by finding a minimizing sequence of a suitable error functional, both with and without any initial guess provided.

Although the necessary assumptions of zero space charge and low injection are both widely used by many semiconductor analysts, the future of semiconductor technology may result in devices so small that such assumptions cannot be made. If this occurs, our work needs to be continued and improved, perhaps using the hydrodynamical model of
Some authors, such as [43], have posed theoretical problems about using bipolar models, while also admitting [43] that electrical current cannot be separated into a component due to electrons versus a component due to holes. It may be possible that mathematical analysis of the bipolar models would still bring improvements to the inverse doping profile problem, and this remains an area for future work.

Perhaps even better results could be obtained with even less data by considering the transient case of the drift diffusion equations (2.3). Time-dependent effects, such as inductance and capacitance, would thus be retained. Even though it may prove very difficult, we think this future analysis would be worthwhile.
Bibliography


