Numerical solution of drift-diffusion and electrostatic potential equations for solar cells

Master thesis in Applied
and Computational Mathematics

Daniel Engelsen

Department of Mathematics

University of Bergen

June 1, 2018

Acknowledgements

First, I would like to thank my supervisor Tor Sørevik who has been an outstanding support and provided excellent help with my work on this thesis. I also need to thank fellow students Einar, Benjamin and Karoline that have also been a part of this solar energy project that Tor Sørevik has initiated. I would also like to pay thanks to the research partners and friends in Ghana, Hisham and Dr. Peter Amoako-Yirenkyi that we have been collaborating with on this solar energy project. Further, I would like to thank BKK for funding the excursions we had to Ghana to work on our research with our collaboration partners in Kumasi. Lastly, I would want to thank family and friends for all support.

Daniel Engelsen
May 2018

You could power the entire United States with about 150 to 200 square kilometers of solar panels. The entire United States! Take a corner of Utah... There's not much going on there, I've been there. There's not even radio stations.

- Elon Musk

Abstract

This thesis gives an introduction to how solar panels work and some of the problems for solar cells that need to be solved. Further, it gives an introduction to some physical equations that describe the electrical behavior inside the panels. We try to solve these equations numerically for different parameter values and with different numerical techniques. An analysis of some of the convergence and stability of each numerical system is given, as well as numerical results with numerical values and plots.

Figure 1: Solar panels at the Department of Engineering, KNUST, Kumasi, Ghana.

Note to reader

In this thesis, lists with explanations of selected physical terms, nomenclature and numerical values of physical constants are noted in Appendix C, D and E, respectively.

All facts that are stated in this thesis that may change with time (for example efficiency of a particular solar cell or its prevalence) are facts at the time of submission: June 1, 2018.

The term "light" refers to photon radiation, regardless of where in the electromagnetic spectrum the photon radiation is.

There are different notations for the groups in the periodic table of elements. Sometimes, group $11-18$ is referred to as group $1-8$, $1B-8B$ or $IB-VIIIB$. I will refer to these groups as group $11-18$.

Computational and numerical considerations are discussed in MATLAB syntax in gray boxes.

Calendar dates are written on the form DD.MM.YYYY.

The bibliography is arranged alphabetically and is made according to the APA 6th style.

Appendix F - The excursions to Ghana, is not to be taken into consideration in the evaluation of this thesis.

Mathematical notation:

I will let log() denote the natural logarithm unless a sub number is noted, then that number will be the base for the logarithm.

I let y' denote spatial differentiation (in one dimension, and let x denote the spatial variable), and $\dot{y}$ denote time differentiation (with t as the temporal variable).

I let $\mathbf{0}_N$ be a zero column vector of length N and 0 be a scalar zero or a zero-matrix of appropriate size. I will let $\mathbf{1}_N$ denote a column vector of length N with all elements equal to 1. The matrix $I_N \in \mathbb{R}^{N\times N}$, will denote the identity matrix.

All vectors are written in bold. We define for a vector $\boldsymbol{v} = [v_1 \dots, v_N] \in \mathbb{R}^N$ that $\hat{\boldsymbol{v}} = [v_0, \dots, v_N] \in \mathbb{R}^{N+1}$ and $\tilde{\boldsymbol{v}} = [v_0, \dots, v_{N+1}] \in \mathbb{R}^{N+2}$, unless otherwise specified.

The element at index (i, j) of matrix A, will be denoted $A_{i,j}$. The notation $\{B\}_{i,j}$ for a block matrix B, will denote the block at index (i, j) of B. Indices of vectors or matrices that are composed of two or more expressions will be denoted using square brackets, for example, the element at (i, j) of the product of matrix A and B will be denoted $[AB]_{i,j}$.

The term "Jacobian" will always refer to the Jacobi matrix. The Jacobian of $\boldsymbol{f}(\boldsymbol{y})$ with respect to $\boldsymbol{x}$ will be denoted $\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{f}(\boldsymbol{y})}$, and evaluated at $\boldsymbol{a}$, will be denoted $\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{f}(\boldsymbol{y})}(\boldsymbol{a})$.

Further notation can be found in Section B.1 and B.2.

Contents

Chapter 1

Introduction

1.1 Introduction to solar panels

A solar panel is an electronic device that converts sunlight into electricity

A *solar panel* is made up of smaller units called *solar cells*. Solar panels and cells are also called *photovoltaic* panels and cells. The term "photovoltaic" is a more precise term, as solar panels sometimes are referred to as devices that use sunlight to heat water. Also, solar panels may use light from other light sources than the sun to operate. I will in this thesis use the terms "solar panel" and "solar cell" to describe devices that convert sunlight into electrical energy.

Solar cells are specialized *semiconductor diodes* that convert light into electrical current. A solar cell can operate in different ways. The most common way of constructing a solar cell is to make it with two layers of *doped silicon* with each layer injected with different *dopants*. The main reasons why silicon is the most used is because silicon-based solar cell technology is the most mature solar cell technology [19]. Silicon devices also exhibit better properties than other semiconductors at room temperature and it is the most abundant semiconductor element on Earth [18].

Another way of constructing a solar cell is to make it with *organic compound*

consisting of layers of materials that have different *electron affinities* and *ionization potentials* [14]. Most organic solar cells are *polymer* cells. These types of panels are currently less efficient than silicon solar panels, but they are much cheaper to produce. The most efficient solar panel, which is a silicon solar cell, has an efficiency of 46% [8]. For comparison is the most efficient organic solar 13.2% efficient [10].

In this thesis, we will consider these two types of cell structures in particular. and how to model different flow, potential and charge characteristics. We start by giving a theoretical introduction to understand the physics of solar panels, and later introduce the governing equations for semiconductor devices that we will utilize. These are then solved numerically for different cases.

1.1.1 Electronic band structure

The electronic band structure describes the range of energies that an electron within a solid may or may not have

This subsection is based on [18]*, Section 1.4.*

One of the most fundamental physical concepts for understanding how solar cells work is the *electronic band structure*. This structure describes the different energy levels in which the electrons in an atom can exist. Due to quantum mechanical effects, there are different energy levels the electrons can lie in. These quantum mechanical effects change depending on how close the atoms are. In a silicon solar cell, however, the silicon atoms are arranged in a lattice. Due to the distance between each atom in the lattice, the energy levels get distributed in such a way that they lie in bands. In other words, there are sections of closely packed energy levels of possible energies in which the electrons around the nucleus can exist. These bands are called *energy bands*. Each physical material has different energy bands and at which energies these bands lie determine a lot of their physical properties.

In an atom, there are different shells that the electrons can be in. The further away the shell is from the atomic center, the higher energy for the electrons that lie in that shell. Each shell of the atom has different energy bands in which the electrons can exist. These shells for the silicon atom can be seen in Figure 1.1.

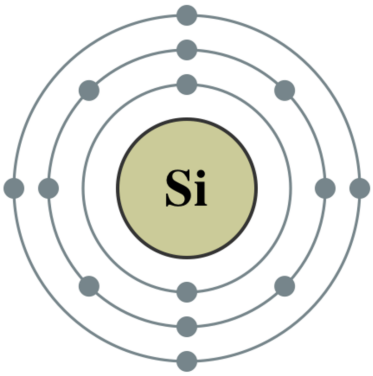


Figure 1.1: The atomic model for the silicon atom. The silicon atom has 4 electrons in its outermost shell, also called valence shell. An atom can have at most 8 electrons in its valence shell when the valence shell is the third shell from the center. Courtesy of Popular Science.

Electrons are negatively charged and lack of electrons, called *holes*, are positively charged. Although the hole is not a particle, the hole acts as a particle in the sense that it takes the same amount of energy to move it a certain distance as it takes moving an electron the same distance. Electrons and holes are *charge carriers*. It is the charge carriers that are responsible for the electricity generation in a silicon solar cell. Only the electrons/holes that lie in the *valence shell*, the outermost shell, can be charge carriers.

There are two different energy bands, in particular. in conducting materials that are of interest. These are the two energy bands nearest to the *Fermi level*, namely the *valence band* and the *conduction band*. The difference in energy between the highest energy level of the valence band, known as the *maximum valence band energy*, E_V, and the lowest energy level of the conduction band, known as the *minimum conduction band energy*, E_C, is defined as the *band gap*, $E_g := E_C - E_V$. The interval in which the electrons cannot exist is sometimes referred to as the *forbidden gap*. An illustration of these energy bands and corresponding bandgaps can be seen in Figure 1.2.

A material with an overlapping valence and conduction band, in other words, a material with a band gap of zero, is called a *metal*. Metal and other materials with a band gap of zero or close to zero are *conductors*. *Insulators* have a band

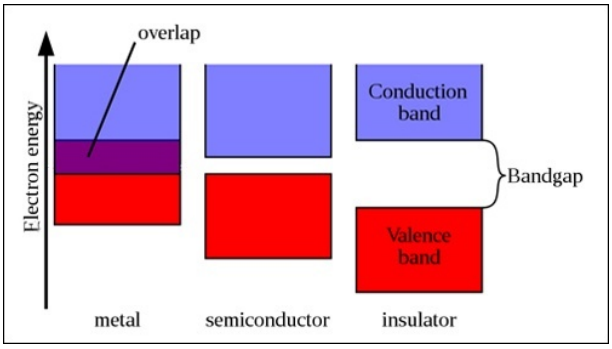


Figure 1.2: An illustration of the band gaps for metals, semiconductors, and insulators. Courtesy of University of Calgary.

gap that is large, and *semiconductors* have a band gap that lies between that of a conductor and an insulator, and is normally on the order of 1 eV.

For organic semiconductors, we do not have valence and conduction bands. Electrons and holes in the conduction band in inorganic semiconductors can be considered free and can thus move freely. In organic semiconductors, however, the electrons and holes lie in energy wells, named *traps*, and can jump from trap to trap. For organic semiconductors, the equivalence to the maximum valence band energy is what we call *highest occupied molecular orbital* (HOMO), and the

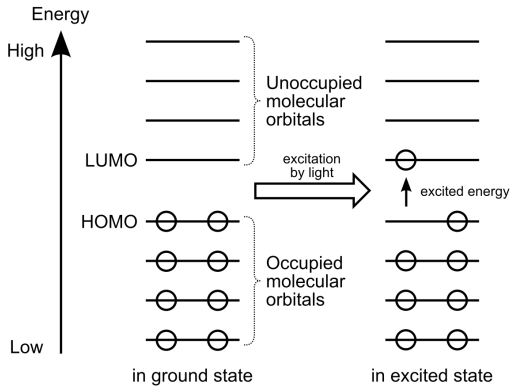


Figure 1.3: An illustration of the HOMO and LUMO orbitals for an atom. Courtesy of Wikipedia.

equivalence to the minimum conduction band energy is what we call *lowest un-*

occupied molecular orbital (LUMO). An illustration of the HOMO and LUMO levels can be seen in Figure 1.3.

1.1.2 Semiconductors

A semiconductor is a material that has an electrical conductivity between that of a conductor and an insulator

A semiconductor is any material whose electrical conductivity is in the range from 10^{-8} to 10^{4} S/cm [18] (S is the SI unit *siemens*, a measure of electric conductance). Although there are some exceptions to this definition, as stated in [17], these are exceptions that will be given no emphasis in this thesis. Also, there are different definitions for the conductivity range of semiconductors; in Figure 1.4, semiconductors are defined as materials in the range from 10^{-8} to 10^{3} S/cm. In any case, where specifically this range lies is not of importance for this thesis.

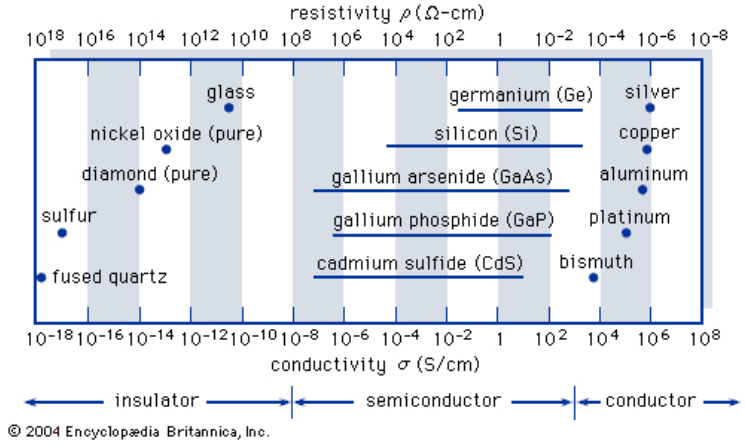


Figure 1.4: Electric conductivity of different materials. Courtesy of Encyclopaedia Britannica.

The electrical conductivity of semiconductors changes with temperature. Semiconductors have the property that the electrical conductivity increases as the temperature increases, opposite to the property of metals.

1.1.3 Radiation from the sun

The sun is the power source for solar cells

A solar cell can turn light into electricity by converting light from the sun into electrical energy. Not all the photon energy that hits the solar panels is converted into electrical energy; some energy is lost due to different physical constraints in the panel construction. Also, some of the sunlight never reaches the panels on Earth due to reflection and diffusion from different gases and particles in the atmosphere, like for example air, clouds, smog, dust and sand [7].

The amount of photon energy that makes it through the atmosphere also depends on how long the distance of atmosphere the light has to go through. If the sunlight rays are perpendicular to Earth's surface, they have the shortest distance of atmosphere to go through, and conversely, when the sun is at the horizon, the light rays have the longest distance of atmosphere to go through. Also, sunlight at different wavelengths has different energies.

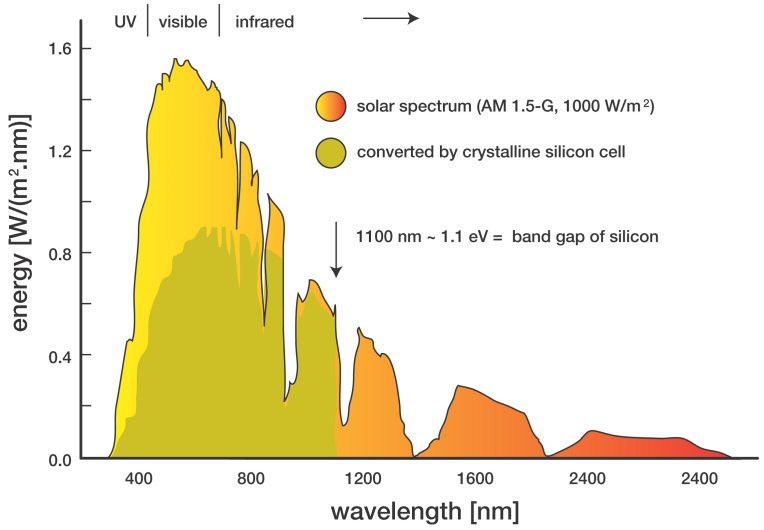


Figure 1.5: The solar spectrum at sea level. The different "valleys" are due to light absorption from different gases in the atmosphere. The green color represents the energy range where a crystalline silicon solar cell can absorb energy. Courtesy of University of Alaska Fairbanks.

The yellow-red-shade part of Figure 1.5 shows, under standard test conditions, the power per square meter of sunlight at sea level, adjusted for the different energies of light in different parts of the spectrum. The green part shows the energy that can be converted by a silicon panel with a band gap of 1.1 eV. As we see, no energy is converted above a certain threshold around 1100 nm in photon wavelength, which is one of the weaknesses of solar panels. As we will see in Subsection 1.1.4, there is an upper limit for the efficiency of a semiconductor solar panel.

1.1.4 Theoretical maximum efficiency

A single-junction solar cell at sea level cannot have an efficiency that exceeds 33.7%

To discuss efficiency, we first need to look at different setups. For this subsection, we will only consider cells that only take in direct sunlight, i.e. that the light has not been concentrated before hitting the cell.

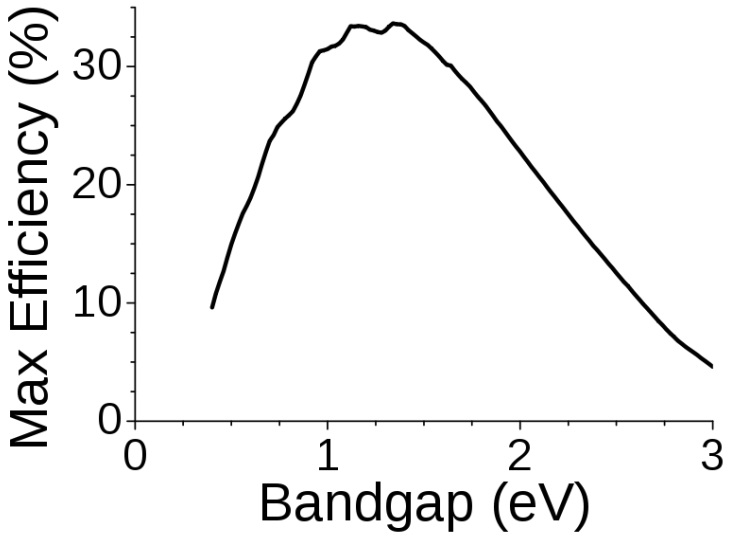


Figure 1.6: Maximum theoretical efficiency vs. band gap for a single junction solar cell. Courtesy of Wikipedia.

For a single-junction semiconductor cell, there is an upper limit for efficiency

known as the Shockley–Queisser limit which states that a single-junction cell may have an efficiency of maximum 33.7%. This theoretical maximum can be attained only if the upper layer of the solar cell is built with a material with a band gap of 1.34 eV [15]. A semiconductor cell with several junctions may have a higher efficiency than a cell with only one junction; with an infinite number of junctions, the theoretical maximum is 86.8% [6]. These theoretical maxima are for cells at sea level under standard test conditions.

The most popular solar cell material, silicon, has a band gap of 1.12 eV, leading to a maximal efficiency of around 32% [18]. These results are experimental and are plotted in Figure 1.6 with real sunlight at sea level. The data is not smooth because of absorption of different spectra due to different gases in the atmosphere. This absorption can also be seen in Figure 1.5.

1.2 How silicon solar cells work

Silicon solar cells work by having two layers of silicon with different doping, with current carriers that flow from one layer to another and do work in an external circuit

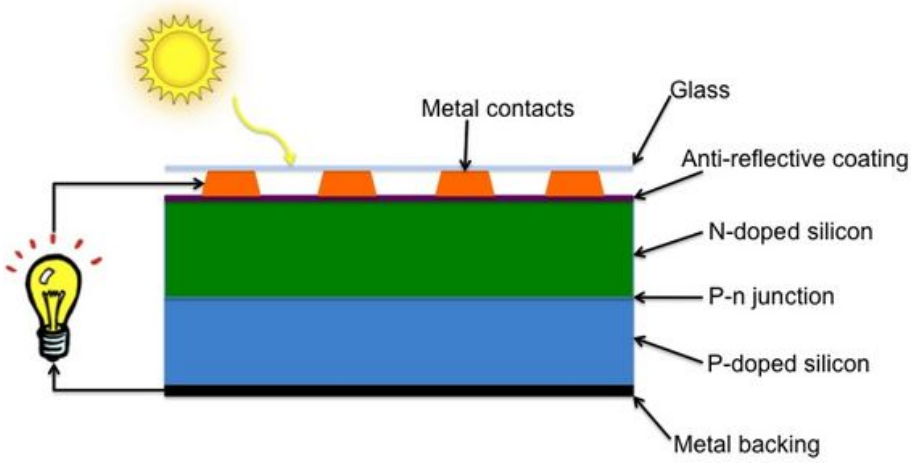


Figure 1.7: An illustration of a typical silicon solar cell. Courtesy of ElProCus.

1.2.1 Doping

Doping is a method of adding an impurity to a pure semiconductor to change its electrical properties

This subsection is based on [18]*, Section 1.6.*

For a semiconductor, doping is a method where a dopant, another chemical material with different chemical properties, is added to a pure semiconductor to change its electrical properties. An undoped material is called *intrinsic* and a doped material is called *extrinsic*.

When a dopant that leads to an increase of electrons is introduced, the material is *n-doped* and the material is called the *n-type*. Conversely, when a dopant that leads to an increase of holes is introduced, the material is *p-doped* and the material is called the *p-type*.

Silicon cells consist of several layers. There are often two layers of silicon. In such a case, it is these two layers that are doped. The layers are doped in order to increase the density of electrons or holes. Silicon is an element in group 14 in the periodic table of elements. A dopant for silicon is typically an element from group

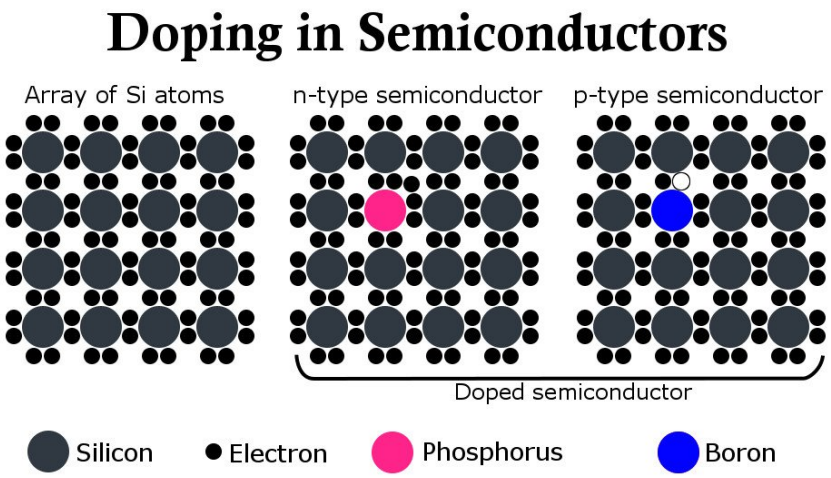


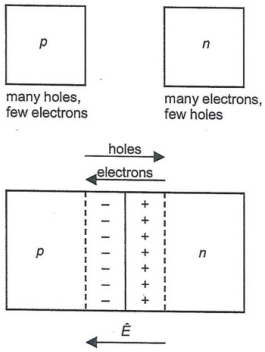
Figure 1.8: An illustration of an intrinsic silicon lattice (left), an n-doped silicon lattice doped with phosphorous (center) and a p-doped silicon lattice doped with boron (right). Courtesy of Science ABC.

13 or group 15. Elements in group 14 have 4 electrons in the valence shell, and elements in group 13 and 15 have 3 and 5 electrons in the valence shell, respectively. If an element from group 13 is introduced, then that material will work as an acceptor, and conversely, if an element from group 15 is introduced, then that material will work a donor. For most silicon solar cells, the two dopants are boron (B) and phosphorus (P). The dopants are very scarce in the silicon material, with a density of around one atom per 10^6 silicon atoms [7].

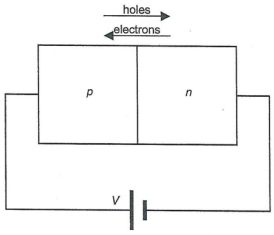
1.2.2 The p−n junction

The p−n junction is the interface between the n-type and the p-type silicon layer

This subsection is based on [19]*, Section 2.5.*



(a) Illustration of a p−n junction. The explanation of the p-type and n-type layers are for charge carriers in the valence shell.



(b) Illustration of a p−n junction with an applied voltage.

Figure 1.9: Courtesy of [19].

A p−n junction is formed by joining one layer of an n-type and one layer of a p-type semiconductor. When the two layers are joined, some of the holes in the p-type will diffuse and flow from the p-type material to the n-type. Conversely, some of the electrons in the n-type will diffuse from the n-type material to the p-type. This leads to a region between the two materials called the *depletion region*, where there are no free charge carriers. This leads to an electrical field, $\hat{E}$, between the two layers. Thus, there will be a built-in electrical potential, V_{bi}, between these two layers. The field, $\hat{E}$, and the potential, V_{bi}, will depend on the materials used. There can also be an externally applied voltage, V, between the two layers, which leads to an effective potential between the two layers given by $V + V_{bi}$. If the externally applied voltage has an electric field in the same direction as the electric field caused by the built-in potential, we call it *forward bias*, and when the electric field caused by the applied voltage has the opposite direction as the electric field caused by the built-in potential, we call it *reverse bias*. The sign of the potential, $V/|V|$, will then determine the direction of the flow.

1.2.3 The process of creating electricity from light

A silicon solar cell converts solar energy into electrical power by converting the energy from photons into excited charges, separating these charges with an electric field, and transporting them through an external circuit

This subsection is based on [19]*, Section 3.1.*

The energy of a photon with frequency, f, is given by hf, where h is Planck's constant. When a photon with an energy, hf, hits an electron in the valence band with an energy lower than the band gap, i.e. when $hf < E_g$, the photon energy cannot be absorbed by the electron. When an electron in the valence band gets hit by a photon with an energy higher than the band gap, however, the electron jumps to the conduction band. There is then a hole in the valence band where there used to be an electron. The electron and hole then make up an electron-hole pair, denoted e−h pair, and also known as an *exciton*. The closer this pair is to the p−n junction when created, the better the chances are for this pair to lead to electrical generation.

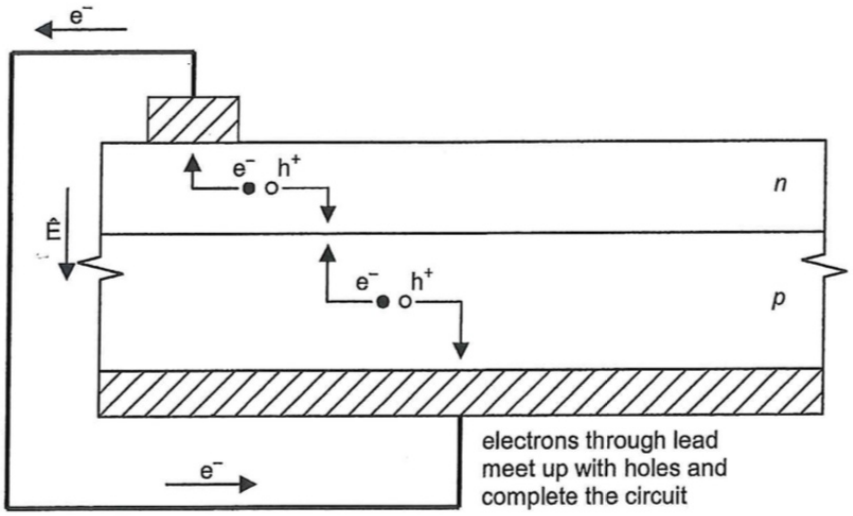


Figure 1.10: Illustration of how a solar cell creates electricity from light. Courtesy of [19].

If the e−h pair is not recombined, the pair may split up due to the electric field, $\hat{E}$. The electron will take the external circuit from the n-type to the p-type, and the hole will go from the n-type to the p-type through the p−n junction, and thus the process is completed. After one such collection, there is still an equal number of electron and holes in each layer, so there is nothing that is "used up" or that the silicon cell can "run out of".

1.2.4 Problems to be solved with silicon solar cells

There is still a long way to go to optimize silicon solar cell efficiency

The following paragraph is based on [18]*, Subsection 10.3.1.*

There are several electrical losses in silicon solar cells. Some of these come from metal-finger coverage of the top surface, top-surface reflection loss, and imperfect light trapping in the cell. There are also losses due to electrical resistance within the cell.

The problem of electrical losses due to electrical resistance within the cell is what we will look into and try to solve equations for in Chapter 2.

1.3 How organic solar cells work

Organic solar cells work by having two layers of organic material with different electron affinity and ionization potential, with current carriers that do work in an external circuit

1.3.1 Electron affinity and ionization potential

Electron affinity and ionization potential describes how easily an electron is added or removed, respectively, from an atom

The following paragraph is based on parts from [14].

A solar panel needs an asymmetry of charge to operate. This asymmetry in an inorganic solar cell is achieved by doping. In an organic solar cell, however, the asymmetry is achieved by using two different materials with different affinities for charge. One material, known as the acceptor, supports negative charge carriers, i.e. free electrons, whereas the other material, known as the donor, supports positive charge carriers, i.e. holes. The important properties of these materials are electron affinity, for acceptors, and ionization potential, for donors.

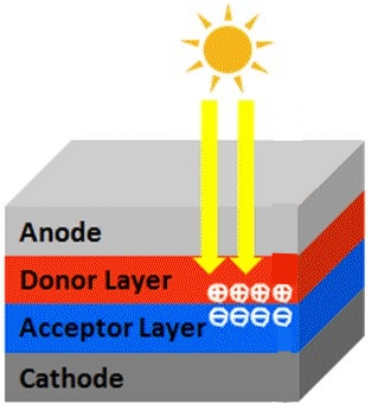
For donor materials, we define the ionization potential as the energy necessary to remove an electron from a neutral atom [12]. Conversely, for acceptor materials the electron affinity is defined as the energy that is released when an electron is added to a neutral atom [5].

1.3.2 The DA interface

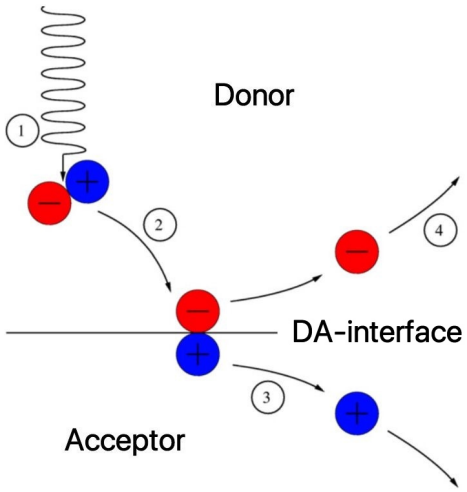
The DA interface is the interface between the donor and acceptor layer in the organic solar cell

This subsection is based on parts from [3].

In an inorganic solar cell, excitons produced has electrons and holes that can move independently within the cell. In polymer solar cells, however, excitons produced have electrons and holes that cannot move independently of each other because they are bound strongly together by coulombic attraction. These excitons must be split at the DA interface in order to generate electricity. Therefore, the location and shape of the DA interface play a crucial role in the performance of an organic solar cell. A useful measure for how long an average charge carrier can move before it is recombined is the *Debye length*, defined by the distance needed to reduce the electric potential by $1/q$, where q is the absolute charge of an electron/hole. The Debye length is often referred to as the distance over which significant charge separation can occur. This measure of length is also useful in modeling of silicon solar cells.



(a) Illustration of the layers in a typical organic solar cell. Courtesy of TCI Europe N.V.



(b) Illustration of a photon hitting a polymer solar cell and creating electricity. 1) photon absorption, 2) exciton diffusion to the DA interface, 3) dissociation of exciton into free charge carriers and 4) drift of free charge carriers to electrodes. Courtesy of [3].

Figure 1.11

1.3.3 The process of creating electricity from light

An organic solar cell converts solar energy into electrical power by converting photons into excited charges, separating these charges at the DA interface, and transporting them through an external circuit

This subsection is based on parts from [3].

When a photon, with sufficient energy, is absorbed by either of the donor or acceptor layer, an exciton is produced. These excitons are strongly bound and can only diffuse over short distances, around 10 nm, prior to recombination. If an exciton can diffuse to a DA interface, then the local variation in electron affinity and ionization potential can split the bound pair. When the exciton has been split, the free charge carriers can be transported by the internal electric field towards the electrodes and, thus, this current can be used in an external circuit and

create electrical energy.

1.3.4 Problems to be solved with organic solar cells

There is still a long way to go in order for organic solar cells to compete with other types of panels in terms of efficiency

This subsection is based on [18]*, Section 10.4.*

In organic solar cells, carrier mobilities are very low because their transport processes are dominated by charge carriers that hop, and therefore the thickness of the upper layer is limited to a few hundred nanometers to achieve low series resistance. However, organic semiconductors show strong absorption of UV-light and visible regions picking up photons with wavelengths $80 - 200$ nm. Thus, only a 100 nm thick upper layer is enough for effective absorption. Currently, the power conversion efficiency is 13.2% ([18], states 5.7% which is no longer true [10]), but organic solar cells are an attractive alternative because of their low cost of production.

Also, in organic solar cells, the e$-$h pair upon absorption of a photon of sufficient energy forms a tightly bound state exciton, as mentioned earlier. In general, only 19% of the excitons dissociate into free carriers, while the remaining exciton decay via recombination after a short time.

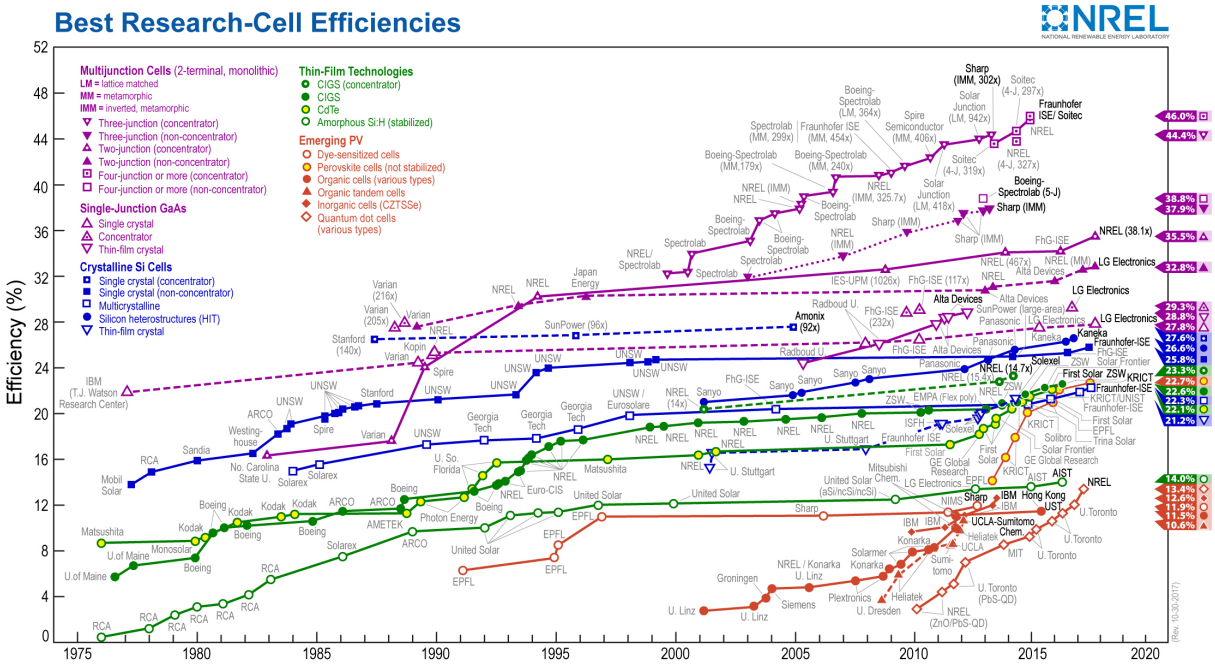


Figure 1.12: Efficiency of different types of panels through the years. Courtesy of Office of Energy Efficiency and Renewable Energy (EERE), U.S. Department of Energy. Reviewed 30.10.2017, not updated with record efficiency for organic cells [10].

1.4 Poisson's equation for electrostatics

Poisson's equation for electrostatics describes the potential field caused by a given charge

This section is based on [18], *Section 2.7.*

We define ρ as the *space charge density* as the sum of charge carrier densities; $\rho := q(p - n)$, where p is the hole concentration and n is the electron concentration. The electric field, $\hat{E}$, caused by caused by these charges can be expressed with the *electrostatic potential*, ϕ, as

$$\hat{E} = -\phi'. \tag{1.1}$$

We have Poisson's equation given by

$$\hat{E}' = \frac{|\rho|}{\epsilon} \tag{1.2}$$

where ϵ is the *dielectric permittivity* for the specific material. Using (1.1), (1.2) and the expressions for ρ, we get

$$\phi'' = -\frac{q|p-n|}{\epsilon}. \tag{1.3}$$

1.5 The drift-diffusion model

Drift and diffusion are the two main processes that make the charge carriers in a solar panel create electricity

This section is based on [18]*, Section 2.1-2.2.*

1.5.1 Drift

Drift current is the transport of carriers when an electric field is applied

The kinetic energy of electrons/holes is given by

$$\frac{1}{2}mv^2 = \frac{3}{2}k_B T \tag{1.4}$$

where m is the *effective mass* of an electron/hole and v_{th} is the *average thermal velocity*. In a semiconductor, a charge carrier will move rapidly in all directions due to thermal motion. Due to this motion, charge carriers collide with each other within the silicon lattice. We define the average time between each collision as *mean free time*, τ.

When an electric field, $\hat{E}$, is applied to a semiconductor, each carrier particle will experience a force, $\pm q\hat{E}$, ($-$ for electrons and $+$ for holes) from the field and

will be accelerated in a direction during the time between each collision. Therefore, there will be an additional velocity component called the *drift velocity*. This velocity (for an electron) can be obtained from

$$-q\hat{E}\tau_n = m_e v_n \tag{1.5}$$

where q is the charge of an electron and $m_n = m_e$ is the electron mass. Equation (1.5) can be rewritten as

$$v_n = -\left(\frac{q\tau_n}{m_e}\right)\hat{E}. \tag{1.6}$$

From this, we define the proportionality factor

$$\mu_n := \frac{q\tau_n}{m_e} \tag{1.7}$$

which is called the *electron mobility*. We can also define the *hole mobility* using $m_p = m_e$ to get

$$\mu_p := \frac{q\tau_p}{m_e}. \tag{1.8}$$

The electron drift current flux, $J_{n,\text{drift}}$, is

$$J_{n,\text{drift}} = qn\mu_n\hat{E} \tag{1.9}$$

where n is the electron density. By using that

$$\hat{E} = -\phi' \tag{1.10}$$

we get

$$J_{n,\text{drift}} = -q\mu_n n\phi'. \tag{1.11}$$

1.5.2 Diffusion

Diffusion current is the transport of carriers from a high concentration of carriers to a low concentration

The diffusion current flow, F, is

$$F = \frac{k_B T}{q}\,\mu_n n'. \tag{1.12}$$

The diffusion current will then be

$$J_{n,\text{diffusion}} = -qF = -q\,\frac{k_B T}{q}\,\mu_n n' = -\mu_n k_B T n'. \tag{1.13}$$

The quantity $k_B T/q$ is known as the *thermal voltage*.

1.5.3 The drift-diffusion model

The drift-diffusion model is the drift model and the diffusion model combined

The drift-diffusion model is simply an expression of the total flux as a sum of the drift and the diffusion flux;

$$\begin{aligned} J_n &= J_{n,\text{drift}} + J_{n,\text{diffusion}} \\ &= -q\mu_n n\phi' - \mu_n k_B T n' \\ &= -\mu_n \left(qn\phi' + k_B T n'\right). \end{aligned} \tag{1.14}$$

We would like to write J independent of charge, because the structure of the solutions of the differential equations is more important than the scale. Hence, we define $J^* := J/q$ and get

$$\begin{aligned} J_n^* &= -\frac{\mu_n}{q}\left(qn\phi' + k_B T n'\right) \\ &= -\mu_n \left(n\phi' + \frac{k_B T}{q} n'\right). \end{aligned} \tag{1.15}$$

Similarly, for holes, we get

$$J_p^* = -\mu_p \left(p\phi' - \frac{k_B T}{q} p'\right). \tag{1.16}$$

1.6 Recombinations

Recombination is when an electron in a high energy state, feasible to conduct electrical energy, loses this energy to another form

This subsection is taken from [18], *Section 2.3.*

For a solar cell in thermal equilibrium, we have that $pn = n_i^2$, where p is the hole density in the valence band and n is the electron density in the conduction band, and n_i is the carrier density for the intrinsic material. If we introduce excess carriers to a solar cell such that $pn > n_i^2$, we have a non-equilibrium. The process of introducing excess carriers is called *carrier injection*. Most solar cell devices operate by the creation of charge carriers in excess of the thermal equilibrium values.

Whenever the thermal-equilibrium condition is disturbed, i.e. $pn \neq n_i^2$, there will be processes that will restore the system equilibrium, so that $pn = n_i^2$. In the case of injection of excess carriers, the mechanism that restores equilibrium is *recombination* of the injected minority carriers with the majority carriers. Depending on the recombination process, the energy released from the recombination process can be emitted as a photon, known as radiative recombination or be turned into as heat to the lattice.

The recombination is expected to be proportional to the number of electrons available in the conduction band and holes available in the valence band; that is

$$R = \beta np \tag{1.17}$$

where β is a proportionality constant and n and p are the charge carrier densities. In thermal equilibrium, the recombination rate must be equal to the generation rate, G. So, for an n-type material in equilibrium, we have

$$G = R. \tag{1.18}$$

The net rate of change for hole concentration is given by

$$\dot{p_n} = G - R. \tag{1.19}$$

In steady state, we thus have $G = R$, which means that the system then is in equilibrium.

1.7 Continuity equation

The continuity equation is a combination of the overall effect of drift, diffusion, and recombination simultaneously

The section is taken from [18]*, Section 2.4.*

The continuity equation for electrostatics is a one-dimensional case of the Divergence Theorem (Gauss's theorem). It explains that the net flow must be equal to the sum of all the sources (and sinks) of the flow. For the net flow, when the sum of sources is zero we have

$$J_n' - q\dot{n} = 0. \tag{1.20}$$

When the sum of forces is the net generation rate, the generation rate minus the recombination rate, we get the general continuity equation for electrostatics

$$J_n' - q\dot{n} = G_n - R_n. \tag{1.21}$$

A similar equation for holes is

$$J_p' + q\dot{p} = G_p - R_p. \tag{1.22}$$

Chapter 2

Solving equations for silicon solar cells - contact resistance

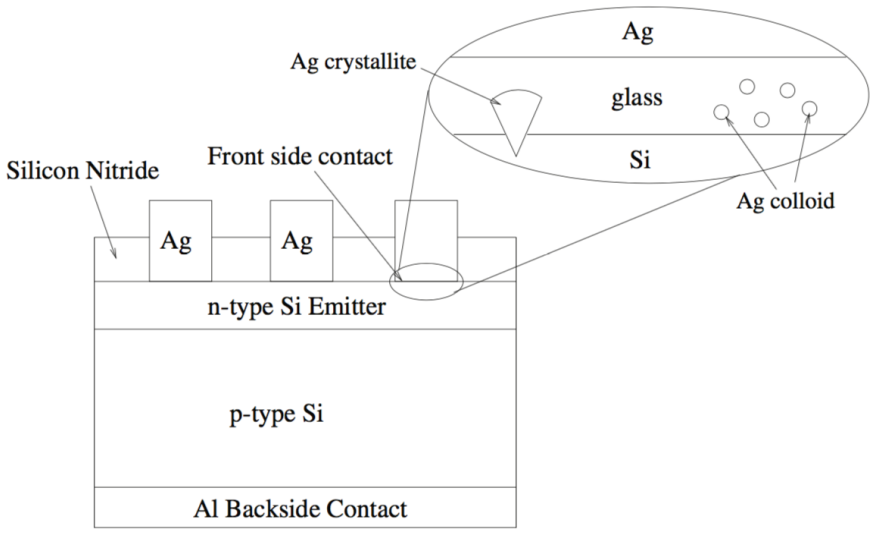


Figure 2.1: Illustration of the solar cell for which we solve equations. Black et al. [2] are considering the front side silver contact between the silver and the n-type silicon emitter. Courtesy of [2].

In this chapter, we look at [2], "Mathematical modeling of contact resistance in silicon photovoltaic cells", by Black et al. In this paper, they consider a silicon solar cell with silver (Ag) contacts above the glass that lies between the silicon

nitride and the n-type silicon. The silicon nitride works like an anti-reflective layer, and the n-type silicon, which they call emitter, is where the electrons that contribute to electron transport are excited. They want to "analyze a simple model for electron transport across [the glass layer] based on drift-diffusion equation equations." They look at solar cells where silver paste is placed on and then thermally etched into the upper n-type silicon layer. A dilemma that occurs at the layer between the silver electrode and the silicon emitter is discussed; "the emitter must be highly doped with phosphorous atoms at the surface" to attain low contact resistance, but this "leads to an increased amount of e−h recombination at the front surface", which reduces the solar cell efficiency. Thus, they seek an "enhanced understanding of the local electron transport mechanism" in order to develop "silver pastes that achieve low contact resistance" and thus pave the way for more efficient solar cells.

Formulation of the problem

We want to look at the cross-section of the panel, as sketched in Figure 2.1. We consider the glass layer that lies between the silicon emitter (n-type) on the bottom and the silver electrode on the top. Since a cell is typically much bigger in order of magnitude than the glass thickness, we can ignore the effect on the edges of the cell and reduce the problem to a one-dimensional spatial problem where the space variable is the vertical axis of the glass in Figure 2.1. We denote the glass thickness, L, and let the applied electric potential difference from the top of the glass layer to the bottom be φ. The panel has silver crystallites embedded into the silicon layer, and silver colloid within the glass itself. These are introduced to see how these two different materials influence electric conductivity. When we look at the case when there is silver on both side of the glass, we expect the electron density to equal of approximately equal on both sides. When we look at when there is silicon on the bottom, however, we expect the electron density to be lower for the silicon than for the silver, as silver has more free electrons to conduct electricity (silver is a conductor, silicon is a semiconductor). By solving different equations, we want to find a dynamic system where we can tune parameters and calculate parameters values depending on the given parameters. Subsequently, we want to look at how different cases for different boundary materials and parameter values change the resistance through the glass.

Governing equations

Black et al. look at the drift-diffusion equations and electrostatic potential equation to study this system. They consider the drift-diffusion equation, Equation (1.15), described in Subsection 1.5.3:

$$J(x) = -\mu_n \left(n(x)\phi'(x) + \frac{k_B T}{q} n'(x) \right) \tag{2.1}$$

and Poisson's equation for electrostatics, Equation (1.3) as described in Section 1.4. They consider these equations on the interval $0 \leq x \leq L$ in space, where L is the thickness of the glass. For Poisson's equation for electrostatics, it is argued that for an n-type material, the electron mobility, exceeds the hole density by a factor of 10^{10} [11], so we can neglect p in Equation (1.3). We thus end up with the following equation:

$$\phi''(x) = -\frac{qn(x)}{\epsilon}. \tag{2.2}$$

From here on, we omit the notation where we write each variable as a function of x. Now, a usual nondimensionalization is done to simplify the variables and the domain; the electrostatic potential is scaled with $\frac{k_B T}{q}$, and the electron density is scaled so that the value at $x = 1$ is unity. The domain is scaled to be $0 \leq x \leq 1$. Also, the parameter, ν, is introduced and denotes the ratio between the Debye length, L_D, and the thickness of the glass, L, where we have $L_D = \frac{\sqrt{\epsilon k_B T}}{q\sqrt{n(1)}}$. Also, the notation is (a bit confusingly) not changed to denote which variables that correspond to the original variables or the nondimensionalized variables. For more details on the nondimensionalization, one can look at [2], Section 2.3. We end up with the following nondimensionalized system:

$$\nu^2 \phi'' = -n \tag{2.3}$$

$$J = -n\,\phi' - n' \tag{2.4}$$

on $0 \leq x \leq 1$ with boundary conditions

$$\phi(0) = \boldsymbol{\varphi};\ \phi(1) = 0;\ n(1) = 1. \tag{2.5}$$

We now have that $n(0)$ is the ratio between the electron concentration at $x = 0$ and at $x = 1$. This value, sometimes denoted α, can be chosen to be given or a value to be computed (I will use the notation $n(0)$ when we want to compute

this value and α when this value is used as a boundary condition). In this system we have $n(1) = 1$, $\phi(1) = 0$ and ν as fixed values for each case we consider. The variables $n(0)$, $\boldsymbol{\varphi}$ and J can be either determined or values that we need to calculate (α for $n(0)$ if it is a given value). Since (2.3) and (2.4) is a set of equations where one equation has a second-order derivative and the other has a first-order derivative, we should have three boundary conditions. We can then choose if $\boldsymbol{\varphi}$ or α should be the fixed on beforehand, and then have J as given value. As we will see later, we can instead have α and $\boldsymbol{\varphi}$ as given values and J as the value computed. It is more realistic that we can know α and $\boldsymbol{\varphi}$ on beforehand so that J is the value that we need to compute. Black et al. make a system where α is given, and then fix $\boldsymbol{\varphi}$ to find J, or vice versa. I will instead first start with a system that has a given $\boldsymbol{\varphi}$ and J so that $n(0)$ is calculated. In reality, $\boldsymbol{\varphi}$ can easily be tuned. By building the panels differently, we can have some control over the values of ν and α, but we can never predetermine J. This means that if we do use a given fixed J, we can see it as an approach where we know what result we want and study how we can tune and tweak parameters that yield the desired J. The approach where we fix four boundary values and then find the flux is considered in Section 2.4.

In all cases, we consider when this system is in equilibrium, that is, when nothing changes with time. We can in such a case use that $\dot{n} = 0$, and that the generation is equal to the recombination, $G_n = R_n$, which by (1.21), means that $J' = 0$, making J constant in space.

Equation (2.3) and (2.4) make up a coupled set of differential equations, where (2.4) is non-linear. These equations cannot easily be solved analytically, so we try to solve them numerically. We discretize in space as described in A.1 by letting $\widetilde{\boldsymbol{x}}$ be the discretization vector of x. We find discrete solutions, $\widetilde{\boldsymbol{\phi}}$ for ϕ, and $\widetilde{\boldsymbol{n}}$ for n, as functions of $\widetilde{\boldsymbol{x}}$, with $N + 2$ discretization points with spacing $\Delta\, x = (N + 1)^{-1}$. We have, as described in A.1, $\widetilde{\boldsymbol{\phi}} = [\phi_0, \ldots, \phi_{N+1}]^T$, $\hat{\boldsymbol{\phi}} = [\phi_0, \ldots, \phi_N]^T$, $\boldsymbol{\phi} = [\phi_1, \ldots, \phi_N]^T$ and $\widetilde{\boldsymbol{n}} = [n_0, \ldots, n_{N+1}]^T$, $\hat{\boldsymbol{n}} = [n_0, \ldots, n_N]^T$, $\boldsymbol{n} = [n_1, \ldots, n_N]^T$. Consequently, we have $\phi_0 = \boldsymbol{\varphi}$, $\phi_{N+1} = 0$ and $n_{N+1} = 1$. We here use hat and tilde notation to increase readability.

We want to find ϕ and n by solving (2.3) and (2.4). We discretize these equations to find and we now seek $\boldsymbol{\phi}$ and $\hat{\boldsymbol{n}}$. Since we have boundary conditions on both sides for ϕ for Equation 2.3, we use second-order finite central differences and hence use Lemma A.1. Since we have a right boundary condition for n in

Equation (2.3), we use first-order finite forward differences and hence use Lemma A.3. By Corollary A.2 and A.4, we get the following system in matrix form:

$$S_0\ \boldsymbol{\phi} + \boldsymbol{r}_{\phi,S_0} = -\frac{(\Delta\, x)^2}{\nu^2}\boldsymbol{n} \tag{2.6}$$

$$\left\{\operatorname{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+}\right\}\hat{\boldsymbol{n}} + \boldsymbol{r}_{n,\hat{F_+}} = -\Delta\, x\ \ J\,\mathbf{1}_{N+1} \tag{2.7}$$

We later consider what happens if we increase the order. We will do this by using second-order finite forward differences for (2.4), with a first-order approximation at x_N. When we mix orders of finite differences in one system, however, we cannot be sure what order we get; we can at best get a method with the order of the highest of the different approximations, and at worst, we get an order of the lowest of the different approximations. When we use different orders for (2.3) and (2.4), we will get an order of the lowest of the two.

We look at three different cases for the parameter values. These are stated in the table below. In the first case, Scenario 1, we look at what happens when we have all values as unity, somewhat for reference. In the next case, Scenario 2, we look at what happens in the case when the electron flux is large, $J \gg 1$, that is, we see how a large flux influences the system. In the last case, Scenario 3, we look at what happens when the Debye length to the glass thickness ratio is small, $\nu \ll 1$, where the Debye length is the length over which significant charge separation can occur. These three scenarios can be considered cases where we look at the problem physically and see what happens in different limits. These three scenarios may not represent typical solar cell parameters, but rather give us an idea for how the electronic properties will be for different limits of the parameters.

Scenario	$\boldsymbol{\varphi}$	ϕ_{N+1}	n_{N+1}	J	ν
1	1	0	1	1	1
2	100	0	1	70	1
3	0.05	0	1	2×10^{-4}	0.01

We first try to solve for $\boldsymbol{\phi}$, $\hat{\boldsymbol{n}}$ with $\boldsymbol{\varphi}$ and J as given values, so that we need to include n_0 as an unknown when finding the normalized electron density.

2.1 Solving using an iteration scheme

We rewrite (2.6) and (2.7), by putting terms on the right side, so that we can solve for $\boldsymbol{\phi}$ and $\hat{\boldsymbol{n}}$:

$$S_0 \ \boldsymbol{\phi} = -\boldsymbol{r}_{\phi,S_0} \ - \left(\frac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n} \tag{2.8}$$

$$\left\{\operatorname{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+}\right\}\hat{\boldsymbol{n}} = -\boldsymbol{r}_{n,\hat{F_+}} \ - \Delta\, x \ \ J\,\mathbf{1}_{N+1} \tag{2.9}$$

This suggests that we can find $\boldsymbol{\phi}$ and $\hat{\boldsymbol{n}}$ using the following iteration scheme:

$$\boldsymbol{\phi}^{(k+1)} = -S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0} \ + \left(\frac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n}^{(k)}\right) \tag{2.10a}$$

$$\hat{\boldsymbol{n}}^{(k+1)} = -\left\{\operatorname{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}}^{(k+1)} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+}\right\}^{-1}(\boldsymbol{r}_{n,\hat{F_+}} \ + \Delta\, x\, J\,\mathbf{1}_{N+1}) \tag{2.10b}$$

For (2.10a), we can consider $\boldsymbol{n}$ as a given value, and for (2.10b), we can consider $\hat{\boldsymbol{\phi}}$ as a given value. Thus, in each step, we actually solve a linear system.

We use the relative error using the 2-norm between each consecutive iteration. We exit the iteration when the largest relative error (either the error for $\boldsymbol{\phi}$ or for $\hat{\boldsymbol{n}}$) is less than 10^{-14}. The reason we choose 10^{-14} is that the rounding error in MATLAB is less than 10^{-16}. This means that for a point, x, which is represented on the computer as the floating point of x, $\text{fl}(x)$, we have that $|\text{fl}(x) - x|/|x| \leq 10^{-16}$. When using the 2-norm, these rounding errors add up, so 10^{-14} makes a decent approximation for the maximum relative error due to rounding errors. If the scheme does not converge in 200 or fewer steps, we also exit the iteration and say that the method does not converge. I will use these conditions for all further iteration methods.

In the calculation in (2.10a) and (2.10b), it is best to let $\boldsymbol{\phi}$ and $\hat{\boldsymbol{n}}$ be the values that solve Equation (2.8) and (2.9), respectively. The system can then be solved with a faster method, as inverting the matrix and multiplying may be slower than other solution methods. MATLAB has a built-in code that tries to find the fastest solution method for a given system. This is implemented as using backslash, \. So to find $\boldsymbol{x}$ that solves $\boldsymbol{x} = A^{-1}\boldsymbol{y}$, it may be faster to solve if implemented as `x = A\y`.

2.1.1 Convergence analysis of the iteration scheme

We can analyze the iteration scheme by writing it into a fixed-point method. We want to see if the given values for ν in the different scenarios will lead to convergence, and which other values of ν will give convergence.

We will consider (2.10b) as the fixed-point method. We then want to write $\hat{\boldsymbol{n}}$ in the form

$$\hat{\boldsymbol{n}}^{(k+1)} = \hat{\boldsymbol{h}}_{2.1}(\hat{\boldsymbol{n}}^{(k)}) \tag{2.11}$$

where $\hat{\boldsymbol{h}}_{2.1} : \mathbb{R}^{N+1} \to \mathbb{R}^{N+1}$ is some mapping that needs to be a contraction mapping for the fixed-point iteration to converge. In this case, $\hat{\boldsymbol{h}}_{2.1}$ is the right side of equation (2.10b). For this right side, we have $\hat{\boldsymbol{\phi}}$ as input argument, but we want to have $\hat{\boldsymbol{n}}$ as input. We want to analyze the following fixed-point iteration:

$$\begin{aligned} &\hat{\boldsymbol{n}}^{(k+1)} = \hat{\boldsymbol{h}}_{2.1}(\hat{\boldsymbol{\phi}}^{(k+1)}) \\ &= -\{\mathrm{diag}(\hat{F_+}\ \hat{\boldsymbol{\phi}}^{(k+1)} + \boldsymbol{r}_{\phi,\hat{F_+}}\ ) + \hat{F_+}\ \}^{-1} \left(\boldsymbol{r}_{n,\hat{F_+}}\ + \Delta\, x\, J\, \mathbf{1}_{N+1}\right) \end{aligned} \tag{2.12}$$

To simplify the analysis, we instead find an expression in the form $\boldsymbol{n} = \boldsymbol{h}_{2.1}(\boldsymbol{\phi})$ and write it in the form $\boldsymbol{n} = \boldsymbol{h}_{2.1}(\boldsymbol{n})$. By doing so, we reduce the $(N+1)\times(N+1)$ system in (2.12) to an $N \times N$ system and replace $\hat{F_+}$, $\hat{\boldsymbol{\phi}}$ and $\mathbf{1}_{N+1}$ with F_+, $\boldsymbol{\phi}$ and $\mathbf{1}_N$, respectively. We then get

$$\begin{aligned} \boldsymbol{n}^{(k+1)} &= \boldsymbol{h}_{2.1}(\boldsymbol{\phi}^{(k+1)}) \\ &= -\{\mathrm{diag}(F_+\ \boldsymbol{\phi}^{(k+1)} + \boldsymbol{r}_{\phi,F_+}\ ) + F_+\ \}^{-1} \left(\boldsymbol{r}_{n,F_+}\ + \Delta\, x\, J\, \mathbf{1}_N\right). \end{aligned} \tag{2.13}$$

To obtain an expression for $\boldsymbol{n}^{(k+1)}$ with $\boldsymbol{n}^{(k)}$ as input, we substitute $\boldsymbol{\phi}^{(k+1)}$ with the expression in (2.10a):

$$\begin{aligned}\boldsymbol{n}^{(k+1)} &= \boldsymbol{h}_{2.1}\left(\boldsymbol{n}^{(k)}\right) \\ &= -\left\{\operatorname{diag}\left(F_+\,\left[-S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0}+\left(\tfrac{\Delta x}{\nu}\right)^2\hat{\boldsymbol{n}}^{(k)}\right)\right]+\boldsymbol{r}_{\phi,F_+}\right)+F_+\,\right\}^{-1} \\ &\qquad\qquad (\boldsymbol{r}_{n,F_+}\,+\Delta x\, J\,\mathbf{1}_N) \end{aligned} \tag{2.14}$$

We now have a mapping in the form $\boldsymbol{n}^{(k+1)} = \boldsymbol{h}_{2.1}(\boldsymbol{n}^{(k)})$, which is a fixed-point iteration. A sufficient and necessary condition for this fixed-point method to converge is that the mapping, $\boldsymbol{h}_{2.1}$, is a contraction mapping, as stated in Theorem A.8. We require $\boldsymbol{n}^{(k)} \in \mathcal{R},\ \ k = 1, 2, \ldots$ where $\mathcal{R} \subseteq \mathbb{R}$ is a complete subset of $\mathbb{R}$. For a mapping to be a contraction mapping, it is sufficient that the norm of Jacobian, $\boldsymbol{h}_{2.1}(\boldsymbol{n})$, is $\left\|\mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}(\boldsymbol{n})\right\| < 1,\ \forall \boldsymbol{n} \in \mathcal{R}$, for some norm.

We now omit the iteration index. Using that diag() is a linear operator and that a minus can be distributed in a matrix inverse, we get

$$\begin{aligned}&\boldsymbol{h}_{2.1}(\boldsymbol{n}) = \\ &\quad\left\{\operatorname{diag}\left(F_+\,S_0^{-1}\boldsymbol{r}_{\phi,S_0}\right)+\operatorname{diag}\left(F_+\;S_0^{-1}\left(\tfrac{\Delta x}{\nu}\right)^2\boldsymbol{n}\right)-\operatorname{diag}\left(\boldsymbol{r}_{\phi,F_+}\right)-F_+\,\right\}^{-1} \\ &\qquad\qquad (\boldsymbol{r}_{n,F_+}\,+\Delta x\, J\,\mathbf{1}_N)\,. \end{aligned} \tag{2.15}$$

It is not straightforward to find the Jacobian of $\boldsymbol{h}_{2.1}$, so we use Claim B.7 to find the j-th column of the Jacobian of $\boldsymbol{h}_{2.1}(\boldsymbol{n})$, for $j = 1, \ldots, N$:

$$\begin{aligned}&\left[\mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}\right]_{:,j} = \\ &\tfrac{\partial}{\partial n_j}\left\{\operatorname{diag}\left(F_+\,S_0^{-1}\boldsymbol{r}_{\phi,S_0}\right)+\operatorname{diag}\left(F_+\;S_0^{-1}\left(\tfrac{\Delta x}{\nu}\right)^2\boldsymbol{n}\right)-\operatorname{diag}\left(\boldsymbol{r}_{\phi,F_+}\right)-F_+\,\right\}^{-1} \\ &\qquad\qquad (\boldsymbol{r}_{n,F_+}\,+\Delta x\, J\,\mathbf{1}_N) \end{aligned} \tag{2.16}$$

We can now use Theorem B.9 to find another expression for the derivative of the inverse:

$$\begin{aligned}\left[\mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}\right]_{:,j} = - \quad & \left\{\operatorname{diag}\left(F_+\, S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0} + \left(\tfrac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n}\right) - \boldsymbol{r}_{\phi,F_+}\right) - F_+\right\}^{-1} \\ & \left(\tfrac{\Delta\, x}{\nu}\right)^2 \tfrac{\partial}{\partial n_j}\left\{\operatorname{diag}\left(F_+\, S_0^{-1}\boldsymbol{n}\right)\right\} \\ & \left\{\operatorname{diag}\left(F_+\ S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0} + \left(\tfrac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n}\right) - \boldsymbol{r}_{\phi,F_+}\right) - F_+\right\}^{-1} \\ & \left(\boldsymbol{r}_{n,F_+} + \Delta\, x\, J\, \mathbf{1}_N\right) \end{aligned} \tag{2.17}$$

where we have used that diag() is linear and that terms with no dependence on $\boldsymbol{n}$ vanish under differentiation. Using Claim B.8 we get

$$\begin{aligned}\left[\mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}\right]_{:,j} = - \quad & \left\{\operatorname{diag}\left(F_+\, S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0} + \left(\tfrac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n}\right) - \boldsymbol{r}_{\phi,F_+}\right) - F_+\right\}^{-1} \\ & \left(\tfrac{\Delta\, x}{\nu}\right)^2 \left\{\operatorname{diag}\left(\left[F_+\, S_0^{-1}\right]_{:,j}\right)\right\} \\ & \left\{\operatorname{diag}\left(F_+\ S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0} + \left(\tfrac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n}\right) - \boldsymbol{r}_{\phi,F_+}\right) - F_+\right\}^{-1} \\ & \left(\boldsymbol{r}_{n,F_+} + \Delta\, x\, J\, \mathbf{1}_N\right). \end{aligned} \tag{2.18}$$

We now find a bound for the norm of the Jacobian. In this case, we have an explicit expression for each column of the Jacobian so it is natural to choose the 1-norm, also known as the "column norm". We then have

$$\left\|\mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}\right\|_1 = \max_{j=1,\ldots,N}\left\|\left[\mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}\right]_{:,j}\right\|_1. \tag{2.19}$$

We calculate the 1-norm of column j:

$$\begin{aligned}\left\|\left[\mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}\right]_{:,j}\right\|_1 = \quad & \Big|\Big| - \left\{\operatorname{diag}\left(F_+\ S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0} + \left(\tfrac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n}\right) - \boldsymbol{r}_{\phi,F_+}\right) - F_+\right\}^{-1} \\ & \left(\tfrac{\Delta\, x}{\nu}\right)^2 \left\{\operatorname{diag}\left(\left[F_+\, S_0^{-1}\right]_{:,j}\right)\right\} \\ & \left\{\operatorname{diag}\left(F_+\, S_0^{-1}\left(\boldsymbol{r}_{\phi,S_0} + \left(\tfrac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n}\right) - \boldsymbol{r}_{\phi,F_+}\right) - F_+\right\}^{-1} \\ & \left(\boldsymbol{r}_{n,F_+} + \Delta\, x\, J\, \mathbf{1}_N\right)\Big|\Big|_1 \end{aligned} \tag{2.20}$$

$$\leq$$

$$\begin{array}{c} \left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \left(\boldsymbol{r}_{\phi,S_0} + \left(\frac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n} \right) - \boldsymbol{r}_{\phi,F_+} \right) - F_+ \; \right\}^{-1} \right\|_1^2 \\ \left(\frac{\Delta\, x}{\nu}\right)^2 \left\| \operatorname{diag}\left(\left[F_+ \, S_0^{-1} \right]_{:,j} \right) \right\|_1 \left\{ \| \boldsymbol{r}_{n,F_+} \; \|_1 + \Delta\, x \| J \, \mathbf{1}_N \|_1 \right\} \end{array} \tag{2.21}$$

$$=$$

$$\begin{array}{c} \left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \left(\boldsymbol{r}_{\phi,S_0} + \left(\frac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n} \right) - \boldsymbol{r}_{\phi,F_+} \right) - F_+ \; \right\}^{-1} \right\|_1^2 \\ \left(\frac{\Delta\, x}{\nu}\right)^2 \left\| [F_+ \, S_0^{-1}]_{:,j} \right\|_\infty \left\{ \| \boldsymbol{r}_{n,F_+} \; \|_1 + \Delta\, x \| J \, \mathbf{1}_N \|_1 \right\} \end{array} \tag{2.22}$$

where we first use (B.23b), (B.23a) and (B.23c) and next Lemma B.12. We can use that for $\left(\frac{\Delta\, x}{\nu}\right)^2 \|\boldsymbol{n}\|_1 \ll \|\boldsymbol{r}_{\phi,S_0}\|_1$, we have

$$\left\| \operatorname{diag}\left(\left(\frac{\Delta\, x}{\nu} \right)^2 F_+ \, S_0^{-1} \boldsymbol{n} \right) \right\|_1 \ll \left\| \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} \right) \right\|_1 . \tag{2.23}$$

Since glass with silver colloid has a lower electron density than both silver and silicon, we have that $0 \leq n(x) \leq 1$, so we can expect that $0 \leq n_i \leq 1$, $i = 1, \ldots, N$. We can then use that $\|\boldsymbol{n}\|_1 \leq N$, $\Delta\, x \sim 1/N$, and $\|\boldsymbol{r}_{\phi,S_0}\|_1 = \boldsymbol{\varphi}$, so (2.23) is satisfied whenever

$$\frac{1}{N\nu^2} \ll \boldsymbol{\varphi}. \tag{2.24}$$

An analysis of (2.24) for different cases is discussed later in this subsection. If (2.24) is satisfied, the expression for the bound for the Jacobian in (2.22) can be approximated with

$$\left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+} \right) - F_+ \; \right\}^{-1} \right\|_1^2 \left(\frac{\Delta\, x}{\nu}\right)^2 \left\| [F_+ \, S_0^{-1}]_{:,j} \right\|_\infty \left\{ \| \boldsymbol{r}_{n,F_+} \; \|_1 + \Delta\, x \| J \, \mathbf{1}_N \|_1 \right\} \tag{2.25}$$

$$\leq$$

$$(1 + \Delta\, x J \, N) \left(\frac{\Delta\, x}{\nu}\right)^2 \left\| [F_+ \, S_0^{-1}]_{:,j} \right\|_\infty \left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+} \right) - F_+ \; \right\}^{-1} \right\|_1^2 \tag{2.26}$$

where we use that $n_{N+1} = 1$, $J > 0$ and $\|1_N\|_1 = N$. We can now substitute $\Delta\, x$ with $1/N$ since $\Delta\, x \sim 1/N$. We then get

$$(1 + J\,) \left(\frac{1}{N\,\nu}\right)^2 \left\| [F_+ \, S_0^{-1}]_{:,j} \right\|_\infty \left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+} \right) - F_+ \; \right\}^{-1} \right\|_1^2 . \tag{2.27}$$

Using MATLAB, I find that $\max_{j=1,\dots,N} \left\| [F_+ \, S_0^{-1}]_{:,j} \right\|_\infty \simeq 1$ for $N = 10^2$, $N = 10^3$ and $N = 10^4$, so we can assume it holds for the values of N that we will use. We now have that a norm bound for the Jacobian of $\boldsymbol{h}_{2.1}$ that does not depend on j;

$$\begin{aligned} \left\| \mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})} \right\|_1 &\leq (1+J) \left(\tfrac{1}{N\,\nu}\right)^2 \left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+}\right) - F_+ \right\}^{-1} \right\|_1^2 \\ &= \tfrac{1+J}{N^2\,\nu^2} \left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+}\right) - F_+ \right\}^{-1} \right\|_1^2 . \end{aligned} \tag{2.28}$$

We now consider the three different cases for different parameters.

For Scenario 1, I used MATLAB to find

$$\left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+}\right) - F_+ \right\}^{-1} \right\|_1 \simeq 0.633N \tag{2.29}$$

so we have that

$$\left\| \mathcal{J}_{\boldsymbol{n}}^{\boldsymbol{h}_{2.1}(\boldsymbol{n})}(\boldsymbol{n}) \right\|_1 \leq \tfrac{1+J}{N^2\,\nu^2} \, (0.633N)^2 \approx \tfrac{1+J}{N^2\nu^2} \, 0.4N^2 = \tfrac{4}{5\nu^2} \tag{2.30}$$

using that $J = 1$. So we get convergence for

$$\frac{4}{5\nu^2} < 1 \iff 4 < 5\nu^2 \iff \nu > \sqrt{\frac{4}{5}} \approx 0.89. \tag{2.31}$$

So we get convergence for $\nu > 0.89$.

Also, we made the assumption in (2.23) that one term will vanish with large enough N, requiring that

$$\frac{1}{N\nu^2} \ll \boldsymbol{\varphi}. \tag{2.32}$$

For Scenario 1, this is true when $N \gg \frac{1}{\boldsymbol{\varphi}\nu^2} = 1.26$ for $\nu_{\min}$, and when $N \gg 1$ for $\nu = 1$, . This is a realistic assumption, as N is generally much larger than ~ 1.

Now, for Scenario 2, MATLAB experiments show that the norm follows the relation

$$\left\| \left\{ \operatorname{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+}\right) - F_+ \right\}^{-1} \right\|_1 \simeq N/100. \tag{2.33}$$

With similar calculations, we get convergence for $\nu > 0.085$.

We need again to investigate if the assumption made in (2.23) is realistic with the given parameter values. We have that (2.24) is fulfilled when $N \gg \frac{1}{\varphi\nu^2} = 0.7$ for $\nu_{\min}$ and when $N \gg 0.01$ for $\nu = 1$. This is, again, a realistic assumption.

Lastly, for Scenario 3, we get

$$\left\| \left\{ \text{diag}\left(F_+ \, S_0^{-1} \boldsymbol{r}_{\phi,S_0} - \boldsymbol{r}_{\phi,F_+} \right) - F_+ \; \right\}^{-1} \right\|_1 \simeq 0.975N. \tag{2.34}$$

For Scenario 3, we get convergence for $\nu > 0.975$.

Again, we need to check if the assumption in (2.23) is realistic. The condition in (2.24) is fulfilled when $N \gg 10$ for $\nu_{\min}$ and when $N \gg 2 \times 10^5$ for ν. The assumption for ν is unrealistic as this is a too great number for the computer to deal with, also, it is not common to have such high discretization numbers for solving such systems. This means that the value of $\nu_{\min}$ for Scenario 3 is not a maximum bound that can be used with confidence.

Solving a system of equations with a matrix system of size $N \times N$ with for example using Gaussian elimination takes on the $\mathcal{O}(N^3)$ operations. So for $N \gg 2 \times 10^5$, say $N = 10^7$, it will take on the order of 10^{21} operations to compute, which requires the computing power and memory of a computer much better than what an average person uses. Thus, the assumption in (2.23) in this analysis is entirely unrealistic in normal computations. With my computer with 4 cores and 2.2 GHz clock speed, it takes around 7 seconds in MATLAB to solve $F_+ \, S_0^{-1}$, which are sparse matrices, when $N = 10^4$. This means that it would take around a thousand years to compute just one Gaussian elimination in the system with my computer, if the computer indeed had the memory for such a computation. When I implemented a $10^7 \times 10^7$ matrix in MATLAB, it stated that a matrix of such a size takes around one petabyte (1 million gigabytes) to store.

These bounds, $\nu_{\min}$, are not strict, but they are bounds that guarantee convergence (maybe not for Scenario 3). We see that for $\nu = 0.01$ in Scenario 3, if we indeed had $\nu_{\min} = 1.38$, the scheme may not converge. Also, we see that the

convergence of the iteration scheme is independent of the discretization number, N, as long as (2.24) is satisfied, which is realistic for Scenario 1 and 2. This means that if the iteration scheme does not converge, we cannot expect that increasing the number of discretization points will give an increased chance for convergence. It is the value of ν that is critical for the convergence.

Results from iteration scheme

When using the iteration scheme, we get convergence for Scenario 1 and 2, but not for Scenario 3 (we exceed 200 iterations). Plots for Scenario 1 and 2 can be seen in Figure 2.2 and 2.3, respectively. For the bounds found in Subsection 2.1.1, real computations show that the bounds we have confidence in, i.e. the bounds for Scenario 1 and 2, fit well. For Scenario 1, computations show that the strict minimum is $\nu_{\min} \approx 0.14$ and for Scenario 2 we get a strict minimum bound of around $\nu_{\min} \approx 0.07$. The bound for Scenario 2 is surprisingly accurate taken into consideration that when we find the bound in 2.1.1 we use norm inequalities. For Scenario 3, we do however get convergence for $\nu_{\min} \approx 0.23$, regardless of the discretization number. A plot of the result for this value of ν for can be seen in Figure 2.4.

By the analysis of the fixed-point method, we have convergence (even for Scenario 3) for values of ν greater than the values for $\nu_{\min}$, and convergence with the actual strict bounds, denoted $\nu^*_{\min}$. These values are stated in the table below.

Scenario	ν	$\nu_{\min}$	$\nu^*_{\min}$
1	1	0.89	0.14
2	1	0.085	0.07
3	0.01	0.975	0.23

If we solve for Scenario 3 using second-order finite differences as described in Lemma A.7, we get convergence (less than 200 iterations) for $\nu = 0.22$, regardless of the discretization number. This is very close to the results that we get for the first-order finite difference method. This means that the order of the method does not influence if the iteration scheme converges or not, and neither does the size of the discretization number. This is, however, greatly influenced by the value of ν.

Plots of the results

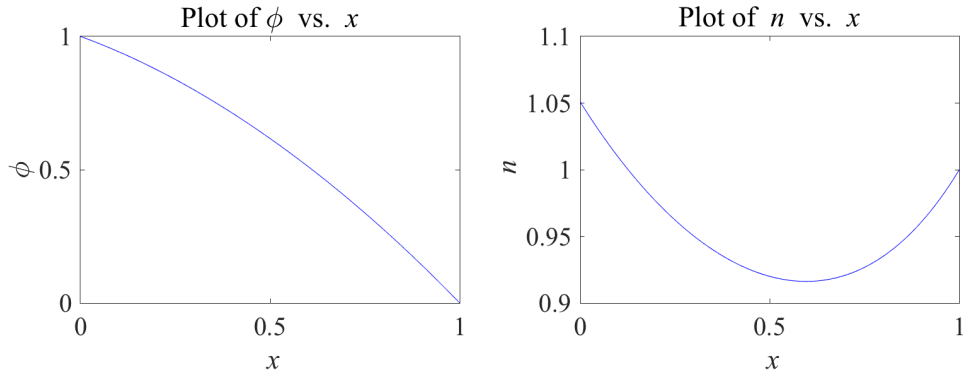


Figure 2.2: Plot of ϕ and n against x for Scenario 1. These are results when using the iteration scheme in Section 2.1 with first-order finite differences.

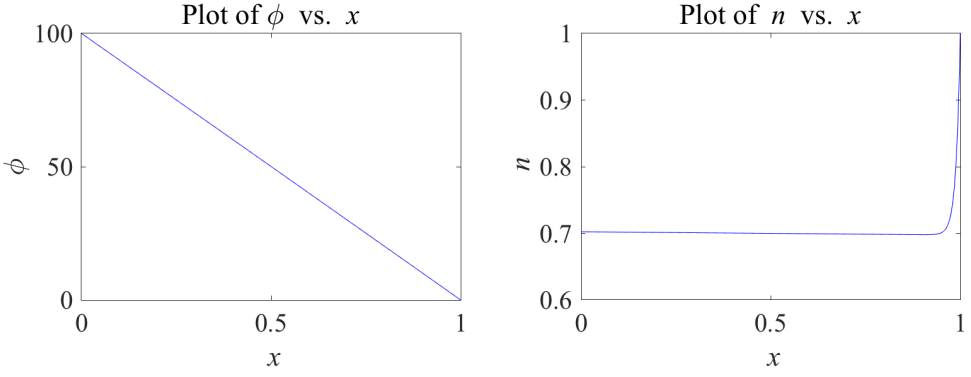


Figure 2.3: Plot of ϕ and n against x for Scenario 2. These are results when using the iteration scheme in Section 2.1 with first-order finite differences.

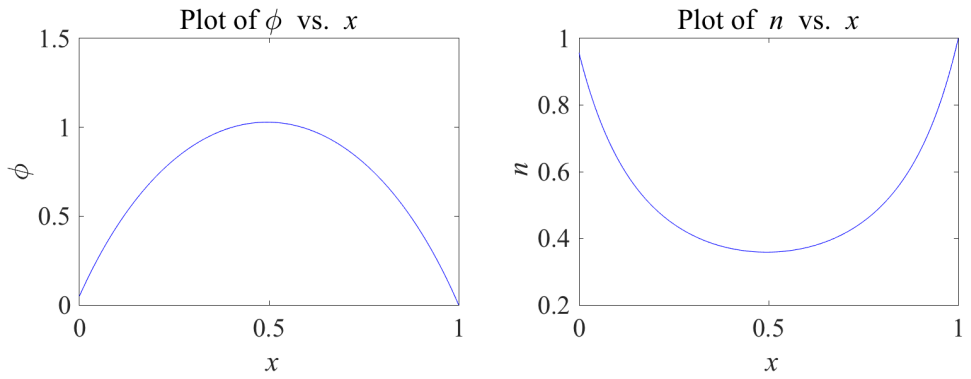


Figure 2.4: Plot of ϕ and n against x for Scenario 3, but with $\nu = 0.23$. These are results when using the iteration scheme in Section 2.1 with first-order finite differences.

Since we do not get convergence for Scenario 3, we instead seek an iteration method that is a more secure approach, leading to Newton-Raphson as a natural choice.

2.2 Solving using Newton-Raphson

The standard formulation of Newton-Raphson is to find the roots of a function. That is find $\boldsymbol{z}$ such that $\boldsymbol{f}_{2.2}(\boldsymbol{z}) = \mathbf{0}_{2N+1}$. We write the system (2.6), (2.7) to be in the desired form to solve with the Newton-Raphson method;

$$S_0\boldsymbol{\phi} + \boldsymbol{r}_{\phi,S_0} + \left(\frac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n} = \mathbf{0}_N \tag{2.35a}$$

$$\operatorname{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}})\hat{\boldsymbol{n}} + \hat{F_+}\,\hat{\boldsymbol{n}} + \boldsymbol{r}_{n,\hat{F_+}} + \Delta\, x\, J\, \mathbf{1}_{N+1} = \mathbf{0}_{N+1} \tag{2.35b}$$

We can write (2.35a) and (2.35b)

$$\begin{aligned} S_0\boldsymbol{\phi} \;+\; \left(\tfrac{\Delta\, x}{\nu}\right)^2 \boldsymbol{n} \qquad\qquad\qquad\qquad &+\; \boldsymbol{r}_{\phi,S_0} &&= \mathbf{0}_N \\ (\operatorname{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+})\hat{\boldsymbol{n}} \;+\; \Delta\, x\, J\, \mathbf{1}_{N+1} \;+\; \boldsymbol{r}_{n,\hat{F_+}} &&= \mathbf{0}_{N+1} \end{aligned} \tag{2.36}$$

We want to solve $\boldsymbol{f}_{2.2}(\boldsymbol{z}) = \mathbf{0}_{2N+1}$, using (2.36), with $\boldsymbol{f}_{2.2}(\boldsymbol{z})$ as

$$\boldsymbol{f}_{2.2}(\boldsymbol{z}) = \underbrace{\left(\begin{array}{c|c} S_0 & \mathbf{0}_N \quad \left(\frac{\Delta x}{\nu}\right)^2 I_N \\ 0 & \operatorname{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+} \end{array}\right)}_{=:G_{2.2}(\boldsymbol{z})} \left(\begin{array}{c} \boldsymbol{\phi} \\ \hat{\boldsymbol{n}} \end{array}\right) + \left(\begin{array}{c} \boldsymbol{r}_{\phi,S_0} \\ J\Delta\, x\mathbf{1}_{N+1} + \boldsymbol{r}_{n,\hat{F_+}} \end{array}\right) \tag{2.37}$$

with

$$\boldsymbol{z} = \left(\begin{array}{c} z_1 \\ \vdots \\ z_{2N+1} \end{array}\right) := \left(\begin{array}{c} \boldsymbol{\phi} \\ \hat{\boldsymbol{n}} \end{array}\right) \iff \boldsymbol{\phi} = \left(\begin{array}{c} z_1 \\ \vdots \\ z_N \end{array}\right), \hat{\boldsymbol{n}} = \left(\begin{array}{c} z_{N+1} \\ \vdots \\ z_{2N+1} \end{array}\right).$$

Here, $G_{2.2}(\boldsymbol{z})$ is a 2×2 block matrix. We can let $\{G_{2.2}\}_{1,1}$ be the upper left block and so on. Then the first N components of $\boldsymbol{f}_{2.2}(\boldsymbol{z})$ will be $\{G_{2.2}\}_{1,1}\boldsymbol{\phi}+\{G_{2.2}\}_{1,2}\hat{\boldsymbol{n}}+\boldsymbol{r}_{\phi,S_0}$. Similarly, the next $N+1$ elements of $\boldsymbol{f}_{2.2}(\boldsymbol{z})$ will be $\{G_{2.2}\}_{2,2}\hat{\boldsymbol{n}}+J\Delta x\mathbf{1}_{N+1}+\boldsymbol{r}_{n,\hat{F_+}}$.

We now find the Jacobian of $\boldsymbol{f}_{2.2}(\boldsymbol{z})$ by using Claim B.6. We find the first N rows of the Jacobian by using Claim B.1. We find the next $N+1$ rows of the Jacobian by using Claim B.1 and Claim B.3. We get the following Jacobian of $\boldsymbol{f}_{2.2}(\boldsymbol{z})$:

$$\mathcal{J}_{\boldsymbol{z}}^{\boldsymbol{f}_{2.2}(\boldsymbol{z})}(\boldsymbol{z}) = \left(\begin{array}{c|c} S_0 & \mathbf{0}_N \quad \left(\frac{\Delta x}{\nu}\right)^2 I_N \\ \operatorname{diag}(\hat{\boldsymbol{n}})[\hat{F_+}]_{:,2:N+1} & \operatorname{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+} \end{array}\right) \tag{2.38}$$

We can now solve for $\boldsymbol{z}$ with the following iteration:

$$\boldsymbol{z}^{(k+1)} = \boldsymbol{z}^{(k)} - \left[\mathcal{J}_{\boldsymbol{z}}^{\boldsymbol{f}_{2.2}(\boldsymbol{z})}\left(\boldsymbol{z}^{(k)}\right)\right]^{-1} \boldsymbol{f}_{2.2}\left(\boldsymbol{z}^{(k)}\right). \tag{2.39}$$

The expression in (2.39) can be solved more quickly by solving for $\boldsymbol{d}^{(k)}$ by letting $\boldsymbol{d}^{(k)}$ be such that

$$\mathcal{J}_{\boldsymbol{z}}^{\boldsymbol{f}_{2.2}(\boldsymbol{z})}\left(\boldsymbol{z}^{(k)}\right)\boldsymbol{d}^{(k)} = -\boldsymbol{f}_{2.2}\left(\boldsymbol{z}^{(k)}\right) \tag{2.40}$$

and then find $\boldsymbol{z}^{(k)}$;

$$\boldsymbol{z}^{(k+1)} = \boldsymbol{z}^{(k)} + \boldsymbol{d}^{(k)}, \tag{2.41}$$

as described in Subsection A.4.2.

Results from Newton-Raphson

As expected, the Newton-Raphson method converges to the same values as the method in Section 2.1, for Scenario 1 and 2. This builds confidence in the code we have developed. We now also get convergence for Scenario 3. However, this convergence only occurs for $N > 200$. For Scenario 3 we expect $n_0 \approx 1$; for $N = 201$, we get $n_0 \approx 883$, and we need around $N = 1500$, before we have $n_0 < 1.2$. This makes this first-order finite difference a bad choice, so we try a second-order finite difference method. For Scenario 3, we get convergence for the second-order approximation for $N > 44$. A plot of the results for Scenario 3 can be seen in Figure 2.5.

When we do get convergence, we get that $n(0) \approx 1.05$, 0.702, 1.05 for Scenario 1, 2 and 3 respectively. This is in accordance with what we expected.

Plots of the results

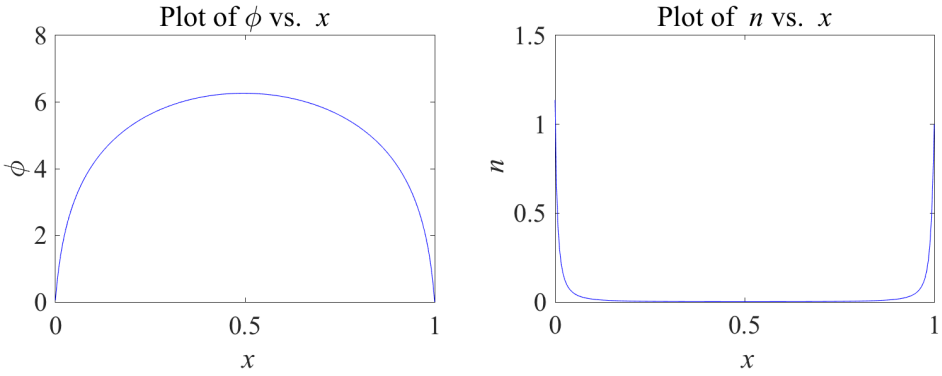


Figure 2.5: Plot of ϕ and n against x for Scenario 3. These are results when using the Newton-Raphson method in Section 2.2 with second-order finite differences.

Now that we have a method that converges for all three scenarios, we can look at different solutions for different values of ν. We can look at $\nu = 0.6, 0.8, 1, 1.2, 1.4$, and the continuous case when $0.6 \leq \nu \leq 1.4$. For Scenario 1 and 3, we get significant changes in the system when varying ν, but for Scenario 2, we see no differences in plots of the electrostatic potential and small differences for the electron density. The results for these cases are plotted in Figure 2.6, 2.7, 2.8 and 2.9. The plot for the continuous case for Scenario 2 is not included since it changes very little for the values of ν used.

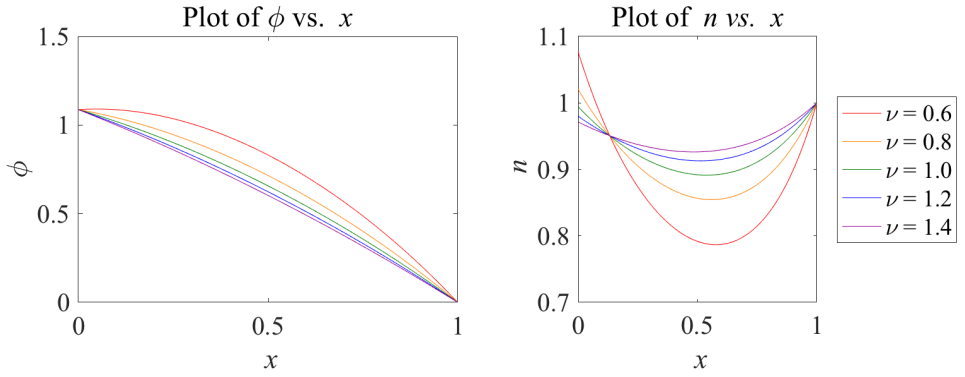


Figure 2.6: Plot of ϕ and n against x for different values of ν for Scenario 1, $\nu = 0.6, 0.8, 1, 1.2, 1.4$. These are results when using the Newton-Raphson method in Section 2.2 with second-order finite differences.

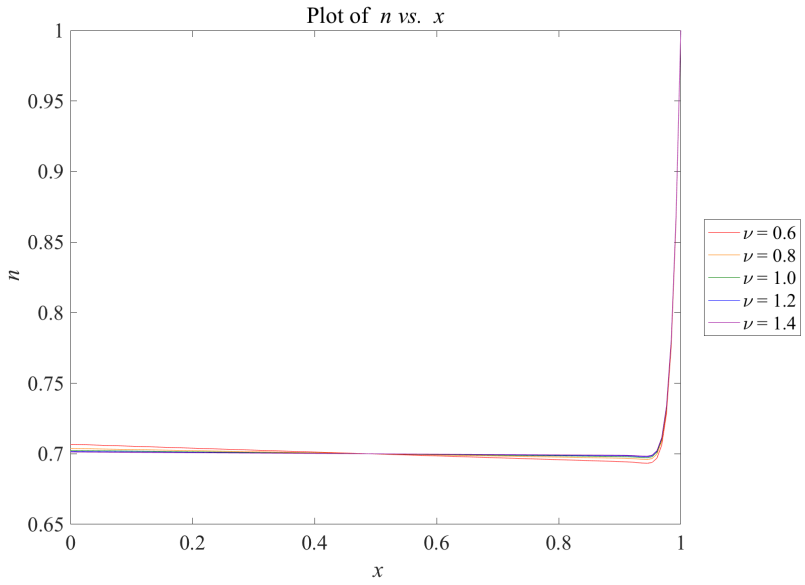


Figure 2.7: Plot of n against x for different values of ν for Scenario 2, $\nu = 0.6, 0.8, 1, 1.2, 1.4$. These values of ν for Scenario 2 change the values of n very little. The changes for ϕ are not visible in a plot and are not included. This result is when using the Newton-Raphson method in Section 2.2 with second-order finite differences.

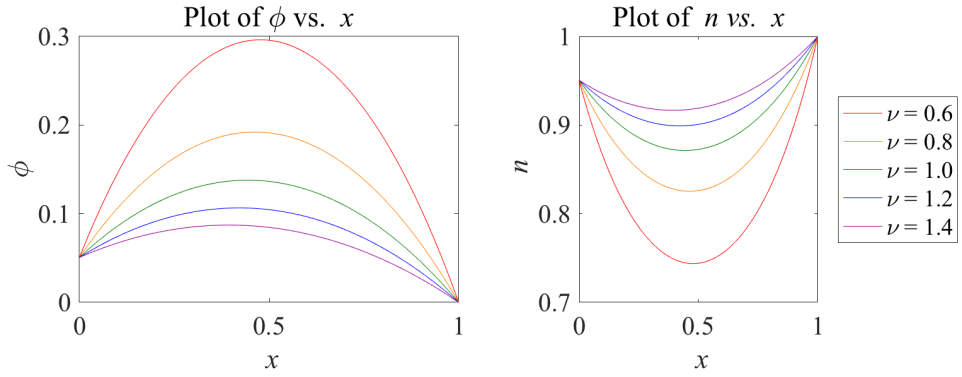
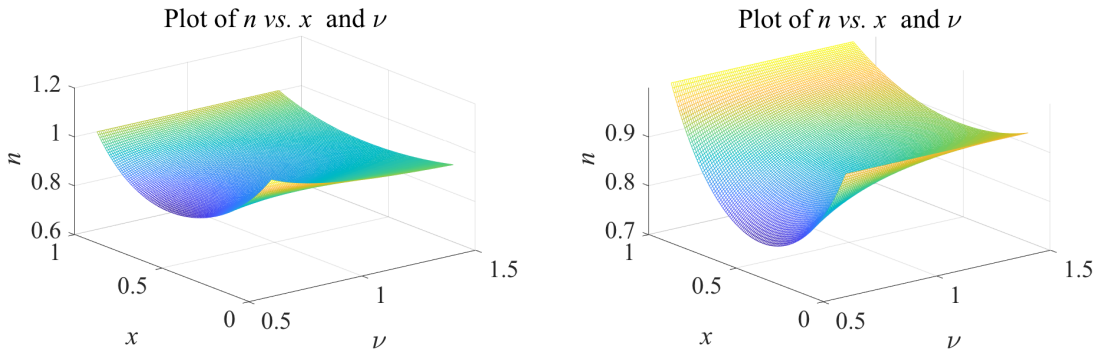


Figure 2.8: Plot of ϕ and n against x for different values of ν for Scenario 3, $\nu = 0.6, 0.8, 1, 1.2, 1.4$. These are results when using the Newton-Raphson method in Section 2.2 with second-order finite differences.



(a) Scenario 1 (b) Scenario 3

Figure 2.9: 3D plot of n as a function of x for different values of $\nu \in [0.6, 1.4]$. These are results when using the Newton-Raphson method in Section 2.2 with second-order finite differences.

We see that changing the values of ν changes the electron density and electrostatic potential distribution quite a lot for Scenario 1 and 2. For Scenario 2, however, there are small changes for the different selected values of ν. We see that for Scenario 1, different values of ν, changes the values of the left boundary for the electron density; higher values of ν yield lower values of $n(0)$. We also see that different values of ν changes the electron density for Scenario 3 quite significantly where lower values of ν yield lower minimum values for n. We also see that lower values of ν leads to a higher maximum for ϕ. This is in accordance with the Poisson's equation.

We also consider how the different values of Φ and J influence the left boundary value of n. We choose to look at a continuous case when $0.5 \leq \boldsymbol{\varphi} \leq 5$ and $0.2 \leq J \leq 1.6$ and $\nu = 1$. These values are chosen so that we see the dependence of the left boundary of n against the other parameters. The result from this consideration is plotted in Figure 2.10. As we see, high flux and low applied potential difference yield a high electron density on the left boundary, whereas, the opposite yield a low electron density on the left boundary. We also notice that the left boundary value of n is linear against J and convex against $\boldsymbol{\varphi}$. Also, decreasing Φ yields an increase in $n(0)$.

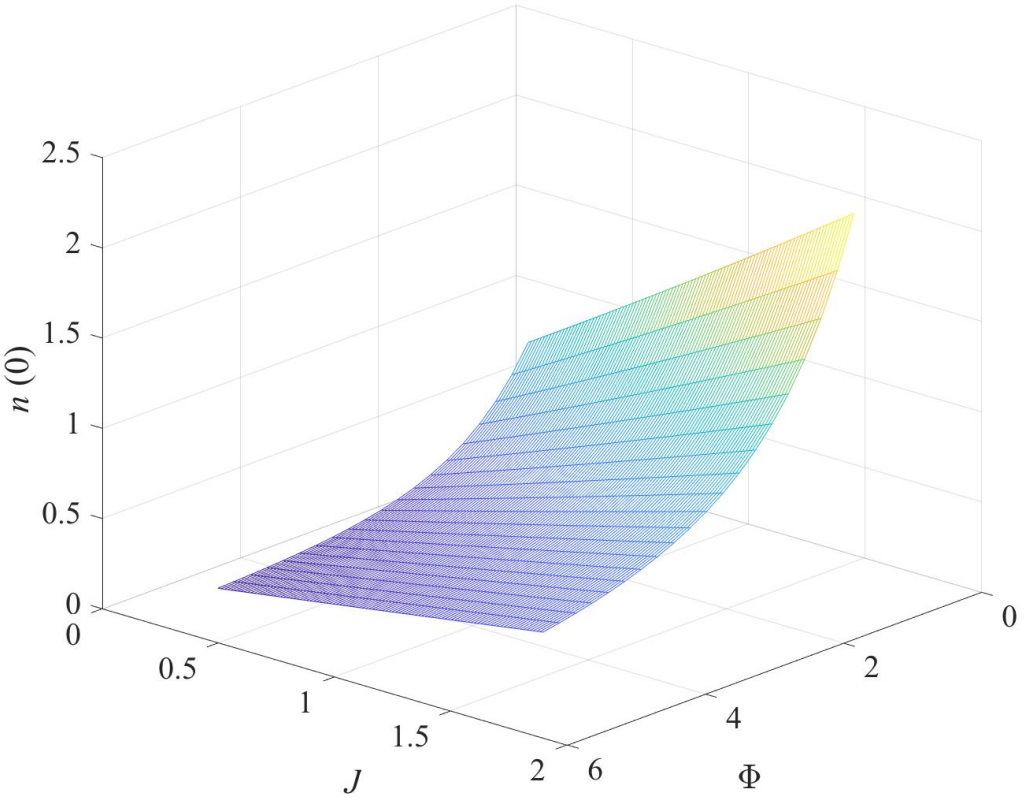


Figure 2.10: 3D plot of $n(0)$ *against* φ *and* J *for* $0.5 \leq \varphi \leq 5$ *and* $0.2 \leq J \leq 1.6$ *for* $\nu = 1$*. These are results when using the Newton-Raphson method in Section 2.2 with second-order finite differences. (In the plot, the variable* φ *is denoted* Φ *because I was not able to implement* φ *in MATLAB)*

Effective error of finite difference approximations

We now look at the effective order of the finite difference approximations we have used. We use the method as described in Section A.6, where we have that the discretization number is given by $N = 2^d - 1$. As mentioned before, the method for Scenario 3 converges only if we have $N > 200$, and is a physical solution for $N > 1000$ when we use first-order approximations. Therefore, in the table with the effective order of the finite difference approximation for the first-order approximation, the column for Scenario 3 is left blank except for $d = 12$. The effective orders for the first-order finite difference method are given in the table below.

Grid points d	**Order of error for ϕ** $\log_2\left(\frac{e_\phi^{[d-1]}}{e_\phi^{[d]}}\right)$			**Order of error for $\hat{n}$** $\log_2\left(\frac{e_n^{[d-1]}}{e_n^{[d]}}\right)$		
	Scenario			Scenario		
	1	2	3	1	2	3
7	0.8238	0.9540		0.9544	0.4810	
8	0.9172	0.9755		0.9763	0.6931	
9	0.9598	0.9872		0.9880	0.8193	
10	0.9802	0.9934		0.9939	0.8962	
11	0.9902	0.9966		0.9970	0.9432	
12	0.9951	0.9982	0.9879	0.9985	0.9701	1.1413

We see that that the effective order for $d = 7$ of n for Scenario 2, is lower than expected, but increases as d increases. For the second-order finite difference method, we do not get convergence for $d = 5$ for Scenario 3, so the space for $d = 7$ is left empty. The effective orders when using second-order finite differences are stated in the table below.

Grid points d	**Order of error for ϕ** $\log_2\left(\frac{e_\phi^{[d-1]}}{e_\phi^{[d]}}\right)$			**Order of error for $\hat{n}$** $\log_2\left(\frac{e_n^{[d-1]}}{e_n^{[d]}}\right)$		
	Scenario			Scenario		
	1	2	3	1	2	3
7	1.9798	1.8985		1.8904	0.7321	
8	1.9970	2.0117	3.0576	1.9448	1.0830	2.8606
9	2.0002	2.0885	2.4851	1.9722	1.3934	1.3466
10	2.0005	2.1047	2.3981	1.9860	1.6351	1.6443
11	2.0004	2.0807	2.3378	1.9930	1.7963	1.7446
12	2.0002	2.0500	2.2504	1.9965	1.8918	1.8431

For this second-order scheme we have used a first-order finite difference near the right boundary of n. For Scenario 2, we notice that for we do not get an effective order of 2, unless the discretization number is close to $N = 10^4$. We even get an effective order lower than 1 for $d = 7$. As we see in Figure 2.11, the results from the first and second-order finite difference method differ for Scenario 2, in particular there is a significant difference in the results for where the values of n make the transition from $n \approx 0.7$ to increase rapidly before reaching $n = 1$. Also, one of the reasons why we get an effective order of n for the second-order finite difference method for Scenario 2 that is lower than expected, may be that we have used a first-order finite difference method at x_N.

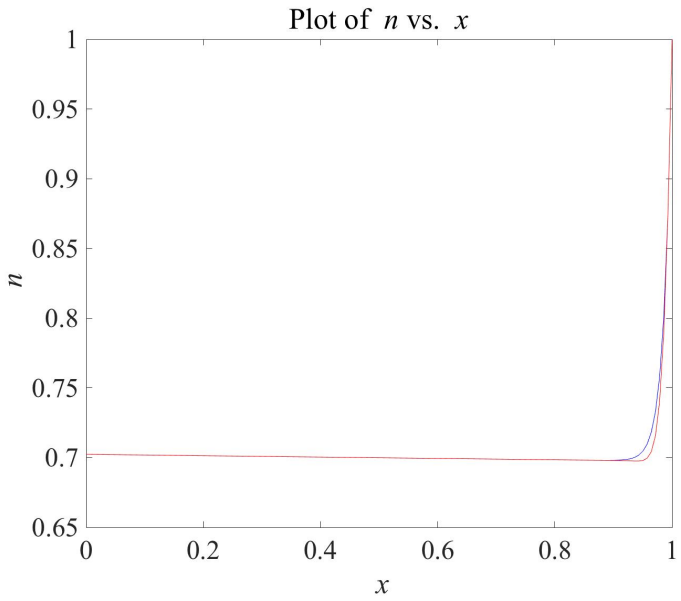


Figure 2.11: Plot of n against x using first (blue) and second-order (red) finite difference methods for Scenario 2 with N=150. These are results when using the Newton-Raphson method in Section 2.2.

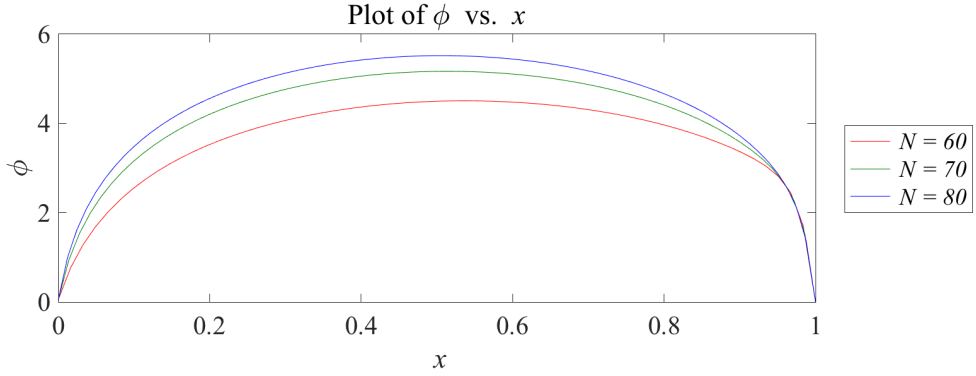


Figure 2.12: Plot of ϕ against x for different values of discretization numbers for Scenario 3. For $N \geq 80$, there are no visible differences in a plot of solutions. These are results when using the Newton-Raphson method in Section 2.2 with second-order finite differences.

For Scenario 3 the solution when using a second-order finite difference method changes quite significantly for discretization numbers from 60 to 80. In the calculation of the effective order when using second-order finite differences for Scenario 3, we use $d = 6 \iff N = 63$, $d = 7$ and $d = 8$. As we see in Figure 2.12, the value $d = 6$, will give a solution that will be very different from the solution for $d = 7$. Therefore, the high effective orders for $d = 8$ for Scenario 3 are not surprising.

We notice that the effect that the different scenarios have on the different systems change the effective order of error. We are therefore in Scenario 3 dependent on using many grid points. Also, for Scenario 2, where n changes rapidly, it would be advantageous to have many grid points. If we were considering a system with more dimensions, a different approach should have been made. For example to account for the big changes in the values of n near the right boundary in Scenario 2. Such an approach could be either a non-equidistant grid spacing or using a finite element method.

For Scenario 3 with $N = 60, 70, 80$, it is hard to see the differences in n; for plots of solutions for Scenario for $N = 60$, we have $n(0) \approx 0.19$, for $N = 70$ we have $n(0) \approx 0.58$ and for $N = 80$, we get $n(0) \approx 0.99$. The solutions all look quite similar except at the left boundary. We therefore only plot the results we get for ϕ with these different discretization numbers.

Runtime and number of iterations for convergence

We here look at the iteration scheme in Section 2.1 vs. the Newton-Raphson method in Section 2.2. We will consider the differences in runtime and difference in number of iterations for convergence. We will of course only consider Scenario 1 and 2, as Scenario 3 does not yield a convergent method for the iteration scheme. We consider the second-order finite difference method for both iteration methods. The results for Scenario 1 and 2 can be seen in the table below. The results for Scenario 2 when using Newton-Raphson for $N = 10^4$ are not included, as these take a long time to calculate.

We have that for the consistent runtimes (where MATLAB uses approximately the same time on one iteration each time I run the code), the Newton-Raphson

method uses around a factor of 10 more seconds to run. We have that the iteration scheme is two iterations, each with a matrix of dimension $\sim N \times N$, and that the Newton-Raphson method has a matrix of dimension $\sim 2N \times 2N$. If the system is solved used Gaussian elimination, it should take $\sim 2N^3$ operations for the iteration scheme and $\sim 8N^3$ operations for the Newton-Raphson to to solve the linear system. However, we mostly have diagonal, bidiagonal and tridiagonal matrices for the iteration scheme, that each take on the $\sim N$ operations to solve. For the Newton-Raphson method, however, we have a $\sim 2N \times 2N$ matrix with elements in the lower left block and upper right block. These elements make the matrix much less sparse and makes the system much more involved to solve. This explains the significant differences in runtime for the two iteration methods.

For the number of iterations for the iteration scheme for $N = 5 \times 10^3$, the number of iterations until convergence is 14, in contrast to 12 for the other discretization numbers. I do not know why we get this result. We need fewer iterations for the Newton-Raphson method to convergence, especially for Scenario 1, where we need 3 times fewer iteration steps for convergence. This is expected as the iteration scheme generally takes more number of iterations to converge compared to the Newton-Raphson method. Also, since the iteration scheme does not converge at all for Scenario 3, we cannot be certain of how well the scheme converges.

Degree of freedom	**Runtime for one iteration**		**Number of iterations to converge**			
	Method		Method			
	Iteration s.	Newton-R.	Iteration s.		Newton-R.	
			Scenario		Scenario	
			1	2	1	2
1×10^2	$\sim 10^{-3}$	$\sim 10^{-3}$	12	6	4	4
5×10^2	$\sim 10^{-3}$	$\sim 10^{-2}$	12	7	4	5
1×10^3	$\sim 10^{-2}$	$\sim 2 \times 10^{-1}$	12	7	4	5
5×10^3	$\sim 8 \times 10^{-1}$	8.4×10^0	14	7	4	5
1×10^4	4.5×10^0	5.7×10^1	12	7	4	

We now want to look at the case where J is not a given value, but rather a value that either is included in the iteration or a value that is calculated after the iteration has converged. We first consider when it is included in the iteration.

2.3 Including the flux in the iteration

If we do not know the value of J beforehand, we can calculate it. We want to find an expression for J. We rewrite (2.4), and get

$$J = -n\phi' - n' \tag{2.42}$$
$$= -n\left(\phi' + \frac{1}{n}n'\right) \tag{2.43}$$
$$= -n\left(\phi + \log(n)\right) \tag{2.44}$$
$$\iff$$
$$\frac{J}{n} = -\left(\phi + \log(n)\right)'. \tag{2.45}$$

We can now integrate from 0 to 1, using that $J' = 0$, $\phi(0) = \boldsymbol{\varphi}$, $\phi(1) = 0$ and $n(1) = 1$:

$$\int_0^1 \frac{J}{n}dx = -\int_0^1 \left(\phi + \log(n)\right)' dx \tag{2.46}$$
$$J\int_0^1 \frac{1}{n}dx = -(\phi + \log(n))\Big|_0^1 \tag{2.47}$$
$$= -\left\{0 + \log[n(1)] - (\boldsymbol{\varphi} + \log[n(0)])\right\} \tag{2.48}$$
$$= \boldsymbol{\varphi} + \log(n(0)). \tag{2.49}$$

We can now write J as

$$J = \frac{\boldsymbol{\varphi} + \log[n(0)]}{\int_0^1 \frac{1}{n}dx}. \tag{2.50}$$

This means that if $J = 0$, then $\boldsymbol{\varphi} = -\log[n(0)]$, which means that $-\log[n(0)]$ is the built-in chemical potential. We can now find the resistance, R, using Ohm's law:

$$R(n) := \frac{\boldsymbol{\varphi} - (-\log[n(0)])}{J} = \frac{\boldsymbol{\varphi} + \log[n(0)]}{J}. \tag{2.51}$$

Using the Trapezoidal Rule, we can find an expression for J. By Theorem A.9 we get

$$J = (\boldsymbol{\varphi} + \log(n_0)) \left[\Delta\, x \left\{ \frac{1}{2} \left(\frac{1}{n_0} + \frac{1}{n_{N+1}} \right) + \mathbf{1}^T \boldsymbol{n}^{\circ(-1)} \right\} \right]^{-1}. \tag{2.52}$$

Now, including the flux, J, in the Newton-Raphson system, we get the following function, that we want to solve for $\boldsymbol{f}_{2.3}(\boldsymbol{z}) = \mathbf{0}_{2N+2}$:

$$\boldsymbol{f}_{2.3}(\boldsymbol{z}) = \underbrace{\left(\begin{array}{c|c|c} S_0 & \mathbf{0}_N \quad \left(\frac{\Delta\, x}{\nu}\right)^2 I_N & \mathbf{0}_N \\ 0 & \operatorname{diag}(\hat{F_+}\, \hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+} & \Delta\, x \mathbf{1}_{N+1} \\ \mathbf{0}_N^T & \mathbf{0}_{N+1}^T & 1 \end{array}\right)}_{=:\, G_{2.3}(\boldsymbol{z})} \left(\begin{array}{c} \boldsymbol{\phi} \\ \hat{\boldsymbol{n}} \\ J \end{array}\right) + \left(\begin{array}{c} \boldsymbol{r}_{\phi,S_0} \\ \boldsymbol{r}_{n,\hat{F_+}} \\ -\omega(\hat{\boldsymbol{n}}) \end{array}\right) \tag{2.53}$$

where $\omega(\hat{\boldsymbol{n}}) := (\boldsymbol{\varphi} + \log(n_0)) \left[\Delta\, x \left\{ \frac{1}{2} \left(\frac{1}{n_0} + \frac{1}{n_{N+1}} \right) + \mathbf{1}_N^T \boldsymbol{n}^{\circ(-1)} \right\} \right]^{-1}$.

We now have

$$\boldsymbol{z} = \left(\begin{array}{c} z_1 \\ \vdots \\ z_{2N+2} \end{array}\right) := \left(\begin{array}{c} \boldsymbol{\phi} \\ \hat{\boldsymbol{n}} \\ J \end{array}\right) \iff J = z_{2N+2}.$$

Here $G_{2.3}(\boldsymbol{z})$ is a 3×3 block matrix. We can let $\{G_{2.3}\}_{1,1}$ be the upper left block and so on, as before. Now, the last element of $\boldsymbol{f}_{2.3}(\boldsymbol{z})$ will be $\{G_{2.3}\}_{3,3} J - \omega(\hat{\boldsymbol{n}})$.

We now find the Jacobian of $\boldsymbol{f}_{2.3}(\boldsymbol{z})$ by extending (2.38) with an extra row and column. The Jacobian of the extra column in the second block row is found with Claim B.1. The last row of the Jacobian is found using Claim B.1 and Corollary

B.11. We get the following Jacobian of $\boldsymbol{f}_{2.3}(\boldsymbol{z})$:

$$\mathcal{J}_{\boldsymbol{z}}^{\boldsymbol{f}_{2.3}(\boldsymbol{z})}(\boldsymbol{z}) = \left(\begin{array}{c|c|c} S_0 & \mathbf{0}_N \quad \left(\frac{\Delta x}{\nu}\right)^2 I_N & \mathbf{0}_N \\ \mathrm{diag}(\hat{\boldsymbol{n}})[\hat{F_+}]_{:,2:N+1} & \mathrm{diag}(\hat{F_+}\,\hat{\boldsymbol{\phi}} + \boldsymbol{r}_{\phi,\hat{F_+}}) + \hat{F_+} & \Delta x \mathbf{1}_{N+1} \\ \mathbf{0}_N^T & -\hat{\boldsymbol{\sigma}}(\hat{\boldsymbol{n}}) & 1 \end{array}\right) \tag{2.54}$$

where

$$\hat{\boldsymbol{\sigma}}(\hat{\boldsymbol{n}}) := \left(\begin{array}{cc} \frac{\frac{1}{2}\left(\frac{1}{n_0}[1+\boldsymbol{\varphi}+\log(n_0)]+\frac{1}{n_{N+1}}\right)+\mathbf{1}_N^T \boldsymbol{n}^{\circ(-1)}}{\Delta x\, n_0\left[\frac{1}{2}\left(\frac{1}{n_0}+\frac{1}{n_{N+1}}\right)+\mathbf{1}_N^T \boldsymbol{n}^{\circ(-1)}\right]^2} & \frac{\boldsymbol{\varphi}+\log(n_0)}{\Delta x\left[\frac{1}{2}\left(\frac{1}{n_0}+\frac{1}{n_{N+1}}\right)+\mathbf{1}_N^T \boldsymbol{n}^{\circ(-1)}\right]^2}\left(\boldsymbol{n}^{\circ(-2)}\right)^T \end{array}\right). \tag{2.55}$$

We can then solve the system as described in (2.39).

Results from Newton-Raphson including the flux

The method "converges" for Scenario 1 and 2, but nor for Scenario 3. With "convergence", we mean that the solutions all tend to one specific value, but we never reach an iteration where the relative error is less than 10^{-14}. It is the error of $\hat{\boldsymbol{n}}$ that does not decrease after a few iterations, and stays at around 10^{-10}, even for the 200th iteration. Also, we get that the relative error increases and decreases in a quite random matter, so that we do not know if next iteration will give us a more accurate solution than the current solution. Since we do not get convergence in the normal sense, we cannot look at the effective order of error for the finite difference approximations. The results for Scenario 1 and 2 are plotted in Figure 2.13 and 2.14.

For Scenario 1, we get a value of J that tends towards $1.71 \approx 1.7\ J_{in}$, where $J_{in} = 1$ is the initial guess of J. This value is reached after around 20 iterations. For Scenario 2, the value of J tends towards $14.8 \approx 0.21\ J_{in}$ when $J_{in} = 70$. This value is reached after around 5 iterations. A plot of the values of J in each iteration for Scenario 1 and 2 can be seen in Figure 2.15. For these plots, the flux, J, fluctuates quite a lot in each iteration, especially in Scenario 1 where we for several iterations have negative flux. This is non-physical when $\boldsymbol{\varphi} > 0$, which we indeed have in Scenario 1.

The results we get are different from what we get when do not change J. For

example, for Scenario 1 we now get that $n(0) > 1.5$, which probably is not correct as the lower boundary is silver, and cannot exceed the silver electron concentration on the top by such an order. Also, we get for Scenario 2 that $n(0) < 0.2$ does indeed fulfill $n(0) < 1$. The feasibility of this result will depend on how doped the n-type layer is.

Plots of the results

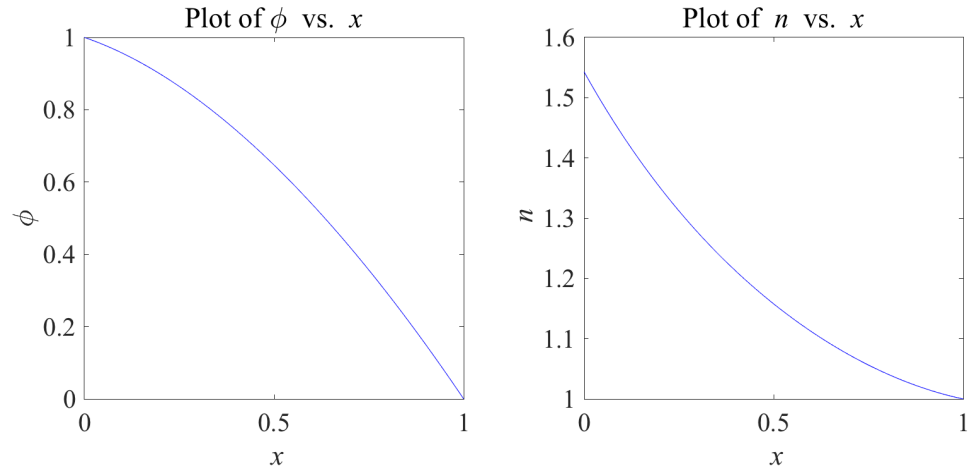


Figure 2.13: Plot of ϕ and n against x for Scenario 1. These are results when using the Newton-Raphson method in Section 2.3 with second-order finite differences.

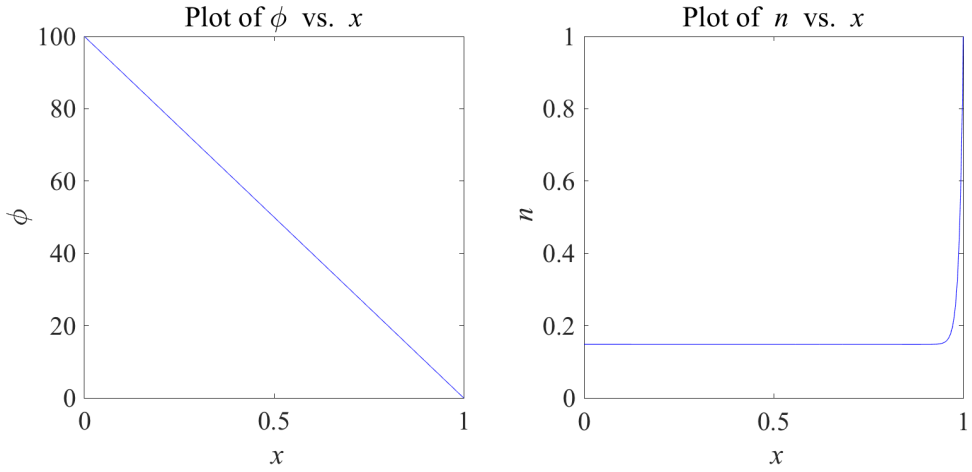
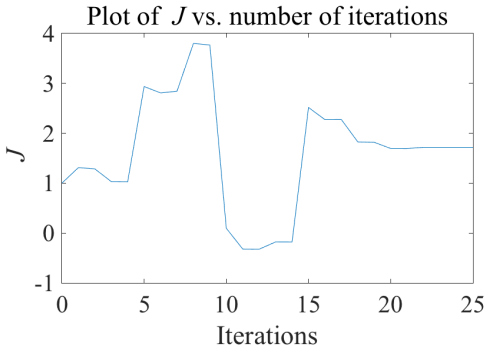
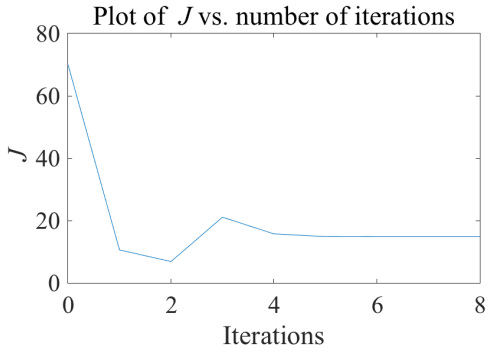


Figure 2.14: Plot of ϕ and n against x for Scenario 2. These are results when using the Newton-Raphson method in Section 2.3 with second-order finite differences.



(a) In Scenario 1, J tends towards 1.71 after 19 iterations.



(b) In Scenario 2, J tends towards 14.8 after 4 iterations.

Figure 2.15: Plot of the value of J vs. the number of iterations. These are results when using the Newton-Raphson method in Section 2.3 with second-order finite differences.

2.4 Differentiating the drift-diffusion equation

We have so far considered the drift-diffusion equation where we need J as a given fixed value, or as an initial guess. We can differentiate Equation (2.4) to arrive at an equation without J, so that we can compute J in a post-process using (2.52). We then need the left boundary for the electron concentration, α, as given and fixed. By differentiating (2.4) and using that $J' = 0$, we get

$$n'' + \phi' n' - \frac{n^2}{\nu^2} = 0 \tag{2.56}$$

where we have substituted the expression for ϕ'' from (2.4) into the equation. As before, we discretize (2.56) and write it in matrix form using Lemma A.23. We then get

$$\left(4S_0 + \operatorname{diag}\left(F_0\boldsymbol{\phi} + \boldsymbol{r}_{\phi,F_0}\right) F_0 - 4\left(\frac{\Delta x}{\nu}\right)^2 \operatorname{diag}(\boldsymbol{n})\right)\boldsymbol{n} \tag{2.57}$$

$$+\operatorname{diag}\left(F_0\boldsymbol{\phi} + \boldsymbol{r}_{\phi,F_0}\right)\boldsymbol{r}_{n,F_0} + 4\boldsymbol{r}_{n,S_0} = 0. \tag{2.58}$$

We now have $\boldsymbol{f}_{2.4}(\boldsymbol{z})$, that we want to solve for $\boldsymbol{f}_{2.4}(\boldsymbol{z}) = \boldsymbol{0}_{2N}$, given by

$$\boldsymbol{f}_{2.4}(\boldsymbol{z}) = \underbrace{\left(\begin{array}{c|c} S_0 & \left(\frac{\Delta x}{\nu}\right)^2 I_N \\ 0 & 4S_0 + \operatorname{diag}\left(F_0\boldsymbol{\phi} + \boldsymbol{r}_{\phi,F_0}\right) F_0 - 4\left(\frac{\Delta x}{\nu}\right)^2 \operatorname{diag}(\boldsymbol{n}) \end{array}\right)}_{=: G_{2.4}(\boldsymbol{z})} \left(\begin{array}{c} \boldsymbol{\phi} \\ \boldsymbol{n} \end{array}\right) \tag{2.59}$$

$$+ \left(\begin{array}{c} \boldsymbol{r}_{\phi,S_0} \\ \operatorname{diag}\left(F_0\boldsymbol{\phi} + \boldsymbol{r}_{\phi,F_0}\right) \boldsymbol{r}_{n,F_0} + 4\,\boldsymbol{r}_{n,S_0} \end{array}\right) \tag{2.60}$$

with

$$\boldsymbol{z} = \left(\begin{array}{c} z_1 \\ \vdots \\ z_{2N} \end{array}\right) := \left(\begin{array}{c} \boldsymbol{\phi} \\ \boldsymbol{n} \end{array}\right) \iff \boldsymbol{\phi} = \left(\begin{array}{c} z_1 \\ \vdots \\ z_N \end{array}\right), \boldsymbol{n} = \left(\begin{array}{c} z_{N+1} \\ \vdots \\ z_{2N} \end{array}\right).$$

We now find the Jacobian of $\boldsymbol{f}_{2.4}(\boldsymbol{z})$ as before by using Claim B.6. The last N rows of the Jacobian can be found by using Claim B.2 and Corollary B.5. We get the following Jacobian of $\boldsymbol{f}_{2.4}(\boldsymbol{z})$:

$$\left[\mathcal{J}_{\boldsymbol{z}}^{\boldsymbol{f}_{2.4}(\boldsymbol{z})}(\boldsymbol{z})\right]_{:,1} = \left(\begin{array}{c} S_0 \\ \left(\operatorname{diag}(F_0\boldsymbol{n}) + \operatorname{diag}(\boldsymbol{r}_{n,S_0})\right) F_0 \end{array}\right) \tag{2.61}$$

$$\left[\mathcal{J}_{\boldsymbol{z}}^{\boldsymbol{f}_{2.4}(\boldsymbol{z})}(\boldsymbol{z})\right]_{:,2} = \left(\begin{array}{c} \left(\frac{\Delta x}{\nu}\right)^2 I_N \\ 4S_0 + \operatorname{diag}\left(F_0\boldsymbol{\phi} + \boldsymbol{r}_{\phi,F_0}\right) F_0 - 8\left(\frac{\Delta x}{\nu}\right)^2 \operatorname{diag}(\boldsymbol{n}) \end{array}\right) \tag{2.62}$$

We can then solve the system as described in (2.39) and then find J using (2.52).

Results from differentiated drift-diffusion equation

When differentiating we may lose information about the system. We introduce this information by adding the left boundary condition for the electron density. For this method, we get convergence for Scenario 1 and 2 and for Scenario 3 only when $N > 86$. However, we do not see a tendency like before where an increase in discretization points above the convergence threshold gives a more accurate result, rather, the results appear identical in a plot for $N > 87$. The results for Scenario 1, 2 and 3 are plotted in Figure 2.16, 2.17 and 2.18.

We get the following values of J for Scenario 1, 2 and 3, respectively: 0.9272, 69.75 and 1.82×10^{-4} when $N = 100$. Since Black et al. find the applied potential difference to be $\boldsymbol{\varphi} = 1.087$ instead of the flux, J, we can see if using $\boldsymbol{\varphi} = 1.087$ instead of $\boldsymbol{\varphi} = 1$ for Scenario 1, and $\boldsymbol{\varphi} = 0.0507$ instead of $\boldsymbol{\varphi} = 0.05$ for Scenario 3, gives values of J that are more in correspondence with what Black et al. find. For Scenario 1, when using $\boldsymbol{\varphi} = 1.087$, we get $J \approx 1.008$, which is more in correspondence for the fixed value, $J = 1$, that is used in Section 2.1 and 2.2. For Scenario 3, when using $\boldsymbol{\varphi} = 0.0507$, we get $J \approx 1.85 \times 10^{-4}$. This is not as close to the fixed value that we used for J in Section 2.1 and 2.2: $J = 2 \times 10^{-4}$. That means that value of J for Scenario 3 changes less when we change $\boldsymbol{\varphi}$ with an relative order 8×10^{-2} as compared to Scenario 1.

Plots of results

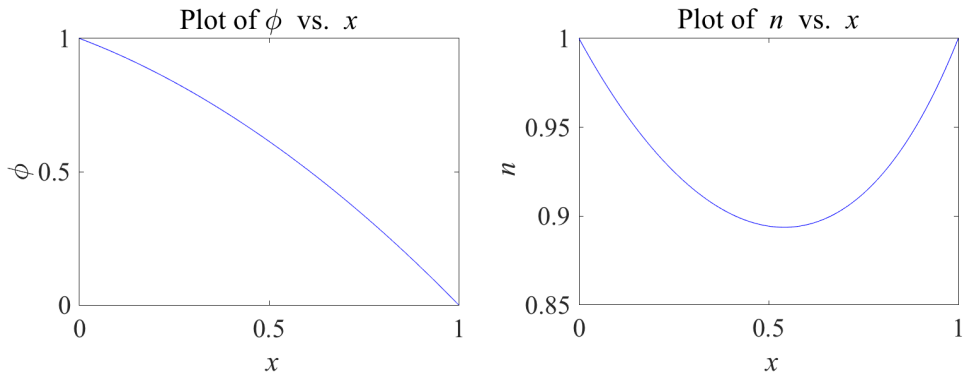


Figure 2.16: Plot of ϕ and n against x for Scenario 1. These are results when using the Newton-Raphson method in Section 2.4 with second-order finite differences. These results are in accordance with [2].

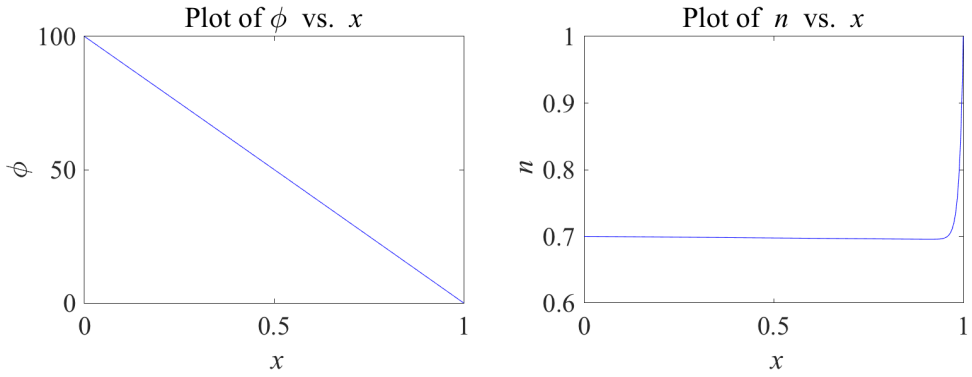


Figure 2.17: Plot of ϕ and n against x for Scenario 2. These are results when using the Newton-Raphson method in Section 2.4 with second-order finite differences. These results are in accordance with [2].

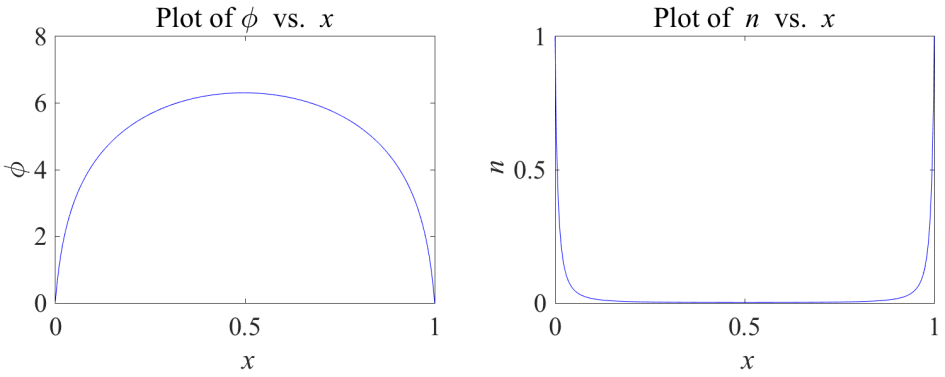
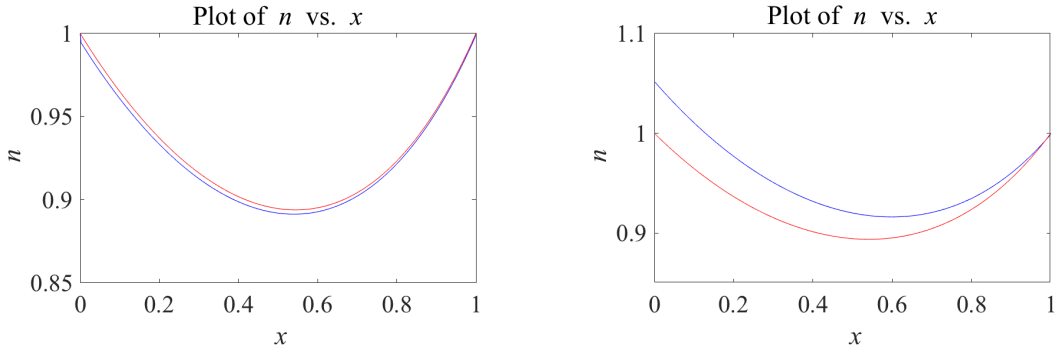


Figure 2.18: Plot of ϕ and n against x for Scenario 3. These are results when using the Newton-Raphson method in Section 2.4 with second-order finite differences. These results are in accordance with [2].

Comparison of results using the non-differentiated vs. the differentiated drift-diffusion equation

For all three scenarios, we get that the values of the electrostatic potential, ϕ, are indistinguishable in a plot, whether we use the non-differentiated or the differentiated drift-diffusion equation. The differences in the values of the electron density, n, for Scenario 1 and 2 are plotted in Figure 2.19 and 2.20. For Scenario 1, we also consider the difference in n when we use $\varphi = 1$, which we have used, and when we use $\varphi = 1.087$, which is what Black et al. finds when solving the system for φ.

For Scenario 1, there are some differences, naturally more when we use $\varphi = 1$ than when we use $\varphi = 1.087$ which is the accurate result for φ when we have specified the left boundary value. The values for n for Scenario 3 are also indistinguishable in a plot and are thus not included. The difference in values of n when using the non-differentiated vs. the differentiated equation are for Scenario 2 quite small.



(a) Scenario 1 when using $\varphi = 1.087$.

(b) Scenario 1 when using $\varphi = 1$.

Figure 2.19: Comparison of the solution of ϕ and n against x for non-differentiated (blue) against differentiated (red) drift-diffusion equation for Scenario 1. The non-differentiated equation has been solved with Newton-Raphson as described in Section 2.2, and the differentiated equation has been solved using the Newton-Raphson method as described in Section 2.4. These are results when second-order finite differences.

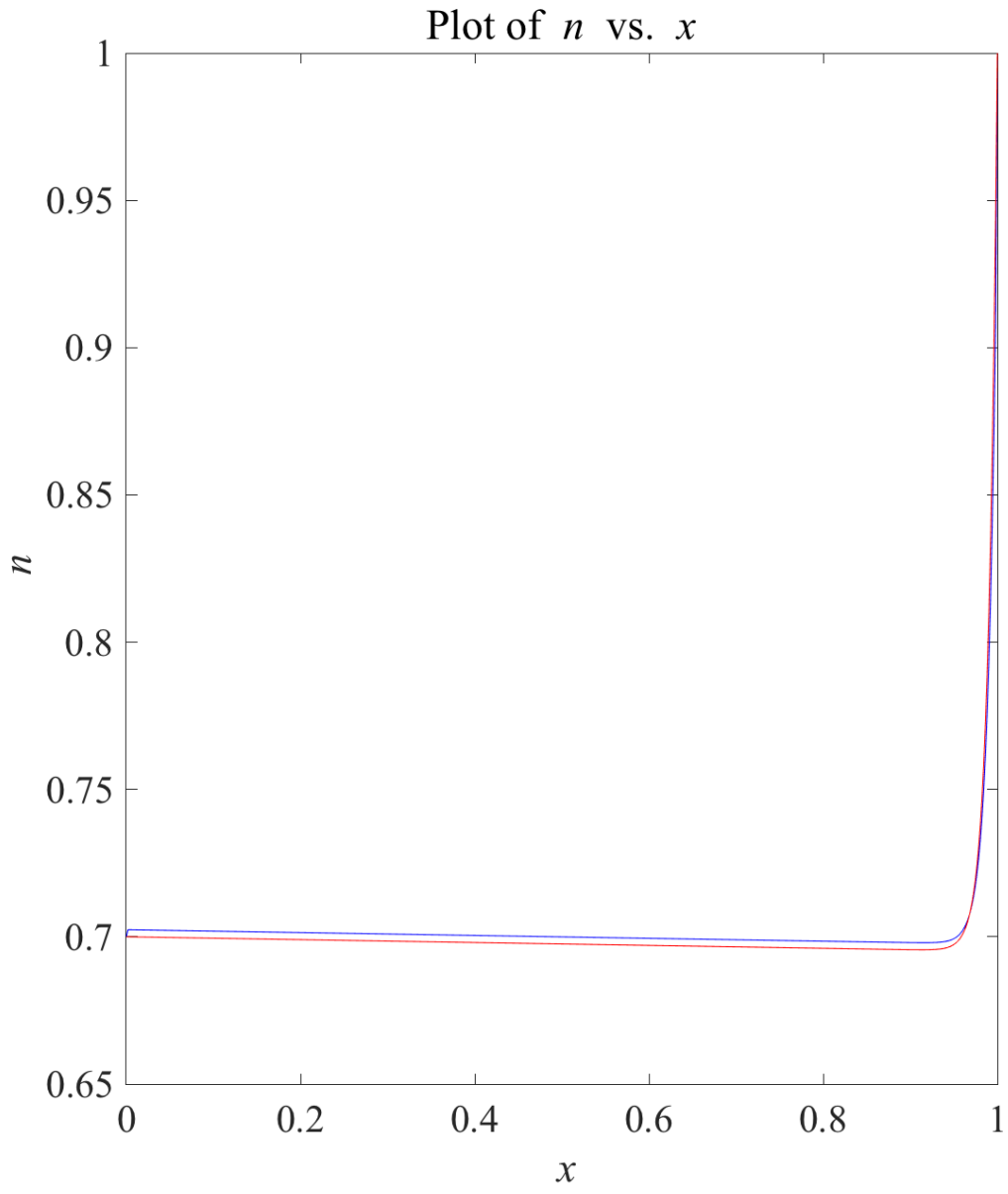


Figure 2.20: Comparison of the solution of ϕ and n against x for non-differentiated (blue) against differentiated (red) drift-diffusion equation for Scenario 2. The non-differentiated equation has been solved with Newton-Raphson as described in Section 2.2, and the differentiated equation has been solved using the Newton-Raphson method as described in Section 2.4. These are results when using second-order finite differences.

2.5 Results for silicon solar cell - contact resistance

We have found a system that takes the applied potential difference and electron flux as given values and then finds the left boundary value for the electron density. We have also considered the electron flux as a variable that can change in the iteration method. Lastly, we have considered where we know all boundary values and calculate the flux in a post-process.

For solving our system, we have considered an iteration scheme and different setups for the Newton-Raphson method. For the iteration scheme, we have found bounds that guarantee convergence with supplementing strict bounds that are found experimentally.

Using the Newton-Raphson method, however, gives better chances for convergence when the ratio between the Debye length and the glass thickness is small. We find that for Scenario 3, the number of discretization points around $N = 70$ makes a big impact on the outcome of the values. It is shown that the value of ν impacts the behavior of the system, especially for Scenario 1 and 3. For Scenario 1, the value of ν changes the left boundary value for n, but not for Scenario 2 and 3. The finite difference approximations that we use do indeed give the expected result for large enough discretization numbers. The Newton-Raphson method also converges in significantly fewer steps than the iteration scheme, but each iteration takes longer time, especially for high discretization numbers. We also find that there are differences between using first and second-order finite differences, but that for large enough discretization numbers, we get that the differences become minuscule.

The system changes behavior when we introduce J as a variable that can change in every iteration. By doing so, we do not get convergence in the sense that the relative error reaches the order of (combined) rounding error (10^{-14}), but we get close to 10^{-10}. For Scenario 3, we get values that diverge. For Scenario 1 and 2, however, we get a system that stabilizes and with values of J that approach a limit. When introducing J, we get different values for the left boundary of n, for both Scenario 1 and 2.

When considering the left boundary of n as given, we have arrived a system where we have differentiated the drift-diffusion equation to make J vanish. We then make a method that can calculate J after the system has converged. We see that this system is advantageous when we know the left boundary value of the electron density, which is realistic when the materials on both boundaries are the same. For Scenario 1 and 2, we get some differences in the solutions when using the differentiated drift-diffusion, as compared to the non-differentiated drift-diffusion. For Scenario 3, we get identical results, in a plot, with this system, as we get with the non-differentiated system.

The results we have found can be used to model the resistance through the glass layer we have considered. For Scenario 1, we get a resistance of $R = 1.078$ when calculating with the method described in Section 2.2, and we get $R = 1.079$ when we use the method from Section 2.4. We see that these two methods give is a resistance that is relatively close. For Scenario 2, however, we get a resistance of $R = 1.424$ when calculating with the method described in Section 2.2, and we get $R = 1.429$ when we use the method from Section 2.4. We see that the method of choice for Scenario 2 is less important for calculating the resistance than it is for Scenario 1. Lastly, for Scenario 3, we get a resistance of $R = 494.0$ when calculating with the method described in Section 2.2, and we get $R = 274.7$ when we use the method from Section 2.4. This means that the method we choose when calculating the resistance for Scenario 3 has a major impact. We also see that it is the setup in Scenario 1 that gives the least resistance, and the setup in Scenario 2 gives a resistance that is around 30% higher. For Scenario 3, however, we get a very high resistance. This means that the thickness of the glass plays a major role for the resistance through it and that thicker glass leads to higher resistance. Physically, this is reasonable.

Chapter 3

Solving equations for organic solar cells - DA-interface

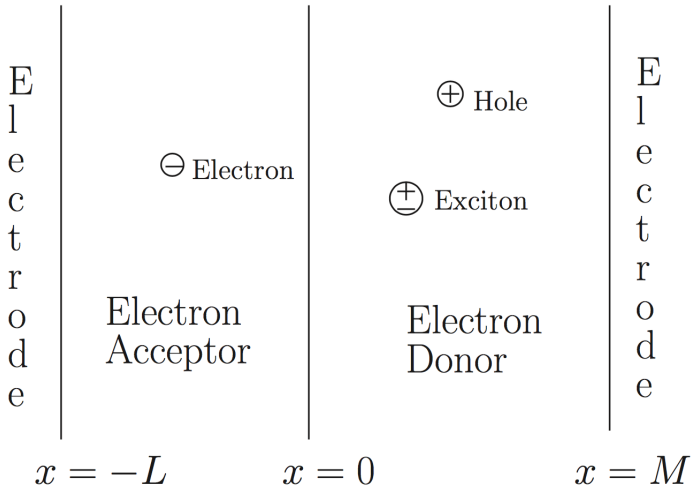


Figure 3.1: Illustration of the domain for which we solve equations. Richardson et al. are considering the acceptor (left) and donor (right) as two domains. Courtesy of [14].

In this chapter, we look at [14], "Asymptotic solution of a model for bilayer organic diodes and solar cells", by Richardson et al. A typical problem for organic solar cells is that the e-h pair can only be separated at the DA-interface as described in Subsection 1.3.2. However, the e-h pairs may only diffuse around 10 nm, while the thickness of an organic solar cell "must be around 200 nm thick in

order to absorb a large proportion of the solar spectrum". To understand how the electron density, flux and electrostatic potential behaves in an organic solar cell with a complex shape, we first need to understand how it behaves when the cell has a simple shape, so they consider a bilayer device where there are "two layers of material so that the device is planar and the interface is flat".

Formulation of the problem

We want to look at the cross-section of the panel. This cross-section is sketched in Figure 3.1. We consider the acceptor and donor layer. Since a cell is typically much bigger in order of magnitude than the thickness of each layer, we can ignore the effect on the edges of the cell and reduce the problem to a one-dimensional spatial problem where the space variable is the horizontal axis of the layers in Figure 3.1. We denote the acceptor layer thickness L and the donor layer thickness M. We let the applied electric potential difference, Φ, go from the left to the right, that is from $x = -L$ to $x = M$, and be 0 at $x = 0$, and of equal magnitude at $x = -L$ and $x = M$. We also include the built-in potential, Φ_{bi}, named chemical potential in Chapter 2. This means that the electrostatic potential will be different from what we had in Chapter 2, where we only used the applied potential difference for the electrostatic potential. So we have a total potential difference of $\Phi + \Phi_{bi}$, with the electrostatic potential taking the value $(\Phi + \Phi_{bi})/2$ at $x = -L$, and $-(\Phi + \Phi_{bi})/2$ at $x = M$. For this model, we consider both positive and negative values for the applied potential difference. So for $\Phi/|\Phi| = \Phi_{bi}/|\Phi_{bi}|$, we have forward bias, and for $\Phi/|\Phi| = -\Phi_{bi}/|\Phi_{bi}|$, we have reverse bias. We expect this transition to lead to a change in direction for the electron/hole flux, that is, the flux value will change sign.

Governing equations

Again, it is the drift-diffusion equation and Poisson's equation that are the governing equation here, as they were in Chapter 2. However, in this case, we let the electron and hole mobility be replaced by the electron affinity, $-\mu_n$, and the ionization potential, μ_p, and these variables are now dependent on space. We now have that the Poisson's equation is dependent on both the electron and hole density. We define the *pseudo band gap* as $E_g := \mu_n - \mu_p$, as the difference between the electron affinity and the ionization potential, instead of the energy difference between the conduction and valence band, like we have for inorganic semiconductors. The governing equations are the Poisson's equation for electrostatics given

by (1.3), and the drift-diffusion equation, given by (1.15) and (1.16), but where the drift-diffusion equations are scaled with the thermal voltage, k_BT/q.

$$\phi''(x) = \frac{q|n(x) - p(x)|}{\epsilon} \tag{3.1}$$

$$J_n(x) = -D_n \left(n'(x) + \frac{1}{k_BT} \left(\mu_n(x) - q\phi(x) \right)' \right) \tag{3.2}$$

$$J_p(x) = -D_p \left(p'(x) - \frac{1}{k_BT} \left(\mu_p(x) - q\phi(x) \right)' \right) \tag{3.3}$$

on $-L < x < M$ where D_n and D_p is the electron and hole diffusivity, respectively. We can split up the interval into two intervals that represent the two different layers. We then have the intervals $-L < x < 0$ and $0 < x < M$. From here on, we omit the notation where we write each variable as a function of x. In reality, the electrostatic potential and electron/hole density at the boundaries are discontinuous [4]. We, therefore, consider $-L^+ \leq x \leq 0^-$ and $0^+ \leq x \leq M^-$, so that the "redefined boundaries" are a part of the domain. Going forward, we write the boundary values without plus/minus superscript, but we keep in mind that the boundaries are not exactly on $x = -L, 0, M$. Richardson et al. now argue that the ionization potential and electron affinity are then not dependent on space since it is at the interface that these values change substantially as functions of space. They also make some arguments for natural choices for boundary values. Then they nondimensionalize the governing equations. In doing so, they assume that the panel is perfectly antisymmetric, which yields $L = M$. They thus $-L \leq x \leq 0$ to $-1 \leq x \leq 0$ and $0 \leq x \leq M$ to $0 \leq x \leq 1$. They also scale the electrical potential by the thermal voltage, k_BT/q, and reduce the electron and hole density in an involved manner, which consequently makes these on the scale of $\sim \exp(\Phi_{bi})$, where Φ_{bi} now has the nondimensionalized value. We can see the analogy to Chapter 2 with the case for silicon solar cells where we had that the built-in chemical potential was on the scale $\sim \log[n(0)]$. For a more detailed derivation of the boundary values and nondimensionalization, the reader can look at [14], Section 2.2-2.4. After nondimensionalizing, we end up with three parameters that play an important role. The first is δ, the ratio of the recombination rate to the diffusion rate. This parameter "is crucial in determining the electrical behavior of the device" as a value of δ that is not very small leads to "current-voltage characteristics [that] does not have the behavior observed in real devices [13]". Next is λ, a parameter that is not explained physically by Richardson et al., but has some analogy to, ν, the ratio between the Debye length and the thickness of the material. It is given by $\lambda := \frac{1}{L}\frac{\sqrt{\bar{\epsilon}k_BT}}{q\sqrt{\Pi_0}}$, where $\bar{\epsilon}$ is "similar to the permittivity of

free space" and Π_0 is the scaling factor for the electron and hole densities. Lastly we have the parameter θ, which is harder to explain what represents. It is given by $\theta := \frac{\exp(-E_g/(2k_BT))}{\lambda^2 N_D(u_1+u_2)}$, where N_D is a constant related to the jump conditions for n and p (see [14] and [4] for closer details), u_1 and u_2 are "appropriately chosen constants" in the calculation of the recombination. Now, with the assumption that the cell is antisymmetric, we can use that $n(-x) = p(x)$ and $\phi(-x) = -\phi(x)$ for $0 < x < 1$. It is then not necessary to look at both n and p, so we consider p. It is also assumed that the system is in a steady state. We can then use Equation (3.1) and (3.3). For (3.1), we use that the hole mobility is much greater than the electron mobility so we can neglect the electron density, similarly to what we argued in Chapter 2. We then get the same equations as in Chapter 2, except that n is replaced with p and ν with λ;

$$\phi'' + \frac{p}{\lambda^2} = 0 \tag{3.4}$$

$$p' + p\phi' + J = 0 \tag{3.5}$$

on $0 \leq x \leq 1$ with boundary conditions

$$p(1) = \exp\left(\frac{\Phi_{bi}}{2}\right); \quad \phi(0) = 0; \quad \phi(1) = -\frac{\Phi + \Phi_{bi}}{2}. \tag{3.6}$$

It is argued that J can be computed from

$$J = -\frac{2\delta(p(0)^2 - 1)}{\lambda^2\theta + p(0)}. \tag{3.7}$$

For this case, we let Φ vary continuously in space with $\Phi_{\min} \leq \Phi \leq \Phi_{\max}$ for two different scenarios. The following parameters are given:

Scenario	$\Phi_{\min}$	$\Phi_{\max}$	Φ_{bi}	δ	λ	θ
1	-20	0	4	e^{-8}	1	0.4
2	-5	10	4	e^{-8}	1	0.4

We now have equations in the same form as before, so we discretize and use the same method as in Chapter 2. We do not, however, have any start value for J, nor do we have the left boundary value of p, so we use the method as described in Section 2.3. We then use another expression for calculating J. We then need

to find the derivative of J with respect to p_0, where p_0 is the discretized value of $p(0)$;

$$\frac{\partial J}{\partial p_0} = -\frac{4\delta p_0(\lambda^2\theta + p_0) - 2\delta(p_0^2 - 1)}{(\lambda^2\theta + p_0)^2} \tag{3.8}$$

$$= -\frac{2\delta(p_0^2 + 2\lambda^2\theta p_0 + 1)}{(\lambda^2\theta + p_0)^2}. \tag{3.9}$$

3.1 Results for organic solar cells - DA-interface

The results for Scenario 1 using the method in as described in 2.3 are plotted in Figure 3.2. As we see, the flux is close to 0 for $-10 \leq \Phi \leq 0$. The results for Scenario 2 are plotted in Figure 3.3. We see that the flux is in the order of 10^{-3} for $-2 \leq \Phi \leq 10$. It is worth noting that for $\Phi = -4$, we do have flux, but not for $\Phi = 0$, which means that the flux is nonzero when the applied potential difference is zero; the holes are driven by the applied potential, not the built-in potential. We also see that the flux is quite small for $\Phi > 0$, which means that the direction of the applied voltage will greatly affect the flow characteristics.

We also look at how different values of the applied potential difference, Φ, affect the values of the electrostatic potential and the hole density. This is plotted in 3D for Scenario 1 and 2 in Figure 3.4, 3.5, 3.6 and 3.7. For Scenario 1 we see that the values of p for $-20 \leq \Phi \leq -12$ are unrealistically high. The values of p that correspond to this domain are non-physical. We see that for a certain value of the applied electrostatic potential difference, we get a hole density that is unrealistically high, and hence have we found a model that can describe where the transition from physical to non-physical values lie. As we expect, the electrostatic potential decreases quite rapidly near $x = 0$ for $\Phi = -20$, which is in accordance with the Poisson's equation. For Scenario 2 we have solutions for Φ and n that look more similar to what we saw for that of the silicon solar cell in Chapter 2. We get that the hole density near the left boundary increases quite rapidly as Φ approaches -5 from above.

We also see some interesting results for how different values of λ, a material prop-

erty parameter, and δ, the ratio of recombination rate to diffusion rate, change the system. We consider when $\lambda \in [0.2, 1]$ for the material constant and $\delta \in [e^{-8}, e^{-2}]$ for the recombination to diffusion rate. The results for these different parameter values are plotted in Figure 3.8 and 3.9. For the different parameter values of λ, we see that the closer λ is to zero, the closer the hole density between the boundaries is to zero. For the different values of the parameter δ, however, we get that values of δ closer to zero yield higher values for the left boundary of p.

Plots of the results

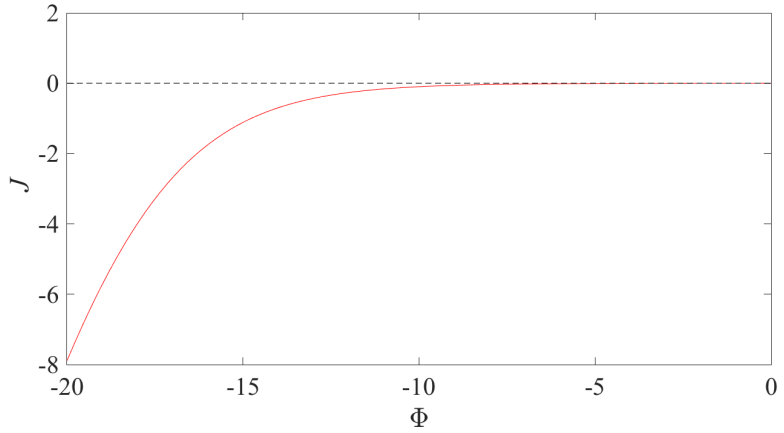


Figure 3.2: Plot of J against Φ for Scenario 1. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences. This result is in accordance with [14].

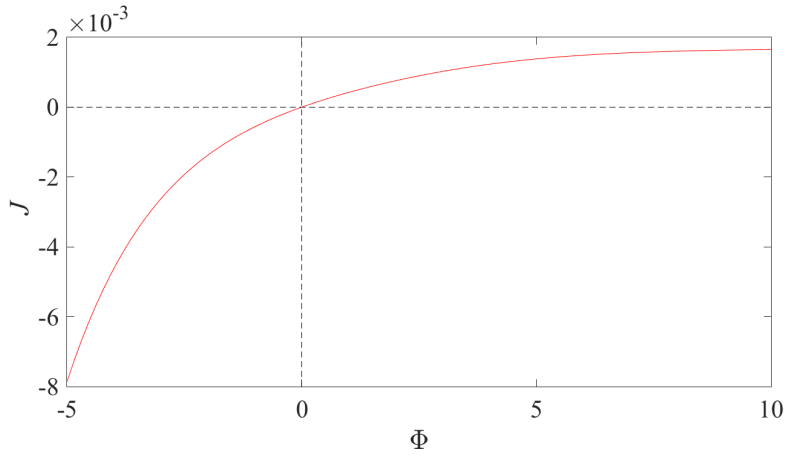


Figure 3.3: Plot of J against Φ for Scenario 2. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences. This result is in accordance with [14].

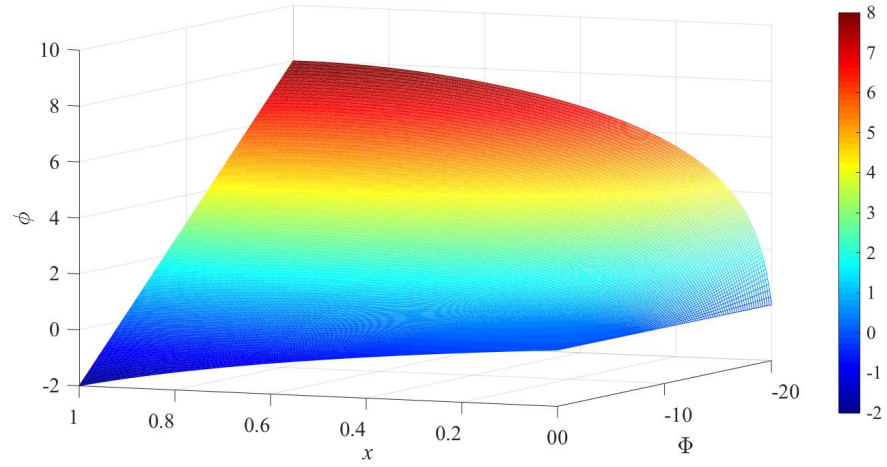


Figure 3.4: 3D-plot of ϕ as a function of x for the different values of Φ in Scenario 1. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences. This result is in accordance with [14].

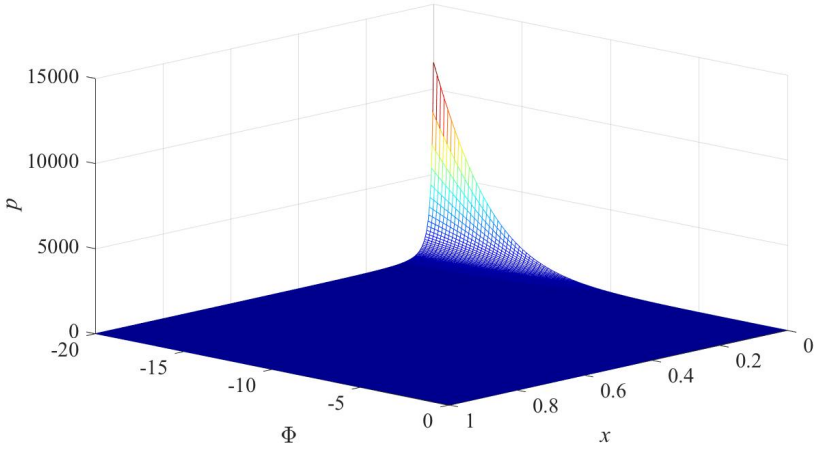


Figure 3.5: 3D-plot of p as a function of x for the different values of Φ in Scenario 1. This results is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences.

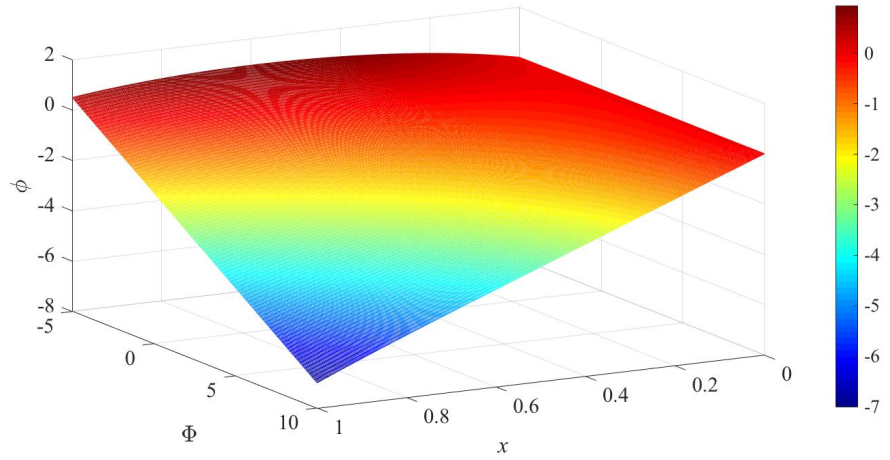


Figure 3.6: 3D-plot of ϕ as a function of x for the different values of Φ in Scenario 2. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences.

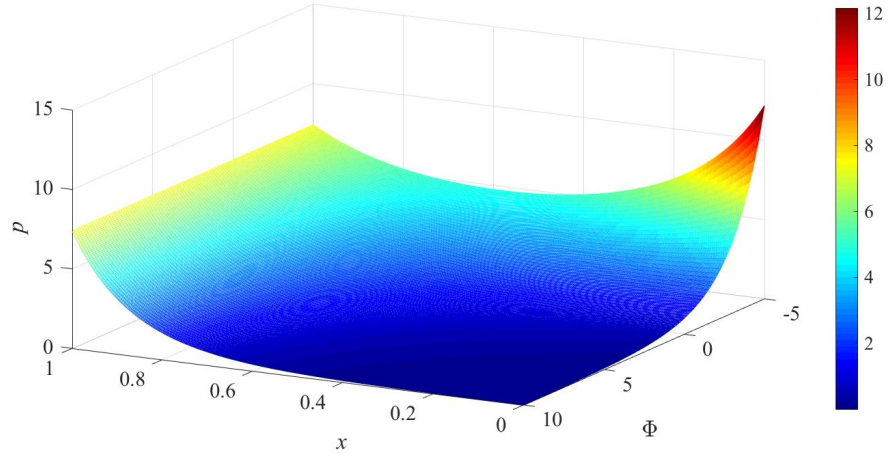


Figure 3.7: 3D-plot of p as a function of x for the different values of Φ in Scenario 2. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences.

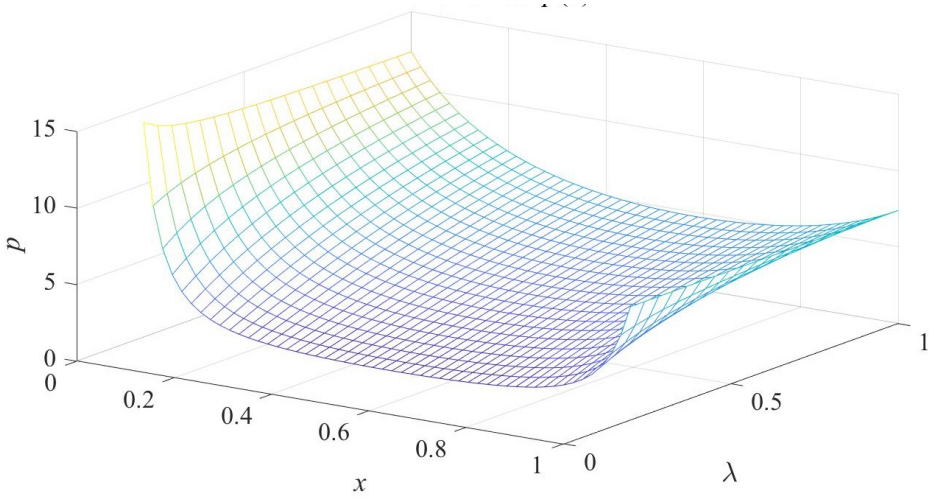


Figure 3.8: 3D-plot of p as a function of x for different values of $\lambda \in [0.2,\ 1]$. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences.

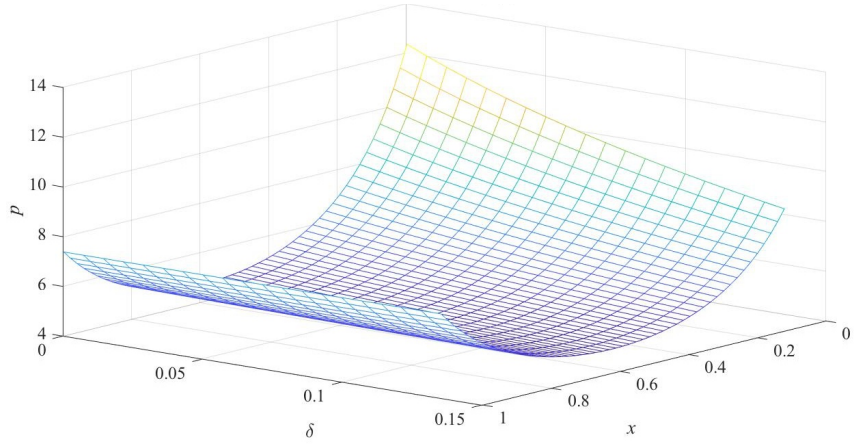


Figure 3.9: 3D-plot of p as a function of x for different values of $\delta \in [e^{-8},\ e^{-2}]$. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences.

We also consider how the applied potential difference influences the boundary value of p and the hole flux. This is plotted in Figure 3.10. Since we have one value of Φ that corresponds with on value of J, the values for the left boundary of p is plotted as a line in 3D. We see that as Φ and J approach zero from above, the left boundary value for p increases. We also see that for $J < 0$, we have almost a linear relationship between $p(0)$ and J where lower values of J yield higher values for $p(0)$. For $J > 0$, we get a decaying function for $p(0)$ that approaches zero. For $p(0)$ vs. Φ, however, we have a convex relationship with a form similar to what we see for the silicon solar case in Chapter 2, as we can see in Figure 2.10.

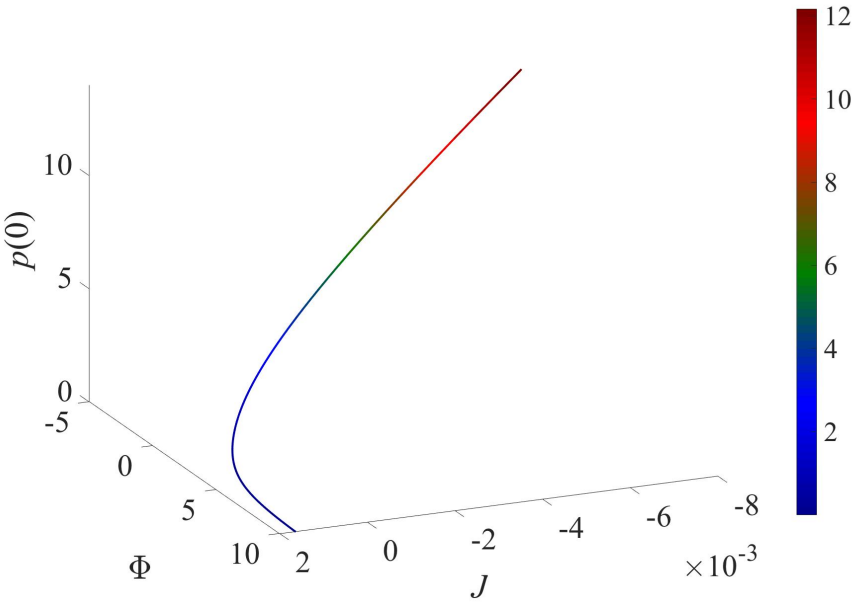


Figure 3.10: Plot of left boundary value of p against Φ and J in Scenario 2. This result is when using the Newton-Raphson method in Section 2.3 (with the necessary replacements for Chapter 3) with second-order finite differences.

Results from the calculations for organic solar cells

We have found a relation between the flux, J, and the applied potential difference, Φ, for different cases for values of Φ. We have discovered that the values of the hole density near the left boundary are very high when we approach $\Phi = -20$ from above. For Scenario 2, we find that the left boundary value of the hole density is close to 0 as Φ approaches 10 from below, and increase quite significantly

when approaching -5 from above. We have also found that there is a continuous relation between $p(0)$ and J and between p and ϕ. Further, it has been shown that the hole density changes as λ changes; we get higher values for the left boundary and lower values for the values in the middle of the layer as λ decreases. It also changes as the ratio of the recombination rate to the diffusion rate, δ, changes, with higher left boundary values for smaller values of δ.

Chapter 4

Conclusion and further work

In this thesis, we have found several results for a coupled system with drift-diffusion and electrostatic potential equations.

For a silicon solar cell, we have looked at the glass layer for a silicon solar cell and found results for different cases. We have found an iteration scheme system that takes the potential difference and electron flux as given values and then finds the left boundary value for the electron density. By using known theorems and claims that are stated, bounds that guarantee convergence have been found with supplementing strict bounds that are found experimentally. We have also found a Newton-Raphson iteration method that we have applied on the same system, which converges in fewer iteration steps and converges for scenarios where the iteration scheme does not. We have looked at how different values of the ratio of the Debye length to the glass thickness influence the desired solutions. We have made comparisons between different orders of finite differences. A system where the flux is a variable in the system has also been considered, but with less fortune than for the other iteration methods. We have also shown that differentiating the drift-diffusion equation can be useful when we do not know the flux. We have compared this differentiated equation with the non-differentiated equation for different cases. We have seen how different scenarios and solution methods influence that value of the resistance that we subsequently want to model.

For the organic solar cell, there has been shown a relation between the flux and the applied potential difference for different values of the latter. The value of the

applied electrostatic potential difference determines many of the properties of the cell. We have also shown how the different values of the parameters λ influence the system. We have also looked at how the ratio of the recombination rate to the diffusion rate changes the values for the hole density.

A good direction for further work would be to prove the norms of the different expressions in the analysis of the fixed-point method in Subsection 2.1.1. This gives a more mathematical and correct way of proving the bounds for the norm of the Jacobian.

There are several other panel types and cases to consider with this coupled system of differential equations. Including several layers of organic material and not assuming the layers yield symmetric/anti-symmetric values will should a more realistic result. Also, the interfaces between the organic solar cell layers is in reality not flat, so a two or three-dimensional consideration with different shapes for the interface would give more realistic results for modeling the current-voltage characteristics.

The cases considered in this thesis have all been on the electrical properties of the panel itself, not when any light is shown on it. Including radiation and using the continuity equation can be useful to see how a panel behaves under operation.

Bibliography

[1] Barnes, R. J. Prof. at the University of Minnesota. (Downloaded 14.02.2018). *Matrix Differentiation.* URL: `https://atmos.washington.edu/~dennis/MatrixCalculus.pdf`

[2] Black, J. P., Breward, C. J. W., Howell, P. D., & Young, R. J. S. (2013). Mathematical modeling of contact resistance in silicon photovoltaic cells. *SIAM Journal on Applied Mathematics, 73*(5), 1906-1925. DOI: 10.1137/130911974

[3] Buxton, G. A. & Clarke, N. (2007). Computer simulation of polymer solar cells. *Modelling and Simulation in Materials Science and Engineering, 15*(2), 13-26. DOI: 10.1088/0965-0393/15/2/002

[4] Crone, B. K., Davids, P. S., Campbell, I. H., & Smith, D. L. (1998). Device model investigation of single layer organic light emitting diodes. *Journal of Applied Physics, 84*(2), 833. DOI: 10.1063/1.368144

[5] De Leon, N. Prof. at Indiana University Northwest. (Downloaded 31.01.2018). *Electron Affinity.* URL: `http://www.iun.edu/~cpanhd/C101webnotes/modern-atomic-theory/electron-affinity.html`

[6] De Vos, A. (1980). Detailed balance limit of the efficiency of tandem solar cells. *Journal of Physics D: Applied Physics, 13*(5), 839-846. DOI: 10.1088/0022-3727/13/5/018

[7] Duffie, J. A. & Beckman, W. A. (2013). *Solar Engineering of Thermal Processes (4th ed.).* Hoboken, New Jersey, USA: Wiley.

[8] Fraunhofer ISE. (01.12.2014). *New world record for solar cell efficiency at 46%* [Press release]. URL: `https://www.`

ise.fraunhofer.de/en/press-media/press-releases/2014/new-world-record-for-solar-cell-efficiency-at-46-percent.html

[9] Gautschi, W. (2012). *Numerical analysis (2nd ed.)*. Basel, Switzerland: Birkhäuser.

[10] Heliatek. (08.02.2016). *Heliatek sets new Organic Photovoltaic world record efficiency of 13.2%* [Press release]. URL: http://www.heliatek.com/en/press/press-releases/details/heliatek-sets-new-organic-photovoltaic-world-record-efficiency-of-13-2

[11] Hughes, R. C. (1975). Hole mobility and transport in thin SiO_2 films. *Applied Physics Letters, 26*(8), 436. DOI: 10.1063/1.88200

[12] Hyperphysics. (Downloaded 31.01.2018). *Ionization Energies*. URL: http://hyperphysics.phy-astr.gsu.edu/hbase/Chemical/ionize.html

[13] Potscavage, W. J., Yoo, S. and Kippelen, B. (2008). Origin of the open-circuit voltage in multilayer heterojunction organic solar cells. *Applied Physics Letters, 93*(19), 193308. DOI: 10.1063/1.3027061

[14] Richardson, G., Please, C., Foster, J., & Kirkpatrick, J. (2012). Asymptotic solution of a model for bilayer organic diodes and solar cells. *SIAM Journal on Applied Mathematics, 72*(6), 1792-1817. DOI: 10.1137/110825807

[15] Rühle, S. (2016). Tabulated values of the Shockley–Queisser limit for single junction solar cells. *Solar Energy, 130*. 139–147. DOI: 10.1016/j.solener.2016.02.015

[16] Sauer, T. (2013). *Numerical Analysis (2nd ed.)*. London, United Kingdom: Pearson.

[17] Seeger, K. (2004). *Semiconductor Physics: An Introduction (9th ed.)*. Berlin, Germany: Springer.

[18] Sze, S. & Lee, M. K. (2013). *Semiconductor devices (3rd ed.)*. Hoboken, New Jersey, USA: Wiley.

[19] Wenham, S. R., Green, M. A., Watt. M. E. & Corkish, R. (2007). *Applied Photovoltaics*. Oxfordshire, United Kingdom / New York, USA: Earthscan / Routledge / Taylor & Francis.

Appendix A

Numerical methods for differential equations

A.1 Discretizing equations

When solving a differential equation numerically for a variable y as a function of x in space, on a domain, $\mathcal{D} := \{x : a \leq x \leq b, x \in \mathbb{R}\}$, with boundary $\partial\mathcal{D} := \{a, b\}$, we discretize by letting the vector $\widetilde{\boldsymbol{y}} = [y_0, \ldots, y_{N+1}] \in \mathbb{R}^{N+2}$ represent the function y in space. We call N the *discretization number*, which is the number of points in space in between the boundaries. We then have $N + 2$ entries of $\widetilde{\boldsymbol{y}}$ where the first and last entry of $\widetilde{\boldsymbol{y}}$ is the boundary values of y on $\partial\mathcal{D}$. When we have $N + 2$ entries in space for $\widetilde{\boldsymbol{y}}$, we get $N + 1$ space steps. For this thesis, we let the step length

$$(\Delta\, x)_i := x_i - x_{i-1}, \quad i = 1, \ldots, N + 1 \tag{A.1}$$

be constant. We then have

$$\Delta\, x = \frac{b - a}{N + 1}. \tag{A.2}$$

We can thus define the spatial discretization vector as

$$\widetilde{\boldsymbol{x}} := [x_0, \ldots, x_{N+1}]^T \quad \text{where} \quad x_i := a + i \cdot \Delta\, x, \quad i = 0, \ldots, N + 1. \tag{A.3}$$

For the boundary values, on $\partial\mathcal{D}$, we have the elements

$$y_0 := y(a); \quad y_{N+1} := y(b) \tag{A.4}$$

and for the elements in between the boundaries, in $\mathcal{D}\backslash\partial\mathcal{D}$, we have the elements

$$y_i := y(x_i), \quad 1 \leq i \leq N. \tag{A.5}$$

We let $\boldsymbol{x}$ be the vector with elements between the boundaries, in $\mathcal{D}\backslash\partial\mathcal{D}$, with $\boldsymbol{y}$ corresponding to the vector $\boldsymbol{x}$. Similarly, we let and $\widetilde{\boldsymbol{x}}$ be the vector with elements in the entire domain, $\mathcal{D}$, with $\widetilde{\boldsymbol{y}}$ corresponding to the vector $\widetilde{\boldsymbol{x}}$. For notation purposes, we also let $\hat{\boldsymbol{x}}$ be the points from and including the left boundary value to the right boundary value, with $\hat{\boldsymbol{y}}$ corresponding to the vector $\hat{\boldsymbol{x}}$. Mathematically, this means that

$$\begin{aligned} \boldsymbol{x} &:= [x_1, \ldots, x_N]^T; & \boldsymbol{y} &:= [y_1, \ldots, y_N]^T; \\ \widetilde{\boldsymbol{x}} &:= [x_0, \ldots, x_{N+1}]^T; & \widetilde{\boldsymbol{y}} &:= [y_0, \ldots, y_{N+1}]^T; \\ \hat{\boldsymbol{x}} &:= [x_1, \ldots, x_N]^T; & \hat{\boldsymbol{y}} &:= [y_1, \ldots, y_N]^T. \end{aligned} \tag{A.6}$$

The discrete solution will then be $\widetilde{\boldsymbol{y}}$ as a function of $\widetilde{\boldsymbol{x}}$.

A.2 Finite difference formulas

In this section we state different finite difference formulas which are approximations of derivatives at discrete points. In this thesis we only look at derivatives in one spatial dimension.

The first-order finite forward difference formula for first-order derivatives is given by

$$y_i' = \frac{-y_i + y_{i+1}}{\Delta\, x} + \mathcal{O}\left(\Delta\, x\right). \tag{A.7}$$

The second-order finite forward difference formula for first-order derivatives is given by

$$y_i' = \frac{-3y_i + 4y_{i+1} - y_{i+2}}{2\,\Delta\, x} + \mathcal{O}\left((\Delta\, x)^2\right). \tag{A.8}$$

The second-order finite central difference formula for first-order derivatives is given by

$$y_i' = \frac{-y_{i-1} + y_{i+1}}{2\,\Delta\,x} + \mathcal{O}\left((\Delta\,x)^2\right). \tag{A.9}$$

The second-order finite central difference formula for second-order derivatives is given by

$$y_i'' = \frac{y_{i-1} - 2y_i + y_{i+1}}{(\Delta\,x)^2} + \mathcal{O}\left((\Delta\,x)^2\right). \tag{A.10}$$

A.3 Matrix expressions of certain discretized differential equations using finite difference approximations

We consider the differential equation

$$y'' = Cz \tag{A.11}$$

in $\mathcal{D} := \{x : a \leq x \leq b, x \in \mathbb{R}\}$ with Dirichlet boundary conditions for y at $x = a$ and $x = b$. We can discretize (A.11) as described in A.1 to get

$$y_i'' = Cz_i \tag{A.12}$$

with $y_i = y(i\Delta\,x), z_i = z(i\Delta\,x)$ for $i = 0, \ldots, N+1$ and $\Delta\,x = \frac{1}{N+1}$. We can use finite differences to approximate the derivatives in the differential equation. With Dirichlet boundary conditions on both sides for y, it is natural to use finite central differences.

Lemma A.1. *Using second-order finite central differences on* (A.12)*, we get*

$$\frac{y_{i-1} - 2y_i + y_{i+1}}{(\Delta\,x)^2} = Cz_i \tag{A.13}$$

$$\Longleftrightarrow$$

$$y_{i-1} - 2y_i + y_{i+1} = (\Delta\,x)^2\,Cz_i \tag{A.14}$$

which is defined for $i = 1, \ldots, N$. At x_1 and x_N, we will have the conditioned boundary values y_0 and y_{N+1} in the discretized equations.

We now want to write this in matrix form. We then want the boundary values separated from the matrix system. The discretized equation at x_1 will be

$$y_0 - 2y_1 + y_2 = (\Delta\, x)^2 C z_i \tag{A.15}$$

and the discretized equation at x_N will be

$$y_{N-1} - 2y_N + y_{N+1} = (\Delta\, x)^2 C z_i. \tag{A.16}$$

We can put y_0 and y_{N+1} in a vector so that they are not a part of the matrix system we want to write.

Corollary A.2. *We now see that the approximations for the differential equation at $x_i, i = 1, \ldots, N$ can be written in matrix form as*

$$S_0\ \boldsymbol{y} + \boldsymbol{r}_{y,S_0} = (\Delta\, x)^2\, C\boldsymbol{z} \tag{A.17}$$

with $S_0, \boldsymbol{r}_{y,S_0}$ as defined in Section B.2. For this system, $\boldsymbol{y}$ has a discretization error of $\mathcal{O}((\Delta\, x)^2)$.

We now consider the differential equation

$$y'z + z' = C \tag{A.18}$$

in $\mathcal{D} := \{x : a \leq x \leq b, x \in \mathbb{R}\}$. We want to solve for z with Dirichlet boundary condition for z at $x = b$. We can discretize (A.18) as described in A.1 to get

$$y_i' z_i + z_i' = C \tag{A.19}$$

with $y_i = y(i\Delta\, x), z_i = z(i\Delta\, x)$ for $i = 0, \ldots, N+1$ and $\Delta\, x = \frac{1}{N+1}$. We can use finite differences to approximate the derivatives in the differential equation. Since we have a boundary condition on the right side for z, we can use finite forward differences.

Lemma A.3. *Using first-order finite forward differences on* (A.19)*, we get*

$$\frac{-y_i + y_{i+1}}{\Delta\, x}\, z_i + \frac{-z_i + z_{i+1}}{\Delta\, x} = C \tag{A.20}$$

which is defined for $i = 0, \ldots, N$. At x_N, we will have the conditioned boundary value, z_{N+1}, in the discretized equation.

We can write (A.20) as

$$(-y_i + y_{i+1})z_i - z_i + z_{i+1} = \Delta\, x\, C. \tag{A.21}$$

We want to write this set of equations in matrix form. We then want the conditioned boundary value, z_{N+1}, separated from the matrix system. The discretized equation at x_N will be

$$(-y_N + y_{N+1})z_N - z_N + z_{N+1} = \Delta\, x\, C. \tag{A.22}$$

We can put z_{N+1} in a vector that so it is not a part of the matrix system we want to find. We now see that the approximations for the differential equation for $x_i, i = 0, \ldots, N$ can be written in matrix form as

$$\text{diag}(\hat{F_+}\,\hat{\boldsymbol{y}} + \boldsymbol{r}_{y,\hat{F_+}})\hat{\boldsymbol{z}} + \hat{F_+}\,\hat{\boldsymbol{z}} + \boldsymbol{r}_{z,\hat{F_+}} = \Delta\, C\mathbf{1}_{N+1}. \tag{A.23}$$

with $\hat{F_+}\,, \boldsymbol{r}_{y,\hat{F_+}}\,, \boldsymbol{r}_{z,\hat{F_+}}$ as defined in Section B.2.

Corollary A.4. *We can factorize* (A.23) *to get*

$$\left\{\text{diag}(\hat{F_+}\,\hat{\boldsymbol{y}} + \boldsymbol{r}_{y,\hat{F_+}}) + \hat{F_+}\right\}\hat{\boldsymbol{z}} + \boldsymbol{r}_{z,\hat{F_+}} = \Delta\, xC\mathbf{1}_{N+1}. \tag{A.24}$$

For this system, $\boldsymbol{z}$ *has a discretization error of* $\mathcal{O}(\Delta\, x)$.

In MATLAB, it is much faster to compute $\boldsymbol{x} \circ \boldsymbol{y}$ rather than $\text{diag}(\boldsymbol{x})\boldsymbol{y}$, since in the latter expression, the program needs to create the diag-matrix before multiplication and then matrix multiply. The first expression can be implemented as

```
x.*y
```

If we, however, need to use the diag-matrix, we can use the integrated function `diag()` in MATLAB and implement the latter expression as

```
diag(x)*y
```

For (A.23), we can use the Hadamard product to compute $\text{diag}(\hat{F_+}\,\hat{\boldsymbol{y}} + \boldsymbol{r}_{y,\hat{F_+}})\hat{\boldsymbol{z}}$ which will be much faster to compute. For (A.24), however, we have factorized the terms in front of $\hat{\boldsymbol{z}}$. In that case, we can not use Hadamard notation because $\hat{F_+}$ needs to be added with the diag-matrix before multiplication.

We now consider the differential equation

$$z'' + y'z' = C\,z^2 \tag{A.25}$$

in $\mathcal{D} := \{x : a \le x \le b, x \in \mathbb{R}\}$. We want to solve for z with Dirichlet boundary conditions for z at $x = a$ and $x = b$. We can discretize (A.25) as described in A.1 to get

$$z_i'' + y_i'z_i' = C\,z_i^2 \tag{A.26}$$

for $y_i = y(i\Delta\,x), z_i = z(i\Delta\,x)$ for $i = 0, \ldots, N+1$ and $\Delta\,x = \frac{1}{N+1}$. We can use finite differences to approximate the derivatives in the differential equation. Since we have boundary conditions on both sides for z, it is natural to use finite central differences.

Lemma A.5. *Using second-order finite central differences on* (A.19)*, we get*

$$\frac{z_{i-1} - 2z_i + z_{i+1}}{(\Delta\,x)^2} + \frac{-y_{i-1} + y_{i+1}}{2\Delta\,x} \cdot \frac{-z_{i-1} + z_{i+1}}{2\Delta\,x} = C\,z_i^2 \tag{A.27}$$

which is defined for $i = 1, \ldots, N$. At x_1 and x_N, we will have the conditioned boundary values z_0 and z_{N+1} in the discretized equations.

We can write (A.27) as

$$4(z_{i-1} - 2z_i + z_{i+1}) + (-y_{i-1} + y_{i+1})(-z_{i-1} + z_{i+1}) - 4C\,(\Delta\,x)^2\,z_i^2 = 0. \tag{A.28}$$

We want to write this system in matrix form. We then want the conditioned boundary values z_0 and z_{N+1} separated from the matrix system. The discretized equation at x_1 will be

$$4(z_0 - 2z_1 + z_2) + (-y_0 + y_1)(-z_1 + z_2) - 4C\,(\Delta\,x)^2\,z_1^2 = 0. \tag{A.29}$$

The discretized equation at x_N will be

$$4(z_{N-1} - 2z_N + z_{N+1}) + (-y_{N-1} + y_{N+1})(-z_{N-1} + z_{N+1}) - 4C\,(\Delta\,x)^2\,z_N^2 = 0. \tag{A.30}$$

We can put z_0 and z_{N+1} in a vector that so it is not a part of the matrix system we want to find. We now see that the approximations for the differential equation for $x_i, i = 1, \ldots, N$ can be written in matrix form as

$$4(S_0\boldsymbol{z} + \boldsymbol{r}_{z,S_0}) + \operatorname{diag}(F_0\boldsymbol{y} + \boldsymbol{r}_{y,F_0})(F_0\boldsymbol{z} + \boldsymbol{r}_{z,F_0}) - 4C(\Delta\,x)^2\boldsymbol{z}^{\circ 2} = 0 \tag{A.31}$$

with $F_0, \boldsymbol{r}_{z,S_0}, \boldsymbol{r}_{z,F_0}, \boldsymbol{r}_{y,F_0}$ as defined in Section B.2.

Corollary A.6. *We can factorize* (A.23) *to get*

$$\{4S_0 + \mathrm{diag}(F_0\boldsymbol{y} + r_{y,F_0})F_0 - 4C\mathrm{diag}(\boldsymbol{z})\}\,\boldsymbol{z} + 4\boldsymbol{r}_{z,S_0} + \mathrm{diag}(F_0\boldsymbol{y} + \boldsymbol{r}_{r_y,S_0})\boldsymbol{r}_{z,F_0} = 0. \tag{A.32}$$

For this system, $\boldsymbol{z}$ *has a discretization error of* $\mathcal{O}\left((\Delta\, x)^2\right)$.

We now look at (A.19) to seek approximations for the derivatives that give errors on the $\mathcal{O}\left((\Delta\, x)^2\right)$. We can use second-order finite forward differences at x_i for $i = 0, \ldots, N-1$, and first-order finite differences at x_N. We then get the following set of equations:

$$\frac{-3y_i + 4y_{i+1} - y_{i+2}}{2\Delta\, x}\, z_i + \frac{-3z_i + 4z_{i+1} - z_{i+2}}{2\Delta\, x} = C \tag{A.33}$$

for $i = 0, \ldots, N-1$, and

$$\frac{-y_N + y_{N+1}}{\Delta\, x}\, z_N + \frac{-z_N + z_{N+1}}{\Delta\, x} = C \tag{A.34}$$

at x_N.

Lemma A.7. *Using second-order finite forward differences on* (A.19)*, we get*

$$(-3y_i + 4y_{i+1} - y_{i+2})\, z_i - 3z_i + 4z_{i+1} - z_{i+2} = 2\Delta\, x\, C \tag{A.35}$$

for $i = 0, \ldots, N-1$*, and*

$$(-y_N + y_{N+1})z_N - z_N + z_{N+1} = \Delta\, x\, C \tag{A.36}$$

At x_{N-1} *and* x_N*, we will have the conditioned boundary value,* z_{N+1}*, in the discretized equation.*

We can write this set of equations in matrix form as

$$\mathrm{diag}\left(K\hat{\boldsymbol{y}} + \begin{pmatrix} \mathbf{0}_{N-1} \\ -y_{N+1} \\ y_{N+1} \end{pmatrix}\right)\hat{\boldsymbol{z}} + K\hat{\boldsymbol{z}} + \begin{pmatrix} \mathbf{0}_{N-1} \\ -z_{N+1} \\ z_{N+1} \end{pmatrix} = \Delta\, x \begin{pmatrix} 2 \cdot \mathbf{1}_N \\ 1 \end{pmatrix} \cdot C \tag{A.37}$$

where the matrix, K, is defined as

$$K := \begin{pmatrix} -3 & 4 & -1 & 0 & \dots & 0 \\ 0 & -3 & 4 & -1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -3 & 4 & -1 \\ 0 & \dots & 0 & 0 & -3 & 4 \\ 0 & \dots & 0 & 0 & 0 & -1 \end{pmatrix} \tag{A.38}$$

A.4 Methods for solving non-linear equations

This section is taken from [9]*, Section 4.9.*

When we solve systems of nonlinear equations, we look at methods that solve equations in the form

$$\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}_N, \qquad \boldsymbol{f}: \quad \mathbb{R}^N \to \mathbb{R}^N. \tag{A.39}$$

In this section, we look two methods that solve systems on this form; the fixed-point method and the Newton-Raphson method.

A.4.1 Fixed-point method

We can rewrite (A.39) to $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{\xi}(\boldsymbol{x})$ and consider the fixed-point iteration

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{\xi}(\boldsymbol{x}^{(k)}), \quad k = 0, 1, 2, \dots \tag{A.40}$$

We say that $\boldsymbol{\xi}: \quad \mathbb{R}^N \to \mathbb{R}^N$ is a contraction map on a set $\mathcal{R} \subseteq \mathbb{R}^N$ if there exists a constant γ with $0 < \gamma < 1$ such that

$$\|\boldsymbol{\xi}(\boldsymbol{x}) - \boldsymbol{\xi}(\boldsymbol{x}^*)\| \leq \gamma \|\boldsymbol{x} - \boldsymbol{x}^*\| \quad \text{for all } \boldsymbol{x}, \boldsymbol{x^*} \in \mathcal{R} \tag{A.41}$$

for some appropriate vector norm. The convergence of the fixed point can be analyzed using the following theorem:

Theorem A.8. (Contraction mapping principle)
Let $\mathcal{R} \subseteq \mathbb{R}^N$ be a complete subset of $\mathbb{R}^N$ (either bounded and closed, or all of $\mathbb{R}^N$). If $\boldsymbol{\xi}: \ \mathbb{R}^N \to \mathbb{R}^N$ is contractive, such as in (A.41)*, and maps $\mathcal{R}$ into $\mathcal{R}$, then*

(i) iteration (A.40) *is well defined for any $\boldsymbol{x}^{(0)} \in \mathcal{R}$ and converges to a unique fixed-point $\boldsymbol{\beta} \in \mathcal{R}$,*

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{\beta} \tag{A.42}$$

(ii) for $k = 1, 2, 3, ...$ there holds

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{\beta}\right\| \leq \frac{\gamma^k}{1-\gamma}\left\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\right\| \tag{A.43}$$

and

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{\beta}\right\| \leq \gamma^k\left\|\boldsymbol{x}^{(0)} - \boldsymbol{\beta}\right\| \tag{A.44}$$

Proof. Proof can be found in [9] □

For applications, it is useful to find an estimate for the γ in theorem A.8. We can find such an approximation by Taylor expanding the mapping function, $\boldsymbol{\xi}$. Taylor expanding $\boldsymbol{\xi}(\boldsymbol{x})$ around $\boldsymbol{x}^*$ yields

$$\boldsymbol{\xi}(\boldsymbol{x}) = \boldsymbol{\xi}(\boldsymbol{x}^*) + \mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{\xi}(\boldsymbol{x})}(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*) + \mathcal{O}\left(\|\boldsymbol{x} - \boldsymbol{x}^*\|^2\right). \tag{A.45}$$

We can thus find an approximation to γ for equation (A.41):

$$\begin{aligned} \|\boldsymbol{\xi}(\boldsymbol{x}) - \boldsymbol{\xi}(\boldsymbol{x}^*)\| &= \left\|\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{\xi}(\boldsymbol{x})}(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*) + \mathcal{O}\left(\|\boldsymbol{x} - \boldsymbol{x}^*\|^2\right)\right\| \\ &\approx \left\|\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{\xi}(\boldsymbol{x})}(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)\right\| \leq \left\|\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{\xi}(\boldsymbol{x})}(\boldsymbol{x}^*)\right\|\|\boldsymbol{x} - \boldsymbol{x}^*\|. \end{aligned} \tag{A.46}$$

Hence, the norm of the Jacobian will give an indicator if the mapping is a contraction mapping or not, and thus give us an indicator for when a fixed-point methods converges or not.

A.4.2 Newton-Raphson method

The Newton-Raphson method can deal with systems of nonlinear equations by reducing the nonlinear problem to an infinite sequence of linear problems. By linearizing the current approximation, we end up with a systems of linear algebraic equations.

We write equation (A.39) more explicitly:

$$f_i(x_1, \ldots, x_N) = 0, \quad i = 1, \ldots, N \tag{A.47}$$

and given an approximation to $\boldsymbol{x}^{(0)}$ to a solution $\boldsymbol{\beta} \in \mathbb{R}^N$, the i-th equation in (A.47) is linearized at $\boldsymbol{x} = \boldsymbol{x}^{(0)}$ by truncating the Taylor expansion of f_i at $\boldsymbol{x}^{(0)}$ after the linear terms which yields the following:

$$f_i(\boldsymbol{x}^{(0)}) + \sum_{j=1}^{N} \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}^{(0)})(x_j - x_j^{(0)}) = 0, \quad i = 1, \ldots, N \tag{A.48}$$

or written in vector form

$$\boldsymbol{f}(\boldsymbol{x}^{(0)}) + \mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{f}(\boldsymbol{x})}(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) = \boldsymbol{0}. \tag{A.49}$$

The solution, $\boldsymbol{x}$, of (A.49) will be taken to be the next approximation. Thus, in general, starting with an initial approximation $\boldsymbol{x}^{(0)}$, Newton-Raphson's method will generate a sequence of approximations $\boldsymbol{x}^{(k)} \in \mathbb{R}^N$ by means of

$$\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{f}(\boldsymbol{x})}(\boldsymbol{x}^{(k)})\boldsymbol{d}^{(k)} = -\boldsymbol{f}(\boldsymbol{x}^{(k)}), \quad k = 0, 1, 2, \ldots \tag{A.50}$$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{d}^{(k)}, \quad k = 0, 1, 2, \ldots \tag{A.51}$$

where we assume that the matrix $\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{f}(\boldsymbol{x})}(\boldsymbol{x}^k)$ in (A.50) is nonsingular for each k. This will be the case if $\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{f}(\boldsymbol{x})}(\boldsymbol{\beta})$ is nonsingular and $\boldsymbol{x}^{(0)}$ is sufficiently close to $\boldsymbol{\beta}$. In that case, the Newton-Raphson's method converges quadratically to $\boldsymbol{\beta}$ in the one-dimensional case, $N = 1$, that is, $\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{\beta}\right\| = \mathcal{O}\left(\left\|\boldsymbol{x}^{(k)} - \boldsymbol{\beta}\right\|^2\right)$ as $k \to \infty$.

Writing (A.51) in the form

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \left[\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{f}(\boldsymbol{x})}(\boldsymbol{x}^{(k)})\right]^{-1} \boldsymbol{f}(\boldsymbol{x}^{(k)}), \quad k = 0, 1, 2, \ldots \tag{A.52}$$

brings out the formal analogy with Newton-Raphson's method that we know for a single equation;

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}, \quad k = 0, 1, 2, \ldots \tag{A.53}$$

However, it is not necessary to compute the inverse of the Jacobian at each step; it is more efficient to solve the linear system directly as in (A.50) & (A.51).

A.5 Numerical integration

This section is based on [16]*, Section 5.2.*

In this section we look at how to integrate a function numerically. In particular we consider the Trapezoidal Rule.

Theorem A.9. (Trapezoidal rule)
Define $\mathcal{D} = \{x : a \leq x \leq b, x \in \mathbb{R}\}$*. Given a discretization vector* $\widetilde{\boldsymbol{y}} = [y_0, \ldots, y_{N+1}]^T \in \mathbb{R}^{N+2}$ *that represents* y *in* $\mathcal{D}$ *with equal spacing,* $\Delta x = \frac{1}{N+1}$*, and the inner points* $\boldsymbol{y} = [y_1, \ldots, y_N]^T \in \mathbb{R}^N$*, then the integral*

$$\int_a^b y(x)dx \tag{A.54}$$

can be calculated numerically by the following:

$$\int_a^b y(x)dx = \frac{\Delta x}{2} \sum_{i=1}^{N+1} y_{i-1} + y_i + \mathcal{O}\left((\Delta x)^2\right) \tag{A.55}$$

$$= \Delta x \left\{ \frac{1}{2}(y_0 + y_{N+1}) + \sum_{i=1}^{N} y_i \right\} + \mathcal{O}\left((\Delta x)^2\right) \tag{A.56}$$

$$= \Delta x \left\{ \frac{1}{2}(y_0 + y_{N+1}) + \mathbf{1}_N^T \boldsymbol{y} \right\} + \mathcal{O}\left((\Delta x)^2\right). \tag{A.57}$$

In MATLAB, it is useful to use the built in `sum()`-function and it can be useful to use the Hadamard power. If we for example want to integrate $1/y$ from a to b, we discretize y with the vector $\widetilde{\boldsymbol{y}} \in \mathbb{R}^{N+2}$ with equal spacing from a to b. Noting that the indices of $\widetilde{\boldsymbol{y}}$ will start from 1 and end at $N+2$ in MATLAB, we can implement the trapezoidal rule for this case as

```
I = Delta_x*(1/2*(1/y(1)+1/y(N+2))+sum((y(2:N+1)).^(-1)));
```

A.6 Analysis of effective order of finite difference approximations

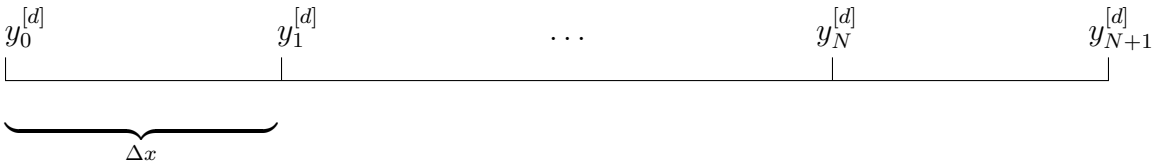
We want to look at the error for the discretization method for the derivatives. We look at how much the error changes by doubling the discretization number, or equivalently, by halving the spacing, Δx. We hence define $N_d := 2^d - 1, d \in \mathbb{N}$. We have $N_d \simeq 2\,N_{d-1}$. The spacing will for N_d be $\Delta\, x = (b-a)(N_d+1)^{-1} = (b-a)2^{-d}$ and for N_{d-1} be $\Delta\, x = (b-a)(N_{d-1}+1)^{-1} = (b-a)2^{-d+1}$, so the spacing is halved as the power of 2 for the discretization number is increased by one.

Let $\widetilde{\boldsymbol{y}}^{[d]} = [y_0^{[d]}, \ldots, y_{N_d+1}^{[d]}]^T$ be the discrete approximation vector of the variable y, with N_d as discretization number, and let $\widetilde{\boldsymbol{y}} = [y(x_0), \ldots, y(x_{N_d+1})]^T$ be the true solution. The relative error is then defined as $e^{[d]} = \frac{\left\|\widetilde{\boldsymbol{y}}^{[d]} - \widetilde{\boldsymbol{y}}\right\|}{\|\widetilde{\boldsymbol{y}}\|}$. As we do not know $\widetilde{\boldsymbol{y}}$, this error cannot be computed. We instead approximate the error. Such an approximation is given later in this section. When using a method of order s, we expect $e^{[d]} = \mathcal{O}\left((\Delta\, x)^s\right)$. This implies that $e^{[d]} \approx \left(\frac{1}{2}\right)^s e^{[d-1]}$. Solving for s gives the following approximation for s:

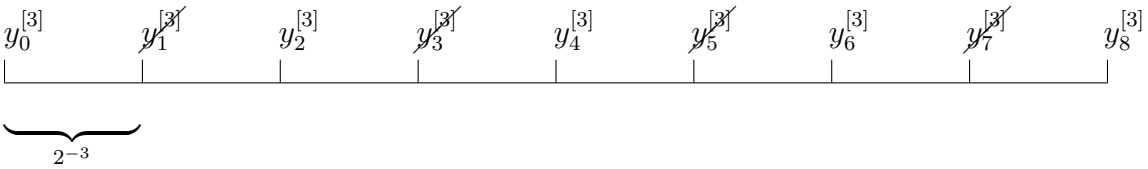
$$s = \log_2\left(\frac{e^{[d-1]}}{e^{[d]}}\right). \tag{A.58}$$

The reason we use $\log_2$, is because the different discretization numbers are powers of 2.

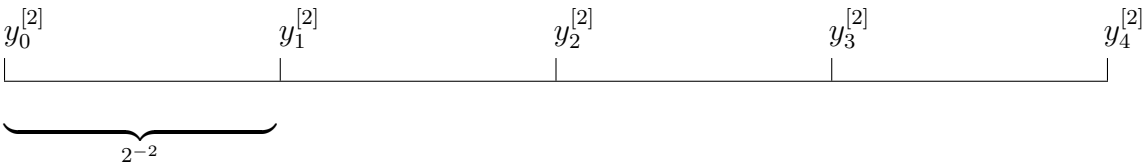
Given the discretization vector $\widetilde{\boldsymbol{y}}^{[d]} = [y_0^{[d]}, \ldots, y_{N_d+1}^{[d]}]^T$ with N_d as discretization number and an equidistant spacing, we can represent the solution for y in space with the following:



If we have for example $d = 3$, we have $N_3 = 2^3 - 1 = 7$. We can represent the solution with such a discretization number by the following:



and for $d = 2$, $N_2 = 2^2 - 1 = 3$, we get the representation



We see that every second element (counting from the first element and onward) of $\widetilde{\boldsymbol{y}}^{[d]}$ corresponds to the elements of $\widetilde{\boldsymbol{y}}^{[d-1]}$ in space. We can now find an expression for the discretization error, by finding the relative error between the solution of every second element of $\widetilde{\boldsymbol{y}}^{[d]}$ and every element of $\widetilde{\boldsymbol{y}}^{[d-1]}$. We use the 2-norm as this is the norm that represents distance in space. We thus get the following expression for discretization error

$$e_y^{[d]} := \frac{\left\| [\, y_0^{[d]}, y_2^{[d]}, \ldots, y_{N_d-2}^{[d]}, y_{N_d}^{[d]} \,] - [\, y_0^{[d-1]}, y_1^{[d-1]}, \ldots, y_{N_{d-1}-1}^{[d-1]}, y_{N_{d-1}}^{[d-1]} \,] \right\|_2}{\left\| [\, y_0^{[d]}, y_2^{[d]}, \ldots, y_{N_d-2}^{[d]}, y_{N_d}^{[d]} \,] \right\|_2} \tag{A.59}$$

We can calculate the discretization error in MATLAB. We have indices $0, \ldots, N+1$ for $\widetilde{\boldsymbol{y}}$, but in MATLAB, indexing starts at 1, so the corresponding vectors will have indices $1, \ldots, N+2$. By using the syntax `1:2:N+2`, we skip every second index. We then find the error by subtracting the solution for N_d from the solution for N_{d-1} in the same place in space. We can implement the error as

```
function [error,d1] = ErrorSolution(yN,yN_2,typeNorm)
% yN   is the solution for N=2^d-1
% yN_2 is the solution for N=2^(d-1)-1
N1 = length(yN)-2;       d1 = log2(N+1);
N2 = length(yN_2)-2;     d2 = log2(N2+1);
% We first see if the vectors are powers of 2 and consecutive
if round(d1) ~= d1 || d2 ~= d1-1
   display('Vectors must be powers of 2 and consecutive')
   return
end
error =...
   norm(( yN(1:2:N1+2) - yN_2(1:N2+2) ),typeNorm)...
   /norm(yN(1:2:N1+2),typeNorm);
end
```

We can now calculate the order for the discretization method, using the 2-norm, with the following code in MATLAB:

```
% yN   is the solution for N=2^d-1
% yN_2 is the solution for N=2^(d-1)-1
% yN_4 is the solution for N=2^(d-2)-1
[errorN  ,d] = ErrorSolution(yN,yN_2,2);
[errorN_2,~] = ErrorSolution(yN_2,yN_4,2);
orderOfError = log2( errorN_2/errorN )

% Then orderOfError is the order of error when we use 2^d-1
% as discretization number
```

Appendix B

On matrices, vectors and specific expressions

B.1 Some operators and notation

In Section B.1, we let $\boldsymbol{x} = [x_1, \ldots, x_N]^T$, $\boldsymbol{y} = [y_1, \ldots, y_N]^T \in \mathbb{R}^N$.

We define the **diag-operator**, $\text{diag}(\boldsymbol{x})$: $\mathbb{R}^N \to \mathbb{R}^{N \times N}$, as

$$\text{diag}(\boldsymbol{x}) := \begin{pmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x_N \end{pmatrix}. \tag{B.1}$$

We notice that the diag()-operator is a linear operator.

> The syntax for diag() in MATLAB is `diag()`, so if we want to diagonalize $\boldsymbol{x}$, that is, find $\text{diag}(\boldsymbol{x})$, we write
> `diag(x)`

We define the **Hadamard product** (elementwise multiplication) as

$$\boldsymbol{x} \circ \boldsymbol{y} := \begin{pmatrix} x_1 y_1 \\ x_2 y_2 \\ \vdots \\ x_N y_N \end{pmatrix}. \tag{B.2}$$

We notice that the Hadamard product is commutative. Also, we have that $\text{diag}(\boldsymbol{x})\boldsymbol{y} = \boldsymbol{x} \circ \boldsymbol{y}$.

> The syntax for Hadamard product in MATLAB is `.*`, so for $\boldsymbol{x} \circ \boldsymbol{y}$, we write `x.*y`

We define the **Hadamard power** (elementwise power) as

$$\boldsymbol{x}^{\circ k} := \begin{pmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_N^k \end{pmatrix}. \tag{B.3}$$

where $k \in \mathbb{R}$. Trivially, $\boldsymbol{x}^{\circ 1} = \boldsymbol{x}$.

> The syntax for Hadamard power in MATLAB is `.^`, so for $\boldsymbol{x}^{\circ k}$, we write `x.^ k`

We define the **Kronecker-delta** tensor as

$$\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & \text{otherwise} \end{cases} \tag{B.4}$$

B.2 Matrices and vectors frequently occurring

In this thesis, the following matrices and vectors frequently occur:

$$S_0 := \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & -2 & 1 \\ 0 & \dots & 0 & 1 & -2 \end{pmatrix}; \; F_0 := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & -1 & 0 & 1 \\ 0 & \dots & 0 & -1 & 0 \end{pmatrix}; \quad \text{(B.5)}$$

$$F_+ := \begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 1 \\ 0 & \dots & 0 & 0 & -1 \end{pmatrix}; \; \hat{F_+} := \left(\begin{array}{c|cc} -1 & 1 & \mathbf{0}_{N-1}^T \\ \hline \mathbf{0}_N & \multicolumn{2}{c}{F_+} \end{array}\right); \quad \text{(B.6)}$$

$$\boldsymbol{r}_{y,S_0} := \begin{pmatrix} y_0 \\ \mathbf{0}_{N-2} \\ y_{N+1} \end{pmatrix}; \; \boldsymbol{r}_{y,F_0} := \begin{pmatrix} -y_0 \\ \mathbf{0}_{N-2} \\ y_{N+1} \end{pmatrix}; \quad \text{(B.7)}$$

$$\boldsymbol{r}_{y,F_+} := \begin{pmatrix} \mathbf{0}_{N-1} \\ y_{N+1} \end{pmatrix}; \; \boldsymbol{r}_{y,\hat{F_+}} := \begin{pmatrix} \mathbf{0}_N \\ y_{N+1} \end{pmatrix}. \quad \text{(B.8)}$$

with $S_0, F_0, F_+ \in \mathbb{R}^{N\times N}$, $\hat{F_+} \in \mathbb{R}^{(N+1)\times(N+1)}$ and with $\boldsymbol{r}_{y,S_0}, \boldsymbol{r}_{y,F_0}$, $\boldsymbol{r}_{y,F_+} \in \mathbb{R}^N$, $\boldsymbol{r}_{y,\hat{F_+}} \in \mathbb{R}^{N+1}$ and where y is defined on $a \le x \le b$ with $y_0 := y(a)$ and $y_{N+1} := y(b)$.

B.3 The Jacobian and derivative of certain expressions

Here, I will give general claims for the Jacobian and derivatives of certain expressions that are noted in the thesis.

In section B.3, we let A, $B \in \mathbb{R}^{N\times N}$ be matrices, $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^N$ be vectors, and let $i, j = 1, \ldots, N$.

Claim B.1. *The Jacobian, $\mathcal{J}$, with respect to $\boldsymbol{x}$ of $A\boldsymbol{x}$ is*

$$\mathcal{J}_{\boldsymbol{x}}^{A\boldsymbol{x}} = A. \tag{B.9}$$

Proof. The i-th row of $A\boldsymbol{x}$ will be

$$[A\boldsymbol{x}]_i = \sum_{k=1}^{N} A_{i,k}x_k.$$

The (i,j)-th element of the Jacobian of $A\boldsymbol{x}$ with respect to $\boldsymbol{x}$ will be

$$\left[\mathcal{J}_{\boldsymbol{x}}^{A\boldsymbol{x}}\right]_{i,j} = \frac{\partial[A\boldsymbol{x}]_i}{\partial x_j} = \sum_{k=1}^{N} A_{i,k}\,\delta_{j,k} = A_{i,j}$$

so we have

$$\mathcal{J}_{x}^{A\boldsymbol{x}} = A.$$

□

Claim B.2. *The Jacobian, $\mathcal{J}$, with respect to $\boldsymbol{y}$ of* $\mathrm{diag}(B\boldsymbol{y})A\boldsymbol{x}$ *is*

$$\mathcal{J}_{\boldsymbol{y}}^{\mathrm{diag}(B\boldsymbol{y})A\boldsymbol{x}} = \mathrm{diag}(A\boldsymbol{x})B. \tag{B.10}$$

Proof. The i-th row of $\text{diag}(B\boldsymbol{y})A\boldsymbol{x}$ will be

$$[\text{diag}(B\boldsymbol{y})A\boldsymbol{x}]_i = \left(\sum_{l=1}^{N} B_{i,l}\, y_l\right)\left(\sum_{m=1}^{N} A_{i,m}x_m\right) = \left(\sum_{m=1}^{N} A_{i,m}x_m\right)\left(\sum_{l=1}^{N} B_{i,l}\, y_l\right).$$

The (i,j)-th element of the Jacobian of $\text{diag}(B\boldsymbol{y})A\boldsymbol{x}$ with respect to $\boldsymbol{y}$ is given by

$$\begin{aligned}\left[\mathcal{J}_{\boldsymbol{y}}^{\text{diag}(B\boldsymbol{y})A\boldsymbol{x}}\right]_{i,j} = \frac{\partial[\text{diag}(B\boldsymbol{y})A\boldsymbol{x}]_i}{\partial y_j} &= \left(\sum_{m=1}^{N} A_{i,m}x_m\right)\left(\sum_{l=1}^{N} B_{i,l}\,\delta_{l,j}\right) \\ &= \left(\sum_{m=1}^{N} A_{i,m}x_m\right) B_{i,j}\end{aligned}$$

so we have

$$\mathcal{J}_{\boldsymbol{y}}^{\text{diag}(B\boldsymbol{y})A\boldsymbol{x}} = \text{diag}(A\boldsymbol{x})B$$

□

Claim B.3. *The Jacobian, $\mathcal{J}$, with respect to $\boldsymbol{y}$ of* $\text{diag}(\hat{A}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}$, *where $\hat{A} \in \mathbb{R}^{(N+1)\times(N+1)}$, $\hat{\boldsymbol{x}} := [x_0, \ldots, n_N]^T$, $\hat{\boldsymbol{y}} := [y_0, \ldots, y_N]^T$ and y_0 is a constant, is*

$$\mathcal{J}_{\boldsymbol{y}}^{\text{diag}(\hat{A}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}} = \text{diag}(\hat{\boldsymbol{x}})\hat{A}_{:,2:N+1}. \tag{B.11}$$

Proof. The i-th, row of $\text{diag}(\hat{A}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}$ will be, for $i = 1, \ldots, N+1$

$$[\text{diag}(\hat{A}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}]_i = \left(\sum_{l=0}^{N} \hat{A}_{i,l+1}\, y_l\right) x_{i-1} = x_{i-1}\left(\sum_{l=0}^{N} \hat{A}_{i,l+1}\, y_l\right).$$

The (i,j)-th element, $i = 1, \ldots, N+1$, $j = 1, \ldots, N$, of the Jacobian of $\text{diag}(\hat{B}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}$ with respect to $\boldsymbol{y}$ is given by

$$\begin{aligned}\left[\mathcal{J}_{\boldsymbol{y}}^{\text{diag}(\hat{A}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}}\right]_{i,j} = \frac{\partial[\text{diag}(\hat{A}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}]_i}{\partial y_j} &= x_{i-1}\left(\sum_{l=0}^{N} \hat{A}_{i,l+1}\delta_{l,j}\right) \\ &= x_{i-1}A_{i,j+1}\end{aligned}$$

so we have

$$\mathcal{J}_{\boldsymbol{y}}^{\mathrm{diag}(\hat{A}\hat{\boldsymbol{y}})\hat{\boldsymbol{x}}} = \mathrm{diag}(\hat{\boldsymbol{x}})\hat{A}_{:,2:N+1}$$

□

Claim B.4. *The Jacobian, $\mathcal{J}$, with respect to $\boldsymbol{x}$ of $\boldsymbol{x}^{\circ k}$ is*

$$\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{x}^{\circ k}} = k\,\mathrm{diag}(\boldsymbol{x}^{\circ(k-1)}). \tag{B.12}$$

Proof. The i-th row of $\boldsymbol{x}^{\circ k}$ will be

$$\left[\boldsymbol{x}^{\circ k}\right]_i = x_i^k.$$

The (i,j)-th element of the Jacobian of $\boldsymbol{x}^{\circ k}$ with respect to $\boldsymbol{x}$ is given by

$$\left[\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{x}^{\circ k}}\right]_{i,j} = \frac{\partial\left[\boldsymbol{x}^{\circ k}\right]_i}{\partial x_j} = k\,x_i^{k-1}\delta_{i,j}$$

so we have

$$\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{x}^{\circ k}} = k\,\mathrm{diag}(\boldsymbol{x}^{\circ(k-1)})$$

□

Corollary B.5. *By Claim B.4, the Jacobian, $\mathcal{J}$, with respect to $\boldsymbol{x}$ of* $\mathrm{diag}(\boldsymbol{x})\boldsymbol{x}$ *is*

$$\mathcal{J}_{\boldsymbol{x}}^{\boldsymbol{x}^{\circ k}} = 2\boldsymbol{x}. \tag{B.13}$$

Proof. We have that $\mathrm{diag}(\boldsymbol{x})\boldsymbol{x} = \boldsymbol{x}^{\circ 2}$. Using $k = 2$ in Claim B.4 proves the corollary. □

Claim B.6. *Let $\boldsymbol{z} = [\boldsymbol{x};\ \boldsymbol{y}]$ with $\boldsymbol{z}_1 := \boldsymbol{x}$ and $\boldsymbol{z}_2 := \boldsymbol{y}$. The Jacobian, $\mathcal{J}$, with*

respect to $\boldsymbol{z}$ *of the block matrix system*

$$\begin{pmatrix} \{A(\boldsymbol{z})\}_{1,1} & \{A(\boldsymbol{z})\}_{1,2} \\ \{A(\boldsymbol{z})\}_{2,1} & \{A(\boldsymbol{z})\}_{2,2} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \tag{B.14}$$

is

$$\mathcal{J} = \begin{pmatrix} \{A(\boldsymbol{z})\}_{1,1} + \frac{\partial \{A(\boldsymbol{z})\}_{1,1}}{\partial \boldsymbol{x}} \boldsymbol{x} + \frac{\partial \{A(\boldsymbol{z})\}_{1,2}}{\partial \boldsymbol{x}} \boldsymbol{y} & \vdots & \{A(\boldsymbol{z})\}_{1,2} + \frac{\partial \{A(\boldsymbol{z})\}_{1,1}}{\partial \boldsymbol{y}} \boldsymbol{x} + \frac{\partial \{A(\boldsymbol{z})\}_{1,2}}{\partial \boldsymbol{y}} \boldsymbol{y} \\ \{A(\boldsymbol{z})\}_{2,1} + \frac{\partial \{A(\boldsymbol{z})\}_{2,1}}{\partial \boldsymbol{x}} \boldsymbol{x} + \frac{\partial \{A(\boldsymbol{z})\}_{2,2}}{\partial \boldsymbol{x}} \boldsymbol{y} & \vdots & \{A(\boldsymbol{z})\}_{2,2} + \frac{\partial \{A(\boldsymbol{z})\}_{2,1}}{\partial \boldsymbol{y}} \boldsymbol{x} + \frac{\partial \{A(\boldsymbol{z})\}_{2,2}}{\partial \boldsymbol{y}} \boldsymbol{y} \end{pmatrix} \tag{B.15}$$

Proof. The vector function of the matrix system will have the following block entry at index l, $\quad l = 1, 2$:

$$\sum_{k=1}^{2} \{A(\boldsymbol{z})\}_{l,k} \boldsymbol{z}_k$$

and hence the Jacobian at block position (l, m), $l, m = 1, 2$, will have the following entry

$$\begin{aligned} \mathcal{J}_{l,m} &= \frac{\partial}{\partial \boldsymbol{z}_m} \sum_{k=1}^{2} \{A(\boldsymbol{z})\}_{l,k} \boldsymbol{z}_k \\ &= \sum_{k=1}^{2} \left(\frac{\partial}{\partial \boldsymbol{z}_m} \{A(\boldsymbol{z})\}_{l,k} \right) \boldsymbol{z}_k + \{A(\boldsymbol{z})\}_{l,k} \frac{\partial}{\partial \boldsymbol{z}_m} \boldsymbol{z}_k \\ &= \sum_{k=1}^{2} \frac{\partial \{A(\boldsymbol{z})\}_{l,k}}{\partial \boldsymbol{z}_m} \boldsymbol{z}_k + \{A(\boldsymbol{z})\}_{l,k} \delta_{m,k}. \end{aligned}$$

which is what we wanted to prove. □

Claim B.7. *Given a matrix,* A*, and a vector,* $\boldsymbol{v}$*, that both depend on a variable vector,* $\boldsymbol{x}$*, then the* j*-th column of the Jacobian of* $A\boldsymbol{v}$ *with respect to* $\boldsymbol{x}$ *will be*

$$\left[\mathcal{J}_{\boldsymbol{x}}^{A\boldsymbol{v}} \right]_{:,j} = \frac{\partial A}{\partial x_j} \boldsymbol{v} + A \frac{\partial \boldsymbol{v}}{\partial x_j}. \tag{B.16}$$

Proof. The Jacobian of $A\boldsymbol{v}$ at position (i, j) is

$$\left[\mathcal{J}_{\boldsymbol{x}}^{A\boldsymbol{v}}\right]_{i,j} = \frac{\partial \left[A\boldsymbol{v}\right]_i}{\partial x_j} = \frac{\partial}{\partial x_j} \sum_{k=1}^{N} A_{i,k}\, v_k = \sum_{k=1}^{N} \frac{\partial A_{i,k}}{\partial x_j}\, v_k + A_{i,k} \frac{\partial v_k}{\partial x_j}.$$

The i-th element of $\frac{\partial A}{\partial x_j}\, \boldsymbol{v} + A\frac{\partial \boldsymbol{v}}{\partial x_j}$ will be

$$\left[\frac{\partial A}{\partial x_j}\, \boldsymbol{v} + A\frac{\partial \boldsymbol{v}}{\partial x_j}\right]_i = \sum_{k=1}^{N} \frac{\partial A_{i,k}}{\partial x_j}\, v_k + A_{i,k} \frac{\partial v_k}{\partial x_j}$$

which is the element of the Jacobian at position (i, j). □

Claim B.8. *The derivative of* $\mathrm{diag}(A\boldsymbol{x})$*, with respect to* x_j *is*

$$\frac{\partial}{\partial x_j} \mathrm{diag}(A\boldsymbol{x}) = \mathrm{diag}(A_{:,j}). \tag{B.17}$$

Proof. The element at index (i, i) of $(\partial/\partial x_j)\, \mathrm{diag}(A\boldsymbol{x})$ is

$$\left[\frac{\partial}{\partial x_j} \mathrm{diag}(A\boldsymbol{x})\right]_{i,i} = \frac{\partial}{\partial x_j} \sum_{k=1}^{N} A_{i,k} x_k = \sum_{k=1}^{N} A_{i,k} \delta_{k,j} = A_{i,j}$$

and the element at index (i, i) of $\mathrm{diag}(A_{:,j})$ is $\left[\mathrm{diag}(A_{:,j})\right]_{i,i} = A_{i,j}$. □

Theorem B.9. This theorem is taken from [1]
The derivative with respect to x *of the inverse of a matrix,* A*, with elements that depend on* x*, is*

$$\frac{\partial A^{-1}}{\partial x} = -A^{-1} \frac{\partial A}{\partial x} A^{-1} \tag{B.18}$$

Proof. The definition of the inverse is $A^{-1}A = I$. We differentiate with respect to x using the product rule and get $\frac{\partial A^{-1}}{\partial x} A + A^{-1} \frac{\partial A}{\partial x} = 0$. By rearranging, we get the desired result. □

Claim B.10. *The derivative with respect to y_i, $i = 0, \ldots, N$, of*

$$f(\hat{\boldsymbol{y}}) = \frac{C + \log(y_0)}{\Delta\, x \left[\frac{1}{2}\left(\frac{1}{y_0} + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}\right]}, \tag{B.19}$$

where $\boldsymbol{y} = [y_1, \ldots, y_N]^T$, $\hat{\boldsymbol{y}} = [y_0, \ldots, y_N]^T$, and y_{N+1} is a given boundary value and C is a constant, is

$$\frac{\partial f(\hat{\boldsymbol{y}})}{\partial y_i} = \begin{cases} \dfrac{\frac{1}{2}\left(\frac{1}{y_0}[1+C+\log(y_0)] + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}}{\Delta\, x\, y_0 \left[\frac{1}{2}\left(\frac{1}{y_0} + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}\right]^2} & , \quad i = 0 \\ \dfrac{C + \log(y_0)}{\Delta\, x\, y_i^2 \left[\frac{1}{2}\left(\frac{1}{y_0} + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}\right]^2} & , \quad i = 1, \ldots, N \end{cases} \tag{B.20}$$

Proof. The derivative of $f(\hat{\boldsymbol{y}})$ with respect to y_0 is

$$\begin{aligned} \frac{\partial f(\hat{\boldsymbol{y}})}{\partial y_0} =& \frac{\Delta\, x \left[\frac{1}{2}\left(\frac{1}{y_0} + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}\right] \frac{\partial}{\partial y_0}\left[C + \log(y_0)\right]}{\left\{\Delta\, x \left[\frac{1}{2}\left(\frac{1}{y_0} + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}\right]\right\}^2} \\ & - \frac{\left[C + \log(y_0)\right] \frac{\partial}{\partial y_0} \Delta\, x \left[\frac{1}{2}\left(\frac{1}{y_0} + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}\right]}{\left\{\Delta\, x \left[\frac{1}{2}\left(\frac{1}{y_0} + \frac{1}{y_{N+1}}\right) + \mathbf{1}_N^T \boldsymbol{y}^{\circ(-1)}\right]\right\}^2} \end{aligned}$$

$$
\begin{aligned}
=&\frac{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]\frac{1}{y_0}-[C+\log(y_0)]\,\Delta\, x\left[-\frac{1}{2}\left(\frac{1}{y_0^2}\right)\right]}{\left\{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]\right\}^2}\\
=&\frac{\frac{1}{y_0}\left\{\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]+[C+\log(y_0)]\left(\frac{1}{2y_0}\right)\right\}}{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2}\\
=&\frac{\frac{1}{2}\left(\frac{1}{y_0}\left[1+C+\log(y_0)\right]+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}}{\Delta\, x\, y_0\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2}.
\end{aligned}
$$

Next, the derivative of $f(\hat{\boldsymbol{y}})$ with respect to y_i, $i=1,\ldots,N$, is

$$
\begin{aligned}
\frac{\partial f(\hat{\boldsymbol{y}})}{\partial y_i}&=-\frac{[C+\log(y_0)]\,\frac{\partial}{\partial y_i}\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]}{\left\{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]\right\}^2}\\
&=\frac{[C+\log(y_0)]\,y_i^{-2}}{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2}\\
&=\frac{C+\log(y_0)}{\Delta\, x\, y_i^2\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2}.
\end{aligned}
$$

□

Corollary B.11. *The Jacobian, $\mathcal{J}$, with respect to $\hat{\boldsymbol{y}}$, of*

$$
f(\hat{\boldsymbol{y}})=\frac{C+\log(y_0)}{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]} \tag{B.21}
$$

is

$$
\left(\begin{array}{cc}
\frac{\frac{1}{2}\left(\frac{1}{y_0}[1+C+\log(y_0)]+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}}{\Delta\, x\, y_0\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2} & \frac{C+\log(y_0)}{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2}\left(\boldsymbol{y}^{\circ(-2)}\right)^T
\end{array}\right) \tag{B.22}
$$

Proof. The $(1,j)-th$ element of the Jacobian of $f(\hat{\boldsymbol{y}})$ with respect to $\hat{\boldsymbol{y}}$ is given by

$$\left[\mathcal{J}_{\hat{\boldsymbol{y}}}^{f(\hat{\boldsymbol{y}})}\right]_{1,j} = \frac{\partial f(\hat{\boldsymbol{y}})}{\partial y_j}$$

By Claim B.10, the derivative of $f(\hat{\boldsymbol{y}})$ with respect to y_0, is the first element in expression (B.22), and thus proves the first expressions of the Jacobian. Again, using Claim B.10, the next N elements of the Jacobian are

$$\left(\begin{array}{ccc} \frac{C+\log(y_0)}{\Delta\, x\, y_1^2\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2} & \cdots & \frac{C+\log(y_0)}{\Delta\, x\, y_N^2\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2} \end{array} \right)$$

$$= \frac{C+\log(y_0)}{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2} \left(\begin{array}{ccc} y_1^{-2} & \cdots & y_N^{-2} \end{array} \right)$$

$$= \frac{C+\log(y_0)}{\Delta\, x\left[\frac{1}{2}\left(\frac{1}{y_0}+\frac{1}{y_{N+1}}\right)+\mathbf{1}_N^T\boldsymbol{y}^{\circ(-1)}\right]^2} \left(\boldsymbol{y}^{\circ(-2)}\right)^T$$

□

B.4 Matrix and vector norms

There are some norm identities that are useful in the analysis of the Jacobian of a fixed-point method. For this section, we let $A, B \in \mathbb{R}^{N\times N}$ be matrices and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ be vectors.

The following inequalities hold for matrix/vector norms:

$$\|A+B\| \leq \|A\| + \|B\|; \qquad \|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \tag{B.23a}$$

$$\|AB\| \leq \|A\|\|B\|; \qquad \left|\boldsymbol{x}^T\boldsymbol{y}\right| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\| \tag{B.23b}$$

$$\|A\boldsymbol{x}\| \leq \|A\|\|\boldsymbol{x}\| \tag{B.23c}$$

Lemma B.12. *Given a vector $\boldsymbol{x} \in \mathbb{R}^N$, the following norm identity holds:*

$$\|\mathrm{diag}(\boldsymbol{x})\|_1 = \|\boldsymbol{x}\|_\infty \tag{B.24}$$

Proof. The definition of the matrix 1-norm of $\|\mathrm{diag}(\boldsymbol{x})\|_1$ is

$$\|\mathrm{diag}(\boldsymbol{x})\|_1 := \max_{i=1,\ldots,N} \sum_{j=1}^{N} |[\mathrm{diag}(\boldsymbol{x})]_{i,j}| = \max_{i=1,\ldots,N} |[\mathrm{diag}(\boldsymbol{x})]_{i,}| \tag{B.25}$$

$$= \max_{i=1,\ldots,N} |\boldsymbol{x}| =: \|\boldsymbol{x}\|_\infty \tag{B.26}$$

□

Appendix C

List of physical terms

Name	Explanation
band	
conduction -	The energy range for a free charge carrier
energy -	Ranges of energy that an electron within a solid may or may not have
- gap	The range between two energy bands
valence -	The highest energy a charge carrier can have without being free
bias	
forward-	When the electric field caused by the applied voltage has the same direction as the electric field caused by the built-in potential
reverse-	When the electric field caused by the applied voltage has the opposite direction as the electric field caused by the built-in potential
charge carrier	A particle with an electric charge that can conduct electricity; for this thesis, they are electrons and holes

Name	Explanation
conductor	A material that conducts electric current well, i.e. a material with a low resistivity
DA-interface	The interface between the donor and acceptor material in an organic diode
Debye length	The distance over which significant charge separation can occur
depletion region	An insulating region between two doped semiconductor materials where the charge carriers have been diffused away, or have been forced away by an electric field
diode	A two-terminal electronic component that conducts electricity primarily in one direction
dopant	The material which is added to the semiconductor to dope the material
effective mass	The mass that an object seems to have when responding to forces
electron	A negatively charged charge carrier
- affinity	The energy released when an electron is added to a material
electronic band structure	The structure that describes the different energy levels in which the electrons can exist
emitter	Another name for n-type semiconductor material in a solar cell
exciton	An electron-hole pair that is bound by Coloumb attraction
Fermi level	The highest energy level for electrons at aboslute zero temperature
hole	A lack of a hole and a positively charged charge carrier

Name	Explanation
insulator	A material that conducts electric current poorly, i.e. a material with high resistivity
ionization potential	The energy necessary to remove an electron from a neutral atom
mean free time	The average time between each collision of a charge carrier
metal	A material with a band gap of zero, and which is a conductor
mobility	Describes how strongly the motion of an electron/hole is influenced by an applied electric field
nucleus	The dense region of protons and neutrons in the center of the atom
organic compound	Virtually any chemical compound that contains carbon
p-n junction	The interface between the p-doped and n-doped material in a semiconductor diode
permittivity	
- of free space	It is the capability of the vacuum to permit electric field lines
dielectric -	The measure of resistance that is encountered when forming an electric field in a particular material
photovoltaic	An adjective that indicates that something converts light into electricity. I use the term "solar" instead of "photovoltaic" in this thesis
polymer	A molecule composed of many repeated subunits
semiconductor	A material which has conductive properties between that of an insulator and a conductor

Name	Explanation
doped -	A semiconductor with different electrical properties than the original semiconductor because a material that has different chemical properties has been added
extrinsic -	A doped semiconductor
intrinsic -	A non-doped semiconductor
n-doped -	A doped semiconductor that has more electrons per atom than the intrinsic type of the material
p-doped -	A doped semiconductor that has more holes per atom than the intrinsic type of the material
shell	The different shells of energy levels of electrons in an atom
valence -	The outermost electron shell of the atom
silicon	An element which is a semiconductor and has symbol chemical Si
solar	
- cell	An electronic device that converts sunlight into electricity
- panel	Several solar cells make up one solar panel
space charge density	The sum of charge carrier densities
thermal voltage	The relationship between the flow of electric current and the electrostatic potential across a p–n junction
velocity	
average thermal -	The average velocity of a charge carrier has due to thermal motion
drift -	The average velocity of a charge carrier has due to an electric field

Appendix D

Nomenclature

Symbol	Explanation
$\hat{E}$	Electric field
$\boldsymbol{f}.$	The function for which we use Newton-Raphson to find the roots
F_0	The matrix that represents the finite central difference for first order derivatives
F_+	The matrix that represents the finite forward difference for first order derivatives
$G.$	The block matrix we use when solving with Newton-Raphson
h	Planck's constant
$\boldsymbol{h}.$	The mapping function we have for the fixed-point method
J	Electron or hole flux
$\mathcal{J}$	Jacobi matrix
k_B	Boltzmann constant
L	Thickness of glass for the silicon solar cell case, thickness of acceptor layer in the organic solar cell case
L_D	Debye length when considering a silicon solar cell

Symbol	Explanation
M	Thickness of donor layer in the organic solar cell
m_n	Effective electron mass
m_p	Effective hole mass
n	Electron concentration
$\boldsymbol{n}$	Discretization vector for electron concentration in between the boundaries in the domain
$\hat{\boldsymbol{n}}$	Discretization vector for electron concentration in the entire domain excluding the right boundary value
$\widetilde{\boldsymbol{n}}$	Discretization vector for electron concentration in the entire domain
N	Discretization number
p	Hole concentration
p_0	Discretization variable for left boundary value for hole concentration
q	Electron/hole charge
$\boldsymbol{r}_{\bullet,\bullet}$	Boundary value vector
S_0	The matrix that represents the finite central difference for second order derivatives
t	Time variable
T	Temperature
v_n	Electron drift velocity
v_p	Hole drift velocity
x	Spatial variable
$\boldsymbol{x}$	Discretization vector for spatial variable in between the boundaries in the domain
$\widetilde{\boldsymbol{x}}$	Discretization vector for spatial variable in the entire domain

Symbol	Explanation
$\boldsymbol{z}$	Variable used when solving systems of equations with Newton-Raphson
α	Ratio between left and right boundary condition for n when considering a silicon solar cell
δ	Ratio of recombination rate to diffusion rate
$\delta_{\bullet,\bullet}$	Kronecker-Delta tensor
$\Delta\, x$	Spatial step length
ϵ	Dielectric permittivity
θ	Parameter for organic solar cell depending on cell properties and temperature when considering an organic solar cell
λ	Parameter for organic solar cell depending on cell properties and temperature when considering an organic solar cell
μ_n	Electron mobility
μ_p	Hole mobility
ν	Ratio between Debye length, L_D, and the glass thickness, L, when considering a silicon solar cell
ρ	The space charge density
τ	Mean free time
ϕ	Electrostatic potential
$\boldsymbol{\phi}$	Discretization vector for electrostatic potential in between the boundaries in the domain
$\hat{\boldsymbol{\phi}}$	Discretization vector for electrostatic potential in the entire domain excluding the right boundary value
$\widetilde{\boldsymbol{\phi}}$	Discretization vector for electrostatic potential in the entire domain
$\boldsymbol{\varphi}$	Applied electric potential when considering a silicon solar cell

Symbol	Explanation
Φ	Applied electric potential when considering an organic solar cell
Φ_{bi}	Built in electric potential when considering an organic solar cell

Appendix E

Values of physical constants

Symbol	Numerical value	Unit
h	6.63×10^{-34}	J s
k_B	1.38×10^{-23}	J K^{-1}
m_e	9.11×10^{-31}	kg
q	1.60×10^{-19}	C

Appendix F

The excursions to Ghana

In my master program at the Department of Mathematics at the University of Bergen, I have had the opportunity to travel to National Insitute of Mathematical Science (NIMS) at Kwane Nkrumah University og Science And Technology (KNUST) in Kumasi, Ghana. NIMS has over a 5 year period in part been funded by the Petroleum Geo Services through the Norwegian Academy of Science and Letters. With funding from BKK, a local power company in Bergen, we visited NIMS two times as a part of the COMPCON project, a collaboration between the Department of Mathematics at the University of Bergen and NIMS in Kumasi. The project aims to initiate research within solar energy, and for use in Ghana in particular.

(a) Photo from the KNUST campus.

(b) From the KNUST campus. On the right are LED street lights which consume less energy than traditional light bulbs.

Figure F.1

At the Department of Engineering at KNUST, they have a research project on solar panels with solar panels and devices that measure irradiance and air temperature. Several different types of solar modules are installed, both different types of silicon solar panels and different types of non silicon solar panels. The measurements from these devices are saved, to see if the energy output is according to the expected output for each particular type of panel, and for other research purposes.

Figure F.2: The different solar installation at at the Department of Engineering at KNUST.

The first excursion to NIMS was in January 2017, with Professor Tor Sørevik, Einar Eimhjellen and I. In addition to conduct research on solar cell theory and statistical data, we visited a 20MW solar park in Onyandze close to Winneba, Ghana. The solar park was installed and operated by the Chinese power company BXC. All the the panels in the park were imported from China and were cheaper than many other options. The low prize was in part due to that fact that the panels were produced in a manner that is more polluting that was allowed in EU, in terms of both production and use, that is, solar panels can not be produced in such a manner in the EU, and these panels were not allowed to be imported into the EU. We visited the park during the Hamattan, a season in Sub-Saharan Africa in which there is dust from the Sahara desert in the atmosphere which blocks much of the radiance, and also settles on the ground, and hence on the panels, to further decrease the irradiance on the panels.

(a) (b)

Figure F.3: Solar panels in the solar park in Onyandze covered with dust during Hamattan.

The second excursion to NIMS was in May 2018, with Professor Tor Sørevik, Einar Eimhjellen, Benjamin Widerøe, Karoline Lekve and I. During this visit, Einar and I presented the results from our master project, and Benjamin and Karoline presented their outlook for their master projects. We also had the opportunity to

(a) A microgrid installation in Kofihwikrom, with the local medical clinic in the background.

(b) The power converter room with batteries for electrical storage.

Figure F.4

visit two microgrid solar panel installations, which are small independent power installation that has solar panels that produce electrical energy and this power

is sent only to a few houses that are close to the source, so that less infrastructure is needed. This can be very useful for small rural villages. In particular we visited two installations close to Kumasi that were both installed by Black Star Energy, who provide energy to 15 local villages in total. We visited one installation in Kofihwikrom which provided energy to 45 houses and one local medical clinic. We also visited another installation in Daban which provided energy for 62 houses. They have installed batteries in the power converter room, so that they could deliver power also at night. The systems are so reliable that the people in these villages have a more reliable power source than the people connected to the state owned grid.

Ghana is a country with around 29 million inhabitants and has a dire and growing need for electric power. During the first visit we had frequent power flickers. We were told not to leave the air-condition on while we were out of the room, and not to leave the water heater on, unless we needed warm water right now. During our second visit, we had to change plans for accommodation, because the guest house we had planned to stay at had a power outage that they were unable to fix. Also, at another guest house, we had several power outages. Ghana has an outdated power production and power grid infrastructure. The power production can be updated with more solar parks, both parks that send electricity to the main grid, and with microgrids that need less infrastructure.

(a) LED street lights powered by solar panels along a main road in Ghana.

(b) Drying of cocoa beans in Kofihwikrom. Most of the houses in Kofihwikrom and Daban are clay huts.

Figure F.5

(a) Group Photo from the second excursion. In the back from the left: Dr. Peter, Benjamin, Isaiah, Michael, Einar, Anas. The middle row from the left: Patricia, Richard, Daniel (author), Joshua, Kofi. In the front from the left: Reindolph, Phebe, Karoline, Frank, Esi, Isaac, Prof. Tor.

(b) Einar and Hisham taking a closer look at the solar panels in the solar park in Onyandze.

Figure F.6

(a) Selling products along the streets in Ghana is common.

(b) Photo from Accra, a city in fast growth and in need of an increasing amount of electric energy.

Figure F.7