Numerical Simulation of a Laterally Confined Double Dot with Tunable Interaction Potential

by

Aaron David Kiyoshi Finck

Submitted to the Department of Physics in partial fulfillment of the requirements for the degree of

BACHELOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2005

©2005 Aaron Finck All Rights Reserved

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author	
	Department of Physics
	May 6, 2005
Certified by	
	Raymond C. Ashoori
	Professor, Department of Physics
	Thesis Supervisor
Accepted by	
	Professor David E. Pritchard
	Senior Thesis Coordinator, Department of Physics

Numerical Simulation of a Laterally Confined Double Dot with Tunable Interaction Potential

by

Aaron David Kiyoshi Finck

Submitted to the Department of Physics on May 6, 2005, in partial fulfillment of the requirements for the degree of BACHELOR OF SCIENCE

Abstract

Recent technological advances have allowed for the construction of small (on the order of 100-1000 nm) systems of confined electrons called quantum dots. Often kept within semiconductor heterostructures, these systems are small enough that the electrons within them occupy states with discrete energy levels. Two single quantum dots can be placed next to each other so as to form a double dot, with a host of special properties. Such properties can be probed if one can design the semiconductor heterostructure containing the double dot so that experimenters can tune the confinement potential of the double dot. To assist in the testing of heterostructures before their actual construction, we have created a numerical simulation program that calculates the electrostatic potential and charge density for a quantum double dot housed in a semiconductor heterostructure. Relaxation techniques were used to solve Poisson's equation for the heterostructure. The Thomas-Fermi approximation was used to calculate the electron density as a function of the spatially varying electrostatic potential. Certain parameters of the simulation, such as the doping density of the semiconductor material and the electron effective mass, were chosen after trial and error such that the behavior of the simulation matched experimentally observable values, including the electron density within and outside of the quantum double dot region and the depletion voltage. Once the free parameters were chosen, the simulation was used to examine various heterostructure geometries with gates that could tune the confinement potential of the double dot. For example, we analyzed a pincher gate geometry that could split the double dot into two isolated quantum single dots by increasing the potential barrier in the channel connecting the two lobes of the double dot. We discovered that such a tunable dot requires a pincher gate that runs continuously across the bottom of the top gate. We also analyzed how two middle gates, each positioned directly above one of the two lobes of the double dot, can shape the relative sizes of the two connected dots.

Thesis Supervisor: Raymond C. Ashoori

Title: Professor, Department of Physics

Contents

1	Inti	roduction	11		
2	Lat	erally Coupled Vertical Double Dot System	15		
	2.1	Vertical Confinement: Band Offsets	16		
	2.2	Geometry of the Semiconductor Heterostructure	18		
	2.3	Double Dot Shape and Coupling Strength	20		
	2.4	Pincher Gate Geometry	21		
3	Tec	hnical Details and Physics of the Simulation	23		
	3.1	Initialization	24		
	3.2	Relaxation	25		
		3.2.1 Using Poisson's Equation	25		
		3.2.2 Overrelaxation	28		
	3.3	Semiconductor Charge Density and Electrostatic Potential	30		
	3.4	Band Offsets	31		
3.5 Reporting Output		Reporting Output	33		
		3.5.1 Sample Output	33		
4	Tes	ting of Parameters	39		
5	Pincher Gate and Middle Gate Simulations: Separating the Double				
	Dot	,	43		
	5.1	Side Gate Geometry	43		
	5.2	Pincher Gate Geometry	44		

	5.2.1	Middle Gates	 	 	 	 	. 46
6 C	Conclusio	n					53

List of Figures

1-1	Diagram of quantum dot used in gated transport spectroscopy. Repro-	
	duced from a figure in [1]. \ldots \ldots \ldots \ldots \ldots \ldots	12
2-1	Diagram of vertically coupled quantum double dot	16
2-2	Diagram of laterally coupled quantum double dot	16
2-3	Diagram of valence and conduction bands of two different semicon-	
	ductors, both while not in contact and in contact with each other.	
	Reproduced from Figure 17 of Chapter 19 in [13]	17
2-4	Diagram of jump in electrostatic potential across an interface between	
	two different semiconductors. Excess negative and positive charges are	
	located on opposite sides of the interface. Adopted from Figure 13 of	
	Chapter 19 in [13]	18
2-5	Cross-section (yz plane) of semiconductor heterostructure. \ldots .	19
2-6	Cross-section (xy plane) of cap layer.	20
2-7	Diagram of upper gates.	22
3-1	Flow chart of the simulation	24
3-2	Central and adjacent grid points, plus cubic cell	26
3-3	Error plotted versus relaxation loop iteration number for a fully three-	
	dimensional geometry.	29
3-4	Electron density as a function of electrostatic potential (with reduced	
	mass = $0.001m_e$ and $\rho_0 = 1.55 * 10^{17} e \text{ per cm}^3$)	31
3-5	Sample plot of two-dimensional lateral confinement potential on a	
	plane within the quantum well layer.	32

3-6	Sample plot of two-dimensional lateral confinement potential on a	
	plane within the quantum well layer.	33
3-7	Sample contour plot of two-dimensional electron density within the	
	quantum well layer. Electron density is in units of e per nm ²	34
3-8	Potential energy plotted versus the Z coordinate (along the center of	
	the geometry). The bottom metal gate is located at $z = 0$ while the	
	top metal gate is located at $z = 200$ nm	35
3-9	Confinement potential plotted versus the X coordinate within the quan-	
	tum well region, for various potential biases applied to the top gate.	
	The listed potential values are equal to the absolute value of the (elec-	
	tron repulsing) top gate potential bias.	36
3-10	Electron density plotted versus the X coordinate within the quantum	
	well region, for various potential biases applied to the top gate. The	
	listed potential values are equal to the absolute value of the (electron	
	repulsing) top gate potential bias	37
3-11	Electron density versus X coordinate plots for single dots of various	
	cap layer radii. For each radius, the electron density is plotted for a	
	set of different voltages applied to the top gate	38
4-1	Electron density far from the quantum dot region versus doping density	
	$ \rho_0. \ldots \ldots$	40
4-2	Electron density within the quantum dot region versus doping density	
	$ \rho_0. \ldots \ldots$	40
4-3	Electron density within the quantum dot region versus electron effec-	
	tive mass	41
5-1	Diagram of side gates.	44
5-2	Electron density contour plot in quantum well layer with original side	
	gate-only geometry (with side gate potential = 50 V)	45
5-3	Electron density contour plot in quantum well layer with pincher gate	
	setup	46

5-4	Zoom-in of contour plots for quantum well electron density for various	
	values of pincher gate voltage bias.	47
5-5	Electron density $\rho(x)$ for various middle gate potential bias values	49
5-6	Zoom-in of contour plots for quantum well electron density for various	
	values of first middle gate potential bias $U_{middlegate1}$	50
5-7	Cartoon of energy levels for a small dot (with widely spaced energy	
	levels) coupled with a large dot (with closely spaced energy levels).	51

Chapter 1

Introduction

Recent advances in semiconductor fabrication technology, especially in terms of fabrication resolution of small structures [1], have allowed for the construction of systems of artificially confined electrons, often called quantum dots or "artificial atoms" [3, 12]. Physicists have been interested in studying the electronic structure of such systems, which are so small that the Fermi wavelength of the confined electrons is comparable to the spatial size of the confinement potential.

There exists a variety of quantum dots types. For example, one common way of fabricating quantum dots is to create lateral confinement of electrons in a 2DEG (2-dimensional electron gas, in which electrons are confined in the z direction due to the band offsets at semiconductor interfaces) within a semiconductor heterostructure (such as GaAs/GaAlAs) [16]. Additionally, heterostructures can be designed so as to produce quantum dots in close proximity of one another, resulting in so-called double dots or quantum dot molecules. For example, Reed et al. vertically coupled two quantum dots together by placing a tunneling barrier between two 2DEG layers [12]. Such vertically coupled dots were also analyzed theoretically by Palacios and Hawrylak [6]. Coupling and tunneling effects were found to produce behavior not observed in isolated quantum dots, including electron ground states that resemble those in diatomic molecules [12]. Furthermore, larger arrays of coupled quantum dots could be used to study solid-state physics, including energy gaps and energy bands, on a lower energy scale than found in atomic crystals [8, 4]. While vertically coupled quantum dots have been the focus of many experimental and theoretical investigations [9], the goal of the program is to simulate relatively large, laterally coupled quantum dots (i.e. quantum dots that exist within the same 2DEG layer of a heterostructure). Also, unlike other quantum dot heterostructures made by etching metal gates with electron beam pattern lithography, the focus of our simulation is a heterostructure with a pillar etched on the top of the structure, which induces a confinement potential within a 2DEG layer. The former type of quantum dot, pictured in figure 1-1, is primarily analyzed through transport experiments (also known as gated transport spectroscopy), in which current is run through the dot and electrons both enter and leave it [1]. The tunneling properties of double dots constructed in this fashion are examined analytically by Matveev et al. in [5]. In the latter type, which we are studying, we create a dot that is in equilibrium with its surroundings and can be studied via capacitance spectroscopy.



Figure 1-1: Diagram of quantum dot used in gated transport spectroscopy. Reproduced from a figure in [1].

In order to design experiments to explore quantum dots, one often resorts to computer simulations to predict how the dots form in various geometries and under numerous conditions. Thus, we have created a program to numerically calculate the electrostatic potential and electron density within a semiconductor heterostructure containing a quantum double dot. The electrostatic potential is determined by Poisson's equation, which is numerically solved through a relaxation method [15]. For the simulation, Poisson's equation is nonlinear due to the inclusion of the electron density term $\rho_e(U)$, which is a function of the electrostatic potential U. This relation between electron density and electrostatic potential is derived from the Thomas-Fermi approximation [13]. The approximation assumes that the electrons form a gas of noninteracting fermions with T = 0 and constant chemical potential $\mu(\vec{r})$ throughout the heterostructure. Also, we assume that the electrostatic potential does not vary appreciably over length scales equal to the Fermi wavelength. The restriction that $\mu(\vec{r}) = \epsilon_F(\vec{r}) - eU(\vec{r})$ is constant for all \vec{r} as well as the fact that the Fermi energy $\epsilon_F(\vec{r})$ is proportional to $\rho_e^{\frac{2}{3}}$ (where ρ_e is the electron density) imply a particular relation between the electron density and the electrostatic potential $U(\vec{r})$.

After testing the simulation and choosing the proper values for the free parameters (such as the electron effective mass) to help ensure that the simulation produced results that matched a previously defined set of experimental observations, the program was used to test out possible heterostructure geometries to determine if any could allow experimenters to control the electrostatic potential in the channel connecting the two lobes of the quantum double dot and thus separate the double dot into two single dots. We discovered that a pincher gate that ran continuously over the region where the two dots intersected could produce a tunable lateral interaction. Furthermore, we analyzed a geometry with two middle gates, each positioned on top of one of the two lobes of the cap layer. We found that the two middle gates, by applying unequal potentials to the two of them, could shrink the size of one of the dots of the double dot. Therefore, the middle gates, combined with the pincher gate, can tune the relative sizes of the two dots and experimentally analyze how the two couple.

Chapter 2

Laterally Coupled Vertical Double Dot System

The double dot system of interest is created within a GaAs/AlGaAs semiconductor heterostructure, with a pillar etched on top to produce the confinement potential within the 2DEG layer (also referred to as the quantum well layer). As opposed to the vertically coupled double dots (seen in figure 2-1) analyzed by Reed [7] (see also the work by Tarucha in [9], our double dot is composed of two laterally coupled quantum dots (seen in figure 2-2), both residing within the same quantum well layer in the heterostructure. Vertical confinement of electrons (i.e. confinement of electrons within the "plane" of the quantum well layer) is achieved by the band offsets between GaAs and AlGaAs layers, which will be described in greater detail below. Lateral confinement is provided by having short pillars etched on top of the heterostructure into the face of the top metal gate. Because of these pillars (to which we also refer as the cap layer), the region of the quantum well layer directly beneath the cap layer is farther from the top metal gate than other regions are (the increase in distance being equal to the height of the pillars). Thus, this region underneath the cap layer feels less of the electrostatic influence of the top metal gate and has more favorable electrostatic potential values for electrons, causing the electrons in the quantum well layer to be corralled into this region.



Figure 2-1: Diagram of vertically coupled quantum double dot.



Figure 2-2: Diagram of laterally coupled quantum double dot.

2.1 Vertical Confinement: Band Offsets

Vertical confinement of the electrons within the quantum well region is carried out via band offsets between two different semiconductor layers (intrinsic GaAs for the quantum well layer and AlGaAs or GaAs/AlGaAs for two adjacent blocking barrier layers). Due to the periodic potential from a crystal lattice, semiconductor materials have an energy gap between their valence band (the band of highest energy states that are occupied by electrons) and their conduction band (the band of lowest energy states that are unoccupied), with the Fermi level (which is closely approximated by the chemical potential) located roughly half-way in between the two. Separated from each other, the two different types of semiconductors have different Fermi levels. When the two layers are placed adjacent to each other, electrons from the layer with the higher Fermi level will jump to the lower energy states in the other layer. The change in electron density will alter the two Fermi levels until they are equal, at which point there is not enough of an energy difference for electrons to move from one layer to another [13, 14]. See figure 2-3 (which contains a picture of the valence and conduction bands for the two semiconductors for before and after they come into contact) for this band bending as well as the charge accumulation near the interface. The imbalance of charge, however, creates an electric field across the interface, leading to a jump in the electrostatic potential at the interface between the two semiconductor layers. This jump (pictured in figure 2-4) is called the band offset. For our particular case, we use GaAs and AlGaAs, more specifically called $Al_xGa_(1 - x)As$, with x = 0.27. The AlGaAs material is produced by replacing some of the gallium atoms in GaAs with aluminum atoms, thus changing the periodic lattice potential felt by the electrons, leading to a shifted, wider band gap as compared to the band gap of GaAs.



Figure 2-3: Diagram of valence and conduction bands of two different semiconductors, both while not in contact and in contact with each other. Reproduced from Figure 17 of Chapter 19 in [13].



Figure 2-4: Diagram of jump in electrostatic potential across an interface between two different semiconductors. Excess negative and positive charges are located on opposite sides of the interface. Adopted from Figure 13 of Chapter 19 in [13].

2.2 Geometry of the Semiconductor Heterostructure

The geometry for the semiconductor heterostructure containing the quantum dot (pictured in figure 2-5) consists of a number of layers of different materials stacked on top of one another, sandwiched between two metal electrode plates. The structure is grown through molecular beam epitaxy, which produces high quality crystal layers. In the experimental heterostructure, the base consists of an intrinsic GaAs substrate with a layer of n+ GaAs conducting substrate. Above this layer is a thin spacer layer, to prevent the diffusion of dopants into the quantum well. In the simulation, these layers are modeled as a layer of insulator 10 nm thick on top of a bottom metal gate (which has constant potential). On top of the insulator layer is a 80 nm thick region of intrinsic GaAs superlattice tunnel barrier material (which is thin enough to allow electrons to tunnel from the quantum well layer into the bottom gate), which is modeled as a semiconductor but with a band offset between it and the intrinsic

GaAs layer. On top of the tunnel barrier is a 18 nm thick layer of intrinsic GaAs. This layer (also known as the quantum well layer) contains the quantum dot. Vertical confinement is provided by the band offsets caused by the sandwiching of the intrinsic GaAs in between the lower GaAs/AlGaAs superlattice tunnel barrier and an upper AlGaAs semiconductor blocking barrier, which is thick enough to prevent tunneling between the quantum well layer and the top gate. This 50 nm thick upper blocking barrier is a layer of AsSi delta doping material that is located 20 nm above the quantum well region. In the simulation, this delta doping layer is modeled as a thin insulator layer with constant, uniform positive charge density (about 10^{17} e per cm³).



Figure 2-5: Cross-section (yz plane) of semiconductor heterostructure.

On top of the upper blocking barrier is the top metal gate. This gate controls the upper potential. A double cap region (which takes the shape of two parallel, partially intersecting pillars, each having a radius on order of 250 nm) is carved into the bottom face of the top metal gate, with a depth of 30 nm. Within the cap layer is intrinsic GaAs. A cross-section of the cap layer (illustrating its characteristic shape in a given xy plane) can be seen in figure 2-6. Because of the cap layer, different points in the quantum well layer have different distances from the top metal gate. The regions in the quantum well layer that fall under the cap layer are farther away from the top metal gate than those that do not fall directly under the cap layer. Thus, the regions that are located directly under the cap layer feel less of the top metal gate's electron repelling effects and will therefore have higher electrostatic potential (i.e. more attractive to negatively charged electrons) than regions that are not under the cap layer. Ultimately, the double-dot shaped cap layer will produce a double-dot shaped blob of charge within the quantum well due to the more favorable electrostatic potential there.



Figure 2-6: Cross-section (xy plane) of cap layer.

2.3 Double Dot Shape and Coupling Strength

Because the two pillars intersect partially with each other to form a double pillar with two lobes, the electrons are confined into a double dot shape, with each of the two lobes directly under one of the two pillars. Ideally, we would like to find a way to separate the double dot into two distinct single quantum dots and also control the strength of the coupling between the two dots by altering the electrostatic potential in the region between the two single dots. This in turn will aid in experiments using single electron capacitance spectroscopy to analyze localized and delocalized states of electrons in the double dot (see [18]). It is expected that under certain conditions, the electron states can become delocalized and form symmetric and anti-symmetric states that span across both dots. The coupling strength of the two double dots, which depends on how much the two dots intersect as well as the strength of the electrostatic potential in the region of intersection, determines the energy splitting of the symmetric and anti-symmetric states [9]. The strength of the coupling or separation of the two dots can be potentially achieved by placing additional metal gates in and around the pillars. By applying a potential bias between these additional metal electrodes and the top gate, one can alter the electrostatic potential within the quantum well region and change the confinement potential of the quantum dot. As we shall see, only the areas in the quantum well directly beneath the relevant gates significantly feel the electrostatic influence of these gates. Thus, in order to control the potential in the area joining the two lobes of the double dot, one must employ a pincher gate which runs along the top gate and the top of the pillars, so as to be directly above the region of the intersection of the two pillars.

2.4 Pincher Gate Geometry

To allow for greater control of the confinement potential and electron density within the double dot, additional gates are located on the undersurface of the top metal gate (see figure 2-7). These gates are metal electrodes that can have different potentials than the top metal gate. The first additional gate is the 200 nm wide pincher gate, which runs across the center of the underside of the top metal gate as well as the top part of the cap layer. Because the pincher gate lies directly above the overlap region of the two intersecting quantum dots, applying a negative (unfavorable for electrons) electrostatic potential to the pincher gate will drive out the electron density in the overlap region. The other two additional gates are called the middle gates. They are located on top of the cap layer, one for each lobe of the cap layer. The potentials of these middle gates can be adjusted so as to vary the depths of the confinement potential for the two conjoined quantum dots, separately. As it will be described below, varying the potential of one middle gate only affects the potential and electron density for the quantum dot directly beneath that middle gate; the other quantum dot is mostly unaffected.



Figure 2-7: Diagram of upper gates.

Chapter 3

Technical Details and Physics of the Simulation

Within this chapter, we will discus the basic flow of the program, as well as describing the physics behind the simulation. For example, we will describe the mathematics and physics behind the technique of relaxation as well as the Thomas-Fermi approximation and band offsets, in addition to describing how they are implemented within the simulation. Furthermore, we will show some examples of the output of the simulation as well note how the size of the quantum dot changes with respect to different potential biases applied to the top gate.

A simple flow chart depicting the path of the simulation is shown in figure 3-1. First, our simulation creates the geometry describing the semiconductor structure based on a collection of parameter variables. Each point in a three-dimensional is assigned spatial coordinates and a material type, which defines the point's physical properties. Then, the program proceeds to relax each point in the grid by updating the electrostatic potential at each point based on the charge at the point and the electrostatic potential of the surrounding points. The simulation iterates through the relaxation loop until the algorithm converges to a solution and the calculated error is below a certain threshold value. Once relaxation is completed, the electrostatic potential and the parameters used to define the semiconductor structure are written to various files in a newly created directory.



Figure 3-1: Flow chart of the simulation.

3.1 Initialization

The basic geometry of the semiconductor structure that we wish to investigate comprises of a series of different GaAs and AlGaAs layers sandwiched between two metal gates. The bottom gate is modeled as just a slab of metal while the top gate contains an indentation that is shaped like two intersecting cylinders whose bases are parallel to the face of the bottom gate. This hollowed out indentation is a double cap containing semiconductor material and breaks the symmetry of the structure such as to create a double potential well within the 2D electron gas layer, thus forming a double quantum dot. The simulation describes this geometry using a three-dimensional grid of points, each having a particular material type. Materials include semiconductor, dielectric, and metal. Each material type has an associated electrical permittivity and different properties during relaxation (for example, metal points are not relaxed because they are at constant electrostatic potential). Each point is also assigned spatial coordinates and material type is assigned based on the location of the grid points.

3.2 Relaxation

After the grid of points are properly defined, the program proceeds to go through a Gauss-Seidel relaxation algorithm with successive overrelaxation[15], which comes from treating Poisson's Equation on a grid of discrete points and relating the electrostatic potential at one grid point to the potential of its nearest neighbors.

3.2.1 Using Poisson's Equation

According to Poisson's equation, $-\nabla \cdot \epsilon \nabla U(x) = \rho(x)$, where U(x) is the electrostatic potential, $\rho(x)$ is the spatially varying charge density, and ϵ is the spatially varying dielectric value. By choosing a small volume V centered on a point in space and integrating over the volume, we get:

$$\int_{V} -\nabla \cdot \epsilon \nabla U(x) dV = \int_{V} \rho(x) dV$$
$$\int_{S} -\epsilon \nabla U(x) \cdot d\vec{S} = Q_{enclosed}$$
$$\int_{s} \epsilon E(\vec{x}) \cdot d\vec{S} = Q_{enclosed}$$
(3.1)

where $\vec{E(x)}$ is the electric field. Thus, the amount of charge within the volume, $Q_{enclosed}$ is equal to the sum of the electric flux across the surface S surrounding the volume. Since the simulation uses Cartesian coordinates, one can imagine splitting up space into a series of small cubes, each one centered on a particular grid point, labeled with the indices i, j, and k. The charge within each cube is simply equal to the volume of the cube times the charge density ρ at that point (ρ is either equal to zero (for dielectrics and metals) or is a function of the electrostatic potential at the point (for semiconductors; the details for this function is described in a section below)). The electric flux, scaled according to the relative permittivity $\epsilon_{rel} = \epsilon/\epsilon_0$ of each point and the surrounding points, across the surface of the cube is equal to the sum of the fluxes across each face of the cube. Each face of the cube lies between the central grid point and one of the six adjacent grid points (so each cube face has a corresponding adjacent grid point). See figure 3-2 for a simple drawing of this setup.



Figure 3-2: Central and adjacent grid points, plus cubic cell.

It is assumed that the relative permittivity across a particular cube face is the average of the relative permittivity of the central grid point and the relative permittivity of the corresponding adjacent grid point. Furthermore, we also assume that the relative permittivity for AlGaAs and GaAs are equal, with the relative permittivity of GaAs equal to 13.13 [13]. The electric field normal to each face is equal to $(U_{central} - U_{adjacent})/dr$, where $U_{adjacent}$ is the electrostatic potential at the corre-

sponding adjacent grid point, $U_{central}$ is the electrostatic potential at the central grid point, and dr is the distance separating the two grid points. The electric flux across each face is equal to the normal electric field times the area of the cube face times ϵ_{rel} .

By setting the sum of the fluxes (which is a function of the electrostatic potentials of the central grid point and the adjacent grid points) equal to the charge contained within the cube, one can see that one can rearrange the resulting equation to describe the electrostatic potential of the central grid point as a function of the electrostatic potential at the adjacent grid points and the charge density at the central grid point. Let $U_{central}$ be the electrostatic potential at the central grid point and U_1 through U_6 are the electrostatic potentials of the six adjacent grid points. In this case, the electric flux through the *i*th face of a cube centered on the central grid point is:

$$\phi_i = dA \frac{U_{central} - U_i}{dr},\tag{3.2}$$

where dA is the area of the face of the cube. This gives us:

$$\sum_{i} \epsilon_{i} \phi_{i} = Q_{enclosed} / \epsilon_{0}$$

$$\sum_{i} \epsilon_{i} dA \frac{U_{central} - U_{i}}{dr} = Q_{enclosed} / \epsilon_{0}$$
(3.3)

where ϵ_i is the relative permittivity at the *i*th face of the cube. One can use equation 3.3 to solve for $U_{central}$ in terms of U_i :

$$U_{central} = \frac{Q_{enclosed}/\epsilon_0 + \sum \frac{\epsilon_i dA U_i}{dr}}{\sum \frac{\epsilon_i dA}{dr}}$$
(3.4)

The essence of the relaxation method is to start with an initial guess for the electrostatic potential and then iterate through each point repeatedly, each time setting the electrostatic potential at that point equal to value described by equation 3.4 (this is referred to as relaxing each point) [15]. One continues to relax each point until

the algorithm converges to a stable answer. The simulation determines when this convergence has taken place by calculating the sum of the error at each point. The error at a given point is equal to $(Q_{enclosed} - \sum \phi_i)^2$, where $Q_{enclosed}$ is the charge present in the small cube surrounding the point and $\sum \phi_i$ is the sum of the fluxes through the surface of the cube. That is, the error at a point is equal to how much the electrostatic potential at the point (and at the surrounding points) deviate from the values expected by Poisson's equation, squared. The simulation stops relaxing points once the total error has decreased below a given threshold value. An example plot of the total error for all grid points (χ^2) versus relaxation loop iteration number (with each relaxation loop step comprising of relaxing all of the grid points once) can be seen in figure 3-3. This data comes from a fully three-dimensional geometry in which one cannot use symmetry to reduce the number of dimensions of the geometry. It is possible to use a geometry with translational symmetry in one direction (for example, along the y-axis) to reduce the simulation to a two-dimensional case, which can be performed much more quickly and enable fast tests of basic properties of the simulation. Such two-dimensional simulations have χ^2 versus relaxation loop iteration number curves that are very similar to those from three-dimensional cases. However, the number of operations required in each relaxation loop iteration is proportional to the number of grid points in the simulation's geometry. The three-dimensional simulations have far more grid points than the two-dimensional simulations have, so each of the relaxation iterations for the three-dimensional simulations takes a proportionally longer time to execute than the relaxation iterations for the two-dimensional simulations. This leads to the speed improvement of two-dimensional simulations over three-dimensional simulations.

3.2.2 Overrelaxation

As a way to speed up the algorithm and prevent nonlinear instabilities, successive overrelaxation/underrelaxation [15] is performed in the simulation. When relaxing a given point, instead of setting its electrostatic potential equal to the value determined by the rearrangement of Poisson's equation via a finite difference method, one



Figure 3-3: Error plotted versus relaxation loop iteration number for a fully threedimensional geometry.

determines the calculated update value (the difference between the new electrostatic potential and the old one) and then instead update the electrostatic potential by an amount equal to the originally calculated update value times a relaxation factor, which varies between 0 and 2. In other words, when relaxing a given point during a particular relaxation cycle, instead of setting the electrostatic potential equal to $U_{central}$ (the new value for the electrostatic potential whose value is defined by equation 3.4), the potential is set to $U_{previous} + \delta * (U_{central} - U_{previous})$, where $U_{previous}$ is the value of the electrostatic potential at the given point from the previous update and δ is the relaxation factor.

When the relaxation factor is greater than 1, the program actually overrelaxes, but this can provide a speed-up by causing the electrostatic potential to reach the final solution more quickly. However, this overrelaxation can cause problems due to the nonlinear nature of the relaxation, caused by the presence of charge density in semiconductor points. The program could possibly continually overshoot a stable solution by overrelaxing points. In this case, the update can be dampened with a relaxation factor of less than 1. In practice, this helps to prevent instabilities. When the algorithm starts out, it overrelaxes points with a relaxation factor equal to 2. Whenever the error starts increasing too much, the program lowers the relaxation factor.

3.3 Semiconductor Charge Density and Electrostatic Potential

The charge density at a semiconductor grid point is calculated using a Thomas-Fermi approximation described in [13]. Let the chemical potential (at $T \approx 0$) be $\mu \approx \epsilon_F^0 = \frac{\hbar^2}{2m_e^*} (3\pi^2 n_0)^{2/3}$, where m_e^* is the effective mass of the electron and n_0 is the electron number density at U = 0. The chemical potential is constant throughout the semiconductor structure (if it were not constant, electrons at locations with higher μ would move about to locations with lower μ until equilibrium was achieved and μ was constant). At areas where $U \neq 0$, we have:

$$\mu = \epsilon_F(\vec{r}) - eU(\vec{r}) \approx \frac{\hbar^2}{2m_e^*} \left(3\pi^2 n(\vec{r})\right)^{2/3} - eU(\vec{r}) \approx \frac{\hbar^2}{2m_e^*} \left(3\pi^2 n_0\right)^{2/3}$$
(3.5)

Using $\rho(\vec{r}) = e * (n_0 - n(\vec{r})), \rho_0 = e * n_0$, and $b = \frac{2m_e^* e^{5/3}}{\hbar^2 (3\pi^2)^{2/3}}$, one can solve the above equation to get an expression for the charge density:

$$\rho(\vec{r}) = \begin{cases}
\rho_0 - \left(\rho_0^{2/3} + bU(\vec{r})\right)^{3/2} & \text{for } U(\vec{r}) > -\frac{\rho_0^{2/3}}{b} = U_{depl} \\
\rho_0 & \text{for } U(\vec{r}) \le U_{depl}
\end{cases} (3.6)$$

In this case, ρ_0 is the charge density provided by doping. The effective mass m_e^* is the effective mass of electrons in the conduction band for the semiconductor. Modifications to this equation can be used in the case of 2D systems and systems with applied magnetic fields; one simply has to calculate the new density of states under each system and calculate the Fermi energy based on the electron density. A plot of electron density as a function of electrostatic potential can be seen in figure 3-4, with



reduced mass = $0.001m_e$ and $\rho_0 = 1.55 * 10^{17} \text{e per cm}^3$.

Figure 3-4: Electron density as a function of electrostatic potential (with reduced mass = $0.001m_e$ and $\rho_0 = 1.55 * 10^{17} e \text{ per cm}^3$).

3.4 Band Offsets

When two different semiconductor materials are adjacent to each other, a band offset at the junction plane is created as electrons transfer from one material to the other so that both materials have the same chemical potential. A band offset is a discontinuous jump in the electrostatic potential which originates from the electric field of the displaced electrons [13]. These sudden drops or rises in electrostatic potential are used to create a confinement potential for electrons, thus trapping them within a two-dimensional electron gas (2DEG). Band offsets are implemented in the simulation by enforcing certain boundary conditions along the junction plane. The electrostatic potential of the grid points just above the junction plane is set to be equal to the electrostatic potential of the grid points just below the junction plane plus a fixed amount (i.e. the band offset). In other words,

$$U_{above} = U_{below} + \Delta \tag{3.7}$$



Figure 3-5: Sample plot of two-dimensional lateral confinement potential on a plane within the quantum well layer.

where U_{above} is the electrostatic potential just above the junction, U_{below} is the electrostatic potential just below the junction, and Δ is the value for the band offset (in simulation, we used values of magnitude 0.26 V) [14]. See figure 3-5 for a diagram of this jump in the electrostatic potential for a continuous case. When relaxing the grid points just below the junction plane, the band offset value is subtracted from the electrostatic potential for the adjacent grid points that are just above the junction plane. A similar adjustment was used for relaxing the grid points just above the junction plane, but instead the band offset value was added to (instead of subtracted from) the electrostatic potential for the adjacent grid points that are just below the junction plane.

3.5 Reporting Output

After relaxation is over, the simulation program makes a new directory whose name is based on the date and time at which the program started. In this directory, simulation creates the output files containing the parameters of the simulation (including the size and spacing of the grid of points, the various dimensions of the geometry, the locations of band offsets, etc) and a series of files containing the electrostatic potential solution. The grid is separated into several planes of points, each plane parallel to the xy plane. An output file is created for each plane (indexed according to the z value of the plane) and holds the electrostatic potential within that plane in matrix format, with each line describing a row of points.

3.5.1 Sample Output



Figure 3-6: Sample plot of two-dimensional lateral confinement potential on a plane within the quantum well layer.

Figures 3-6, 3-7, and 3-8 contain sample plots of the output of the simulation for the standard double dot geometry and zero applied external potential. Figure 3-6 is a three-dimensional plot of the two-dimensional lateral confinement potential on



Figure 3-7: Sample contour plot of two-dimensional electron density within the quantum well layer. Electron density is in units of e per nm².

a plane within the quantum well layer. The shape of this potential demonstrates how the electrons in the quantum well are laterally confined. Figure 3-7 is a contour plot of the two-dimensional electron density within the quantum well layer; its shape illustrates how electrons distribute themselves in response to the two-dimensional confinement potential. Finally, figure 3-8 is a plot of the potential energy (equal to -|e| times the electrostatic potential) versus Z coordinate along a line that goes through the center of geometry, from the bottom metal gate (at z = 0) to the top metal gate (at z = 200 nm). In this plot, the band structure and band offsets are clearly seen. The quantum well layer resides near the z = 98 nm to z = 116 nm region, in between the lower blocking barrier layer (roughly z = 90 nm to z = 98 nm) and the upper blocking barrier layer (roughly z = 118 nm to z = 160 nm).

Size of Single Quantum Dot

As a side note, we also examined how the size of a single quantum dot varied with the application of an external potential bias to the top gate. We changed the double cap to a single pillar shape and ran the simulation with different negative potential



Figure 3-8: Potential energy plotted versus the Z coordinate (along the center of the geometry). The bottom metal gate is located at z = 0 while the top metal gate is located at z = 200 nm.

biases applied to the top gate, which slowly depleted electrons in the single dot. We recorded the distribution of electron density in the quantum well layer as well as the confinement potential for each external potential bias value. See figure 3-9 for the plots of the confinement potential and figure 3-10 for the plots of the electron density. Figure 3-11 has similar plots of quantum well electron density for different potential values, but shows plots for four different cap layer radii: 150 nm, 250 nm, 350 nm, and 450 nm. The electron density plots illustrate how the electron density within the quantum dot decreases as one increasing the repelling top gate potential bias, but the radius of the dot does not change significantly, particularly for larger cap radius values. Furthermore, while the confinement potential far from the quantum dot region changes in response to the different top gate potential biases, the confinement potential at the center of the dot responds much less drastically to changes in the top gate potential. One possible explanation for this is that the confinement potential within the dot is mostly determined by the potential at the top of the etched pillar, where the top gate meets the cap layer. However, because the top gate is farther from the quantum well layer in the regions of the double dot (which are directly below the etched pillar) than it is for the region beyond the double dot, where the top gate is closer to the quantum well layer, the roughly 1/r dependence of the electrostatic potential decreases the magnitude of the changes in the quantum dot potential in response to changes to the top gate potential. Furthermore, nonlinear effects caused by the presence of the semiconductor material in the cap layer could also minimize the effect of changes of the top gate potential on the electrostatic potential in the quantum dot region of the quantum well layer.



Figure 3-9: Confinement potential plotted versus the X coordinate within the quantum well region, for various potential biases applied to the top gate. The listed potential values are equal to the absolute value of the (electron repulsing) top gate potential bias.



Figure 3-10: Electron density plotted versus the X coordinate within the quantum well region, for various potential biases applied to the top gate. The listed potential values are equal to the absolute value of the (electron repulsing) top gate potential bias.



(a) cap radius = 150 nm



Figure 3-11: Electron density versus X coordinate plots for single dots of various cap layer radii. For each radius, the electron density is plotted for a set of different voltages applied to the top gate.

Chapter 4

Testing of Parameters

Although in theory some of the parameters of the simulation can be determined a priori (such as the effective mass of the electron within the semiconductor and the doping density in the semiconductor layers), the various approximations that were made encouraged tweaking the values of these parameters until the simulation produced results that matched that seen in experiments. Once the simulation behaves properly for a pre-defined set of conditions, one can then explore other geometries to anticipate the properties of other quantum dots that have yet to be examined experimentally.

The primary measurable responses of the simulation on which we focused were the depletion voltage (V_{depl} , the value of the top gate potential at which the quantum dot region is completely depleted of electron density), the electron density in the center of the quantum dot when the top gate is at zero potential (ρ_{dot}), and the electron density in the quantum well layer far from the quantum dot region (ρ_{far}). The main parameters that we adjusted were the effective mass of the electron (m_e^*), the delta doping density in the AsSi doping layer (ρ_d), and the doping density in the intrinsic GaAs (ρ_0). The first and third of these parameters are key features of the electron density as a function of electrostatic potential, $\rho(U)$. These adjustable parameters were altered until the response variables matched experimental values: $V_{depl} \approx 0.5$ V [18], $\rho_{dot} \approx 10^{-3} - 10^{-2}$ electrons per nm² [17], and $\rho_{far} \approx 0$ [18]. After much trial and error during which we tested several combinations and ranges of parameters,

we found that the best set of parameters that produced these response values were $\rho_d = \rho_0 = 1.55 \times 10^{17}$ electrons per cm³ and $m_e^* = 0.001 m_e$. Once these values were found, we were ready to move on to the testing of sample heterostructure geometries, including the pincher gate geometry.



Figure 4-1: Electron density far from the quantum dot region versus doping density ρ_0 .



Figure 4-2: Electron density within the quantum dot region versus doping density ρ_0 .

To get a feel of how we chose the values for the adjustable parameters, figures 4-1, 4-2, and 4-3 contain plots of data from some of the simulations that were run in order to test values for doping density ρ_0 and effective mass m_e^* . Figure 4-1 is a plot of quantum well layer electron density far from the quantum dot region (ρ_0) versus doping density ρ_0 (keeping all other parameters constant). It demonstrates



Figure 4-3: Electron density within the quantum dot region versus electron effective mass.

the appropriate range of values for ρ_0 for which $\rho_{far} = 0$. As expected, decreasing the amount of free electrons in the system decreases ρ_{far} in a roughly linear fashion, until it reaches the minimum value of 0. Figure 4-2 plots the electron density within the dot (ρ_{dot}) versus ρ_0 . Like the case for ρ_{far} , increasing ρ_0 causes a increase in ρ_{dot} . There is thus a tradeoff in adjusting ρ_0 between keeping ρ_{far} close to zero while keeping ρ_{dot} sufficiently large. Lastly, figure 4-3 shows the relation between ρ_{dot} and the effective mass of the electron. Appropriate values of ρ_{dot} results from making the effective mass near the reasonably small value of 0.001.

Chapter 5

Pincher Gate and Middle Gate Simulations: Separating the Double Dot

The simulation was first used to investigate double dot geometries to determine which could, after applying the appropriate external voltages to the pincher gates and middle gates, separate the double dot electron density into two distinct blobs of charge within the 2DEG layer that partially overlap with each other. A geometry involving side gates, which were a pair of metal electrodes that approached but did not overlap with the double pillar, was initially tried, but we discovered that the side gates had limited influence on the lateral interaction because the side gates were situated too far from the region where the two dots overlapped. Thus, the side gate geometry was replaced by the pincher gate geometry, which could effectively manipulate the electrostatic potential and electron density in the overlap region of the double dot.

5.1 Side Gate Geometry

The first geometry that was tested involved two side gates, which were two separate electrodes located on the underside of the top metal gate near the overlap region of the cap region, but not actually extending into the cap region. See figure 5-1 for a



Figure 5-1: Diagram of side gates.

diagram of the setup for the side gates. Theoretically, if one were to apply a negative electrostatic potential to the side gates, the unfavorable potential would drive out electrons in the overlap region of the double quantum dot in the quantum well region by raising the potential energy of electrons residing there. The double quantum dot would thus become separated into two unconnected single dots and theoretically one could tune the lateral interaction of the two dots by controlling the potential applied to the side gates.

5.2 Pincher Gate Geometry

However, test simulations showed that due to the large distance between the overlap region of the double quantum dot and the tips of the side gates (on order of 100 nm), the side gates had relatively little influence on the electron density in the quantum dot. Separation of the double dot could only be achieved with side gate electrostatic potentials of over 50 V (a contour plot of the electron density within the quantum well layer for this geometry can be seen in figure 5-2). However, it is expected that this



Figure 5-2: Electron density contour plot in quantum well layer with original side gate-only geometry (with side gate potential = 50 V).

value is unacceptably high due to avalanche breakdown, in which the intense electric fields caused by the high potential bias accelerates electrons so much that they lead to the creation of electron-hole pairs and a large current between the side gates and the surrounding materials [19].

Therefore, the side gate geometry was discarded in favor of a pincher gate setup, as described in section 2.2. Unlike the side gates, the pincher gate extends over the overlap region of the quantum double dot, while remaining on the underside of the top metal gate. Simulations indicated that the pincher gate could separate the double dot with applied electrostatic potentials of 0.3 to 0.5 V below that of the top metal gate. Figure 5-3 shows a contour plot of the electron density within the quantum well layer with the pincher gate (with a voltage that is 0.3 V lower than the top metal gate), demonstrating the separation of the double quantum dot into two separate dots. Figure 5-4 shows various contour plots for electron density within the quantum well layer after applying different pincher gate potential bias values, illustrating how one can use the pincher gate to tune the coupling strength between the two lobes



Figure 5-3: Electron density contour plot in quantum well layer with pincher gate setup.

of the double quantum dots or even separate them completely. Some amount of separation of the two lobes is achieved at a pincher gate potential bias of 0.1 V. At around 0.4 V to 0.5 V, the two lobes become completely separate. One can further raise the potential bias of the pincher gate to increase the barrier between the two dots even more so and control the lateral interaction.

5.2.1 Middle Gates

The purpose of the addition of the middle gates along with the pincher gate is to duplicate the function of the top gate. They work to both deplete electron density within the quantum well region and to create the confinement potential for the double dot as well as to act as electrodes for capacitance measurements. Furthermore, one can apply a different potential bias value to one of the middle gates and measure its effects on the overall electron density and the electron density within the dot underneath the other middle gate. For example, figure 5-5 shows plots of quantum well electron density ρ versus x (along a line that cuts through the middle of the two quantum dots)





Figure 5-4: Zoom-in of contour plots for quantum well electron density for various values of pincher gate voltage bias.

for different values of potential bias in the first middle gate. A constant potential bias is also applied to the pincher gate, slightly separating the two dots. One can see the electron density in the left quantum dot decrease as one increases the potential bias to the first middle gate, which depletes the electron density from the quantum dot directly beneath it while having little effect on the other quantum dot. Contour plots of the corresponding electron density within the quantum well region can be seen in figure 5-6.

Variable Dot Size and other Tunable Features

Ultimately, one can use the pincher gates and middle gates to tune the confinement potential of the double dot system. From the electron density plots for variable middle gate potentials, one can see that by partially depleting the electrons from one of the two dots, the middle gates can effectively shrink the size of one dot while the other dot remains mainly unaffected. Thus, one can tune the relative sizes of the two dots and examine how a smaller dot couples with a large dot. It is expected that as one reduces the size of one dot, the energy spacings of states for electrons confined to that dot will increase. By altering the potential bias for the other middle gate, one can also raise and lower the more closely spaced energy levels of the larger dot. It is expected that interesting phenomena will occur when the energy levels line up, such as resonance tunneling of an electron from one dot to another. This resonance tunneling was examined by van der Vaart et al. and Waugh et al. via transport experiments using double and triple quantum dots made from lithographic etching[10, 11]. This matching of energy levels is illustrated in figure 5-7. Furthermore, a potential applied to the middle gate is expected to distort the associated dot's confinement potential, breaking it's quasi-cylindrical symmetry.



Figure 5-5: Electron density $\rho(x)$ for various middle gate potential bias values.





Figure 5-6: Zoom-in of contour plots for quantum well electron density for various values of first middle gate potential bias $U_{middlegate1}$.



Figure 5-7: Cartoon of energy levels for a small dot (with widely spaced energy levels) coupled with a large dot (with closely spaced energy levels).

Chapter 6

Conclusion

The simulation was able to model the behavior of a double quantum dot geometry via its relaxation algorithm and Thomas-Fermi approximation. It also allowed for the testing of various geometries that would conceivably allow experimenters to control the shape and properties of the quantum double dot, such as a pincher gate geometry that could separate the double dot into two single quantum dots or manipulate the coupling strength between the two lobes of the double dot. The simulation demonstrated the pincher gate's ability to modify the lateral interaction of the two dots. Also, the simulation showed how the middle gates, in combination with the pincher gate, could alter the relative sizes of the two dots and allow for experiments that analyzed the effects of coupling a smaller quantum dot with a larger dot. With this combination of upper gates, one can tune various properties of the confinement potential of the dot, allowing for a host of experiments on quantum dots and confined electron systems.

The double dot geometry that was used to test the simulation is only one possible geometry that can be examined by the simulation. Other geometries could be tested by the program, thus allowing for the exploration of other ways to control the properties of the double quantum dot and analyze its rich physics. For example, one could use a modified heterostructure and analyze vertically coupled quantum dots.

Various improvements could still be made to the simulation, including replacing some of the simplifying assumptions that we have made about the physics of the electron system and heterostructure with more sophisticated models. The relaxation algorithm could be optimized instead of merely relying on a simple over-relaxation scheme. Fully three-dimensional simulations can take minutes or hours to run, so an acceleration of the program would be useful. For example, one might implement multigrid relaxation techniques in the simulation [15]. Furthermore, the Thomas-Fermi approximation could be expanded upon to include electron-electron interactions. Presumably, the electrons in the two separate quantum dots could couple with each other via the Coulombic interaction, in addition to quantum mechanical coupling. Additionally, a self-consistent Poisson-Schrödinger solver might also be incorporated into a variant of the simulation. Such an improvement would help incorporate more quantum effects into the simulation without depending so much on the Thomas-Fermi approximation, whose accuracy breaks down if the electrostatic potential varies drastically over length scales equal to the Fermi wavelength. Finally, one could replace the implicit assumption that the ionized donor atoms present in the semiconductor materials are evenly distributed with a more complicated distribution that is a function of spatial coordinates, gate voltage, and electron density within the quantum dot [2].

Bibliography

- [1] R.C. AShoori. Electrons in artificial atoms. *Nature*, **379**, 413. 1996.
- [2] S. Bednarek, B. Szafran, K Lis, and J. Adamowski. Modeling of electronic properties of electrostatic quantum dots. *Phys. Rev. B*, 68, 155333. 2003.
- [3] M. Kastner. Artificial atoms. *Physics Today*, 46:24. 1993.
- [4] L.P. Kouwenhoven, F.W.J. Hekking, B.J. van Wees, C.J.P.M. Harmans, C.E. Timmering, C.T. Foxon. Transport through a Finite One-Dimensional Crystal. *Phys. Rev. Lett.* 65, 361-364. 1990.
- [5] K.A. Matveev, L.I. Glazman, and H.U. Baranger. Coulomb blockade of tunneling through a double quantum dot. *Phys. Rev. B* 54, 5637-5646. 1996.
- [6] J.J. Palacios and P. Hawrylak. Correlated few-electron states in vertical doublequantum-dot systems. *Phys. Rev. B* 51, 1769-1777. 1995.
- [7] M.A. Reed, etal., Festköperprobleme/ Advances in Solid State Physics, 29, 267.
 1989.
- [8] C.A. Stafford and S. Das Sarma. Collective Coulomb blockade in an array of quantum dots: A Mott-Hubbard approach. *Phys. Rev. Lett.* 72, 3590-3593. 1994.
- [9] S. Tarucha, T. Honda, D.G. Austing, Y. Tokura, K. Muraki, T.H. Oosterkamp, J.W. Janssen, L.P. Kouwenhoven. Electronic states in quantum dot atoms and molecules. *Physica E*, 3, pp. 112. 1998.

- [10] N.C. van der Vaart, S.F. Godijn, Y.V. Nazarov, C.J.P.M. Harmans, J.E. Mooij, L.W. Molenkamp, and C.T. Foxon. Resonant Tunneling Through Two Discrete Energy States. *Phys. Rev. Lett.* **74**, 4702-4705. 1995.
- [11] F.R. Waugh, M.J. Berry, D.J. Mar, R.M. Westervelt, K.L. Campman, and A.C. Gossard. Single-Electron Charging in Double and Triple Quantum Dots with Tunable Coupling. *Phys. Rev. Lett.* **75**, 705-708. 1995.
- [12] T. Chakraborty. Quantum Dots: A survey of the properties of artificial atoms. Elsevier Science B.V. Amsterdam. 1999.
- [13] C. Kittel. Introduction to Solid State Physics: Seventh Edition. John Wiley and Sons, Hoboken, NJ. 1996.
- [14] W. Mönch. Semiconductor Surfaces and Interfaces. Springer-Verlag Berlin Heidelberg, Germany. 2001.
- [15] Press et. al. Numerical Recipes in C. Cambridge University Press. 1992.
- [16] K. Seeger. Semiconductor Physics: An Introduction. Springer-Verlag Berlin Heidelberg, Germany. 2004.
- [17] R. Ashoori. The Density of States in the Two-Dimensional Electron Gas. Cornell University. 1991.
- [18] M. Brodsky. Charging of Small Two-Dimensional Electron Puddles. MIT, Cambridge, MA. 2000.
- [19] URL: http://ece-www.colorado.edu/~bart/book/book/chapter4/ch4_5.htm