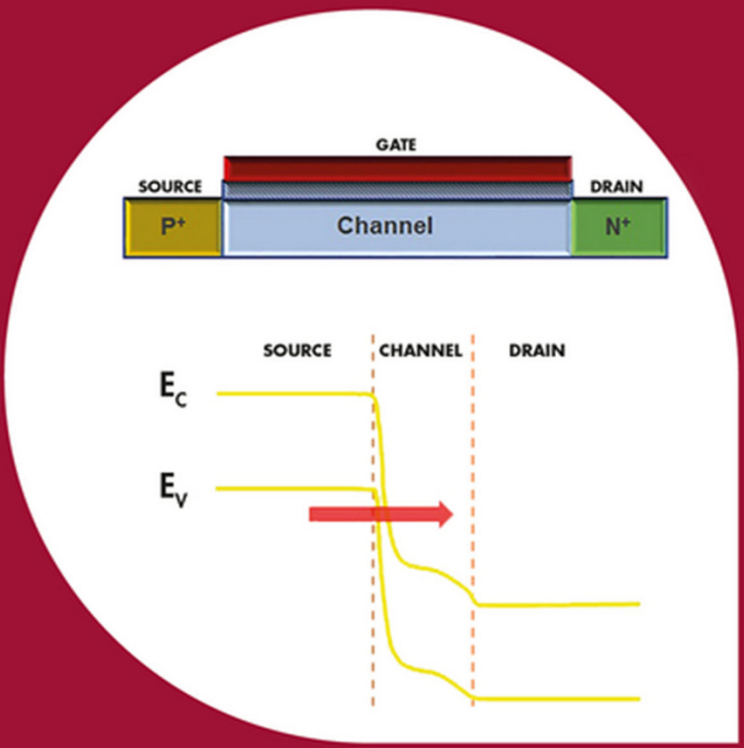

TUNNEL FIELD-EFFECT TRANSISTORS (TFET):

Modelling and Simulation

JAGADESH KUMAR MAMIDALA
RAJAT VISHNOI
PRATYUSH PANDEY

WILEY

Tunnel Field-Effect Transistors (TFET)

Tunnel Field-Effect Transistors (TFET)

Modelling and Simulation

Jagadesh Kumar Mamidala
Indian Institute of Technology (IIT), Delhi, India

Rajat Vishnoi
Indian Institute of Technology (IIT), Delhi, India

Pratyush Pandey
University of Notre Dame, Indiana, USA

WILEY

This edition first published 2017

Registered Office
John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

Names: Kumar, Mamidala Jagadesh, author. | Vishnoi, Rajat, author. | Pandey, Pratyush, author.
Title: Tunnel field-effect transistors (TFET) : modelling and simulations / Jagadesh Kumar Mamidala, Rajat Vishnoi and Pratyush Pandey.
Description: 1 | Hoboken : Wiley, 2016. | Includes bibliographical references and index.
Identifiers: LCCN 2016023165 (print) | LCCN 2016041464 (ebook) | ISBN 9781119246299 (hardback) | ISBN 9781119246282 (pdf) | ISBN 9781119246305 (epub)
Subjects: LCSH: Tunnel field-effect transistors. | Integrated circuits–Design and construction. | Nanostructured materials. | Low voltage integrated circuits. | BISAC: TECHNOLOGY & ENGINEERING / Electronics / Semiconductors.
Classification: LCC TK7871.95 .K86 2016 (print) | LCC TK7871.95 (ebook) | DDC 621.3815/284–dc23
LC record available at https://lccn.loc.gov/2016023165

A catalogue record for this book is available from the British Library.

Set in 10.5/13pt Times by SPi Global, Pondicherry, India

10 9 8 7 6 5 4 3 2 1

Contents

Preface

Overview

Just as a human body is made up of millions of biological cells, an integrated circuit is made up of millions of transistors. Transistors are the basic building blocks of all modern electronic gadgets. Ever since the advent of CMOS circuits, the dimensions of the transistor have been continuously scaled down in order to pack more logic on to a silicon wafer and also to reduce power consumption in the circuits. In recent years, with mobile devices becoming popular, the search for low power devices with steep switching characteristics has become important. Highly scaled MOSFETs are rendered unsuitable for low power applications due to a thermal limit on their switching. Hence, the Tunnel Field Effect Transistor (TFET) is being explored extensively for low power applications. A TFET has a steep switching characteristic as it works on the phenomena of band-to-band tunnelling. Over the past few years, TFETs have been heavily researched by various notable groups in the field of semiconductor devices across the globe.

This book provides a comprehensive guide for those who are beginning their study on TFETs and also as a guide for those who wish to design integrated circuits based on TFETs. The book covers the essential physics behind the functioning of the TFETs and also the device modelling of TFETs, for the purpose of circuit design and simulation. It begins with studying the basic principles of quantum mechanics and then builds up to the physics behind the quantum mechanical phenomena of band-to-band tunnelling. This is followed by studying the basic functioning of TFETs and their different structural configurations. After explaining the functioning of TFETs, the book describes different approaches used by researchers for developing the drain current models for TFETs. Finally, to help new researchers in the field of TFETs, the book describes the process of carrying out numerical simulations of TFETs using the TCAD tool Silvaco ATLAS. Numerical simulations are helpful tools for studying the behaviour of any semiconductor device without getting into the complex process of fabrication and characterisation.

Key feature in relation to existing literature

This book is the first comprehensive literature on TFETs, which are very popular transistors and have been extensively studied in recent years; they are going to be important building blocks for low power solid state circuits in the future. It is a one-stop volume for studying TFETs for someone who has a basic knowledge of MOSFET physics. It covers the physics behind the phenomena of tunnelling as well as the device physics of TFETs. It also has a unique feature of describing device simulation along with device physics so as to enable readers to do further research on TFETs.

The presentation of the book is clear and accurate and is written in simple language. The book endeavours to explain different phenomena in the TFETs using simple and logical explanations so as to enable the reader to get a real feel for the functioning of the device. Also, each and every aspect of the TFET has been compared to that of the MOSFET so that the facts presented in the book make more sense to the entire semiconductor device fraternity and help in the integration of the TFET with the prevailing technology in the industry. The book also attempts to cover all the recent research articles published on TFETs so as to make sure that, along with covering the basics, it also covers state of the art work on TFETs.

1

Quantum mechanics

1.1 Introduction to quantum mechanics

Before attempting to investigate the workings of a tunnelling field-effect transistor, it is essential to be familiar with the concept of tunnelling. Tunnelling is a quantum phenomenon, with no counterpart in the everyday physics one encounters, or the physics that one applies while dealing with devices a few hundred nanometres in length. The initial two chapters will, therefore, help us develop an understanding of quantum phenomena. In this chapter, we will present an introduction to the field of quantum mechanics and the next chapter will discuss the phenomenon of tunnelling in detail.

The chapter begins with a description of a landmark experiment that conclusively proved the wave nature of particles, after which we will study the concept of wavefunctions and how to use Schrodinger's equation to obtain them. A few basic problems will be presented so that the readers may familiarise themselves with basic quantum concepts.

1.1.1 The double slit experiment

There are many experiments that led to the conception of quantum mechanics – blackbody radiation, the Stern Gerlach experiment, the photoelectric effect, the line spectrum, etc. However, for our purposes we will concentrate on one of the landmark experiments, that is the double slit experiment, which demonstrated the fundamental quantum nature (i.e. both wave and particle) of electrons.

Tunnel Field-Effect Transistors (TFET): Modelling and Simulation, First Edition. Jagadesh Kumar Mamidala, Rajat Vishnoi and Pratyush Pandey.

You would have read that only waves can undergo superposition, and not particles. Superposition is the fundamental principle behind the occurrence of interference – therefore, if something exhibits interference, it must have a wave nature. The double slit experiment is famously associated with Thomas Young, who used it for the first time in the early nineteenth century to prove the wave nature of light. Before this experiment was performed, light had been associated with a particle nature (since the times of Newton), and the fact that it underwent interference was conclusive proof of its wave nature.

However, the behaviour of light that led Newton and others to believe that it had a particle nature could not be reconciled with this newly formed wave picture. It took another century of research and experiments to establish a rather astonishing result regarding the behaviour of light – that it displays both particle and wave natures. The particle nature leads to phenomena such as the photoelectric effect and rectilinear propagation of light in ray optics; the wave nature explained the interference and diffraction of light.

While this dual nature (that is both particle and wave natures) of light was being worked out, many people were, independently, studying the behaviour of subatomic particles. Phenomenon like the discrete line spectrum of hydrogen, the observed distribution of blackbody radiation, etc., could not be explained by any established theory. Theoretical physicists were in a quandary. At this point, de Broglie hypothesised that, just like light, particles possess a dual nature as well. When de Broglie made this hypothesis, there was little evidence to support his claim. A few years later, Davisson and Germer experimentally observed that electrons underwent diffraction just as light did. These were landmark moments in the history of physics – de Broglie received the Nobel Prize in physics (the second time it was awarded for a PhD thesis) and, later, so did Davisson and Germer. While the Davisson–Germer experiment was the first to establish the dual nature of matter, the double slit interference experiment is far easier to conceptually grasp and visualise, which is why we will use it to embark on our study of quantum phenomena.

The setup of an electron interferometer used in the double slit experiment is conceptually quite similar to that of a light interferometer (Figure 1.1). A parallel beam of electrons is incident on a screen with two slits. The electrons that pass through the slits impinge upon the optical screen, where their incidence is captured by a visible spot. First, let us think of these electrons as if they were the kind of particles we observe in our daily lives (classical particles) and see how they should behave. All the electrons in the initial beam have the same speed and direction of motion and they are heading towards the screen with two slits. All the electrons that hit this screen are blocked, except for the ones passing right through the slits. These electrons that passed through the slits should have no reason to change either their speed or the direction of their motion. They do not “know” that there

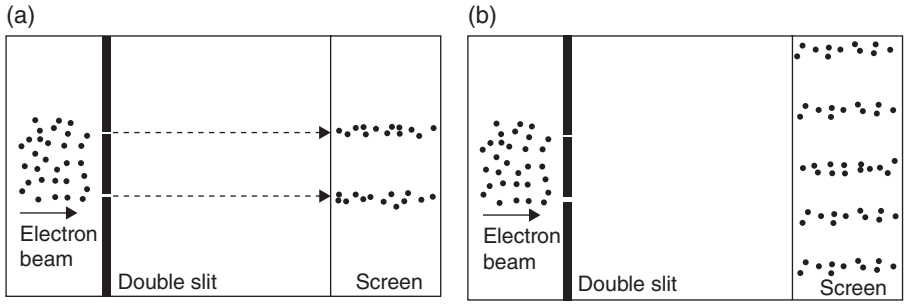

Figure 1.1 (a) Classically predicted electron pattern. It can be seen that interference fringes are experimentally observed, as opposed to the classically predicted pattern. This establishes the wave-like behaviour of electrons, (b) experimentally observed electron pattern.

was a screen in the first place – they pass through unaffected. They subsequently keep heading straight and hit the final screen as illustrated in Figure 1.1(a). Two narrow bands are formed on the screen, corresponding to the two thin beams of electrons that passed unaffected through the two small slits.

Now let us take a look at what was actually observed in the experiment. There was an interference pattern on the screen, as shown in Figure 1.1(b), a pattern uncannily similar to what is observed when we perform the same experiment with light instead of electrons. At this juncture, you might hypothesise this behaviour to result from some sort of statistical phenomenon due to the large number of electrons. However, the experiment is far from finished, and further strangeness lies ahead.

Let us now adjust the electron source so that instead of a beam of electrons it sends a single electron at a time. This time, we find something even more extraordinary – after a lot of electrons have hit the screen, the same interference pattern builds up as in the case of a beam of many electrons. There is no way this electron "knows" that it has been preceded by, or it will be followed by, another electron. What, then, could be happening? The answer is even more puzzling than the question, and will take you quite a while to come to terms with – *each and every electron is undergoing interference with itself*. This is what leads to the final conclusion that not just aggregates of particles but each and every particle exhibits a wave nature. To make this point clear, let us modify the experiment such that we are able to find out through which slit each electron passes. Independent of *how* we find out which slit each electron passes through, we get *exactly* the same result, that is the interference pattern vanishes and we get the pattern shown in Figure 1.1(a), as predicted by classical mechanics. Think about this very carefully, because this

point merits serious investigation. For a wave to show interference, there need to be two sources – the two slits in this case. Thus, for a single electron to show interference, *it must be passing through both slits*. However, this is not possible! At the very least, we cannot imagine such a situation. It is only reasonable to assume that the electron either goes through one slit or the other, but the moment we impose such a restriction on the electron, we are thinking of it as a classical particle. By just *knowing* which slit the electron is going through, and thereby imposing the condition that it will pass through either one slit or the other, we are restricting it to behave like a classical particle. While the mathematical foundations will be laid later in this chapter, for now the reader should try and grasp the underlying concept – the quantum electron passes through *both* the slits; it is a *superposition* of these two states (corresponding to passing through the upper or lower slit). You may think that the electron *actually* passes through either of the two slits and due to limitations of our experimental techniques, we do not know which slit it passes through. This is not the case – the electron is indeed passing through both the slits. This counterintuitive phenomenon is at the very root of quantum mechanics and it will take some time for us to be familiar with this kind of approach. You *cannot* ask of the *quantum* electron (or any general quantum particle), "Which slit does it pass through?" The question in itself is wrong. It passes through both. It should be noted that this wave nature of a particle becomes appreciable only at very small sizes, such as a few nanometres.

1.1.2 Basic concepts of quantum mechanics

1.1.2.1 Wavefunctions

The behaviour of classical particles can be fully explained by describing how their position changes with time. This information would be sufficient to give us the trajectory, the velocity, the momentum and the acceleration of the particle. However, what of the quantum particle? Surely, the electron that passed through both slits of the double slit experiment cannot be assigned a precise location. This leads us to the realisation that we need some new method to describe the quantum particle. The rest of this chapter is devoted to formulating a mathematical picture that is able to capture the unusual behaviour of quantum particles.

The search for this new method of description was helped by the knowledge that the quantum behaviour of particles closely resembled the behaviour displayed by waves. Waves of many kinds – electromagnetic waves, sound waves, etc. – had been extensively studied, and all these waves were described by wave equations. These equations described the behaviour of a wave at every point in space, and at all times. For example, in the case of sound waves, the wave equation described the displacement ($\Delta\vec{r}$) of each particle as a function of time:

$$\Delta \vec{r} = \vec{\psi}\,(x,y,z,t) \tag{1.1}$$

Similarly, for an electromagnetic wave, the wave equations described the electric ($\vec{E}$) or magnetic ($\vec{B}$) field at each and every point as a function of time:

$$\vec{E} = \overrightarrow{\psi_E}\,(x,y,z,t) \tag{1.2}$$

$$\vec{B} = \overrightarrow{\psi_B}\,(x,y,z,t) \tag{1.3}$$

Taking the cue from these equations, physicists assigned a similar wave equation to the quantum particle. This equation was called the wavefunction of the quantum particle and was usually denoted by the Greek symbol ψ (psi). Just like in the case of classical waves, this wavefunction contained all the information about the particle – its current state and the variation of its behaviour with time. It is important to note that while the previously described wave equations (1.1) to (1.3) were real functions, the wavefunction of a quantum particle is a complex function.

1.1.2.2 Born interpretation

While the wavefunction-based formulation of quantum mechanics was proposed by Erwin Schrodinger quite early, he was at a loss to ascribe any physical meaning to it. The theory he built up described what sort of mathematical operations one needed to perform on the wavefunction to get information relating to its various properties, such as its position, its momentum, its energy, etc. However, what this wavefunction *itself* meant was a mystery, especially because it was a complex function. Many interpretations were proposed as to the meaning of this wavefunction, but the one that is most widely accepted was proposed by Max Born. It is known as the "Born interpretation of quantum mechanics" and is one of the fundamental principles of quantum mechanics. According to this interpretation, the wavefunction ψ is the "probability amplitude" of the quantum particle, the square of whose magnitude gives us the probability density ρ of finding that particle at any point:

$$\rho = |\psi|^2 = \psi^* \psi \tag{1.4}$$

where $\psi*$ is the complex conjugate of ψ. Using this interpretation, the probability P of finding the particle in a volume V at any time t would be

$$P(V,t) = \iiint_V \rho \; dx \; dy \; dz \tag{1.5}$$

Since the probability of finding the particle in the entire space should always be unity, we can say that

$$\int\int\int_{-\infty}^{\infty} \psi^{*}\psi \, dx \, dy \, dz = 1 \tag{1.6}$$

A wavefunction that displays this property is called a "normalised" wavefunction.

It is very important to realise that the probabilistic behaviour that follows from the Born interpretation is different from the probabilistic behaviour encountered in statistical mechanics. For example, consider an ensemble of particles in a chamber each occupying a particular position. This allows us to calculate the probability of finding a particle at any position. If there were only a single classical particle in this chamber, we could always precisely identify its position. However, in quantum mechanics, every single particle is "spread out" in space, and its position is uncertain. Even if there is only a single electron, we cannot say, "The electron is at this particular point". We can only talk about the probability of finding the electron at any given point once we measure its position. Thus, the Born interpretation provided a physical meaning to the wavefunction that was compatible with the fundamentally probabilistic behaviour of a quantum particle, and gave a mathematical approach to calculate the probability of finding a quantum particle at any region in space.

1.1.2.3 Measurement

Measurement is a fundamental process in our lives, yet it is so much a part of our instincts that we barely pay any attention to it. However, if you think carefully, most of the information you get is by the process of measurement. When you look at a tree, your eyes measure the frequency and amplitude of the incoming electromagnetic waves, giving you information regarding the colour and brightness of the tree. Subsequently, your eyes measure the angular difference between the signals received by the two eyes, and calculations by your brain tell you how far away this tree is. You may hear a bird chirping on this tree – once again, due to your ears measuring the frequency and location of the pressure waves (sound) impinging upon them. Similarly, any information we get about a quantum particle is by the process of measurement – measuring the position, energy, momentum, etc.

However, there is a very fundamental difference between measurement in classical physics and quantum physics. While the state of a classical particle is independent of measurements performed on it, in quantum mechanics, the state of the quantum particle is intricately linked to measurements performed on it. We shall go back to the double slit experiment to illustrate this point. When

we measured which slit the electron passed through, that is when we measured its position, it stopped showing interference. The electron, before measurement, exhibited interference. After we carried out the measurement, it no longer showed interference. This shows that *measurement changed the state of the electron*. In general, measurement changes the state of a quantum particle, and its final state (after measurement) depends both on its initial state and the kind of measurement being performed. Do not be worried if the picture is not completely clear yet – to fully understand the process of measurement, we will have to know about operators and eigenvalues, which we will do in the next two sections.

1.1.2.4 Operators

The Born interpretation told us that we can obtain the probability of finding a quantum particle at any given point if we know its wavefunction. However, the wavefunction contains far more information than this. If you remember, the wavefunction was supposed to contain *all* the information about the quantum particle. How, then, do we extract this information from the wavefunction?

Since the wavefunction is a mathematical function, it is clear that we will be performing certain mathematical operations on it to get the information we desire. This mathematical operation must be different, depending on the specific kind of information – energy, momentum, position, etc. – we need to obtain. This, indeed, is the case.

Corresponding to every physically observable parameter (also called observables) of a quantum particle, such as position, momentum, energy, we have mathematical operators. The operators for certain common observables are listed below in Table 1.1, where $\imath$ (iota) is the square root of negative unity and $\hbar$ (h-cross or h-bar) is the reduced Planck's constant.

To understand the use of these operators, let us imagine an experiment where we have a large number of quantum particles with the same wavefunction ψ. We wish to measure a particular observable, the mathematical operator corresponding to which is O. The outcome of each measurement is o. As the behaviour of

Table 1.1 Quantum mechanical operators corresponding to physical observables.

Observable	Operator
Position (x)	x
Momentum ($\vec{p}$)	$-\imath\hbar\vec{\nabla}$
Energy (E)	$\imath\hbar\frac{\partial}{\partial t}$

quantum particles is probabilistic, measuring O for every particle will give a different outcome o. Looking back to our example of the double slit experiment, all the incoming electrons were exactly similar. However, when we start measuring which slit they pass through, sometimes we find that an electron passes through the upper slit and at other times through the lower slit. We can, therefore, only discuss the expectation value $\langle o \rangle$ after taking an average of all the measurements. This expectation value of the observable o is given as

$$\langle o \rangle = \frac{\int\int\int_{-\infty}^{\infty} \psi^* O\psi \; dx \; dy \; dz}{\int\int\int_{-\infty}^{\infty} \psi^* \psi \, dx \; dy \; dz} \tag{1.7}$$

The above equation tells us about the expectation value when we perform a large number of measurements, all on particles with the same wavefunction ψ. However, if we have only one particle, it would be useful to know the probability of obtaining a particular result. For us to know this, we must find the eigenfunctions of the operator in question.

1.1.2.5 Eigenfunctions

Let us recollect from the section on measurement (Section 1.1.2.3) that the state of a quantum particle changes upon measurement, and the final state is dependent on both the initial state and the kind of measurement being performed. However, there are certain very special states corresponding to every observable that do not change when it is measured. These special states are the eigenfunctions of that observable. If, for an operator O, the wavefunction ψ_o behaves as

$$O\psi_o = \lambda\psi_o \tag{1.8}$$

where λ is a constant, then ψ_o is an eigenfunction (also referred to as an eigenstate or an eigenvector) of the operator O and λ is the corresponding eigenvalue. Suppose that we measure the observable corresponding to the operator O on a particle having the wavefunction ψ_o. We will find the value of this observable to be λ. This can be proven by substituting the value of $O\psi_o$ from Equation (1.8) into Equation (1.7) that gave us the expectation value corresponding to any operator. Moreover, the wavefunction ψ_o will remain unchanged. Therefore, for a particle whose wavefunction is an eigenfunction of an observable, we can, with absolute certainty, state the result of measurement. To understand this, let us consider the energy operator (Table 1.1) as an example. Let us assume $\psi_{E_i} (i = 0, 1, 2, \ldots)$ to be the eigenfunctions of the energy operator, having eigenvalues E_i:

$$\imath\hbar \frac{\partial \psi_{E_i}}{\partial t} = E_i \psi_{E_i} \tag{1.9}$$

If we take a particle with wavefunction ψ_{E_i}, we *know* that its energy is E_i. There is no probability involved in this.

Furthermore, even if we have a wavefunction that is not an eigenfunction of the operator in question, the result of every measurement can *only* be one of the eigenvalues of the operator. Let us again take the energy operator as an example. Suppose that we measure the energy of a particle having a wavefunction Ψ that is *not* one of the eigenfunctions ψ_{E_i} of the operator. The result will *always* be one of the eigenvalues E_i. Note that every measurement will result in a *different* energy being observed each time. You will now say that once you have measured the energy, and it is found to be a particular E_i, you *know* the energy of the particle to be E_i. However, it was stated earlier in this section that we can only know (with absolute certainty) the energy of the eigenfunctions of an observable. Yet Ψ is not an eigenfunction of the energy operator. What happens is that, after measurement, the wavefunction Ψ "collapses" into the wavefunction ψ_{E_i} corresponding to the observed energy E_i. Remember that measurement changes the state of a quantum particle. Now we can say that *measuring an observable leads us to observe one of the eigenvalues of that observable, and the state of the quantum particle being measured changes to the corresponding eigenfunction*.

We now face the problem of finding the probability of this "collapse" into a particular eigenfunction. Linear algebra provides us with a very handy solution to this problem. Any general wavefunction can be written in terms of the eigenfunctions of an operator. Let us clarify this point. Every operator has a set of eigenfunctions. If we use a linear combination of all these eigenfunctions, we get a set of states that includes every possible state that the quantum particle can have. That is, every wavefunction can be decomposed into a linear superposition of the eigenfunctions of any given operator. Let us once again go back to the energy operator and its eigenfunctions that we discussed in Equation (1.9). *Any* general wavefunction Ψ can be written in terms of the eigenfunctions ψ_{E_i} of the energy operator as

$$\Psi = \Sigma a_i \psi_{E_i} \tag{1.10}$$

where a_i are coefficients corresponding to every wavefunction ψ_{E_i} and are complex numbers.

Unlike in the case of the eigenfunctions ψ_{E_i}, we cannot discuss the energy of this general particle as having wavefunction Ψ. It is a linear superposition of states ψ_{E_i} of different energies E_i. When the energy of this particle is measured, one obtains any one of the energies E_i, and the particle is found to be in the state ψ_{E_i} after the measurement. However, we cannot say that the particle had energy E_i because the measurement may very well have led to the observation of a different energy E_j. Now the state of the particle changes after measurement from