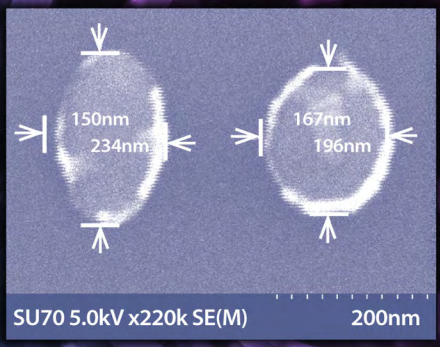
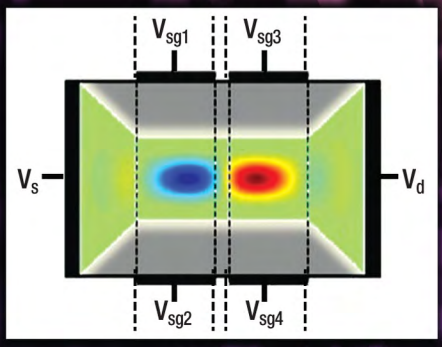
SECOND EDITION

Introduction to SPINTRONICS





Supriyo Bandyopadhyay
Marc Cahay

CRC Press
Taylor & Francis Group

SECOND EDITION

Introduction to SPINTRONICS

This page intentionally left blank

SECOND EDITION

Introduction to SPINTRONICS

Supriyo Bandyopadhyay
Marc Cahay

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

The cover artwork shows two figures. On the left is a scanning electron micrograph of two closely spaced nanomagnets—one more shape-anisotropic than the other – which together act as a nanomagnetic inverter or NOT gate because of their mutual dipole interaction. These nanomagnets are made of FeGa alloy and have been delineated on a piezoelectric substrate with electron beam lithography at Virginia Commonwealth University. The magnetizations of both nanomagnets are bistable and can orient along either of the two mutually anti-parallel directions along the ellipses' major axes. These two orientations encode the binary bits 0 and 1. The magnetization of the more shape anisotropic nanomagnet encodes the input bit of the inverter and that of the other encodes the output bit. Because of dipole coupling between them, the magnetizations of the two nanomagnets will be anti-parallel in the ground state, which makes the output bit the logic complement of the input bit, thus realizing a nanomagnetic inverter. Micrograph is provided by Hasnain Ahmad.

On the right is a schematic of an all-electric spin polarizer based on a dual quantum point contact formed in a two-dimensional electron gas. A set of four in-plane side gates is used to control the amount of spin polarization in the narrow portion of the device. The blue and red regions represent accumulations of spin-down and spin-up electrons. The spin polarization configuration can be altered by changing the bias configurations on the four side gates. The onset of spin polarization is accompanied by the presence of anomalies in the conductance of the dual quantum point contact. This figure is the result of simulations and theoretical calculations carried out at University of Cincinnati.

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20150821

International Standard Book Number-13: 978-1-4822-5557-7 (eBook - PDF)

S. B. dedicates this book to Bimalendu, Bela, Anuradha and Saumil Bandyopadhyay.

M. C. dedicates this book to his wife Janie, thanking her for her patience and encouragement, and to the memory of his sister Michèle.

This page intentionally left blank

Contents

This page intentionally left blank

Preface

This is a textbook intended to introduce a student of engineering, materials science and/or applied physics to the field of *spintronics*. While the term "spintronics" may have different connotations for different people, in this textbook it deals primarily with the science and technology of using the spin degree of freedom of a charge carrier to store, encode, access, process and/or transmit information in some way. That role had been traditionally delegated to the "charge" of an electron, not its "spin." Over the last two decades or so, there has been burgeoning interest in augmenting the role of charge with spin, or even replacing charge with spin in information processing devices.

Interest in spintronics was motivated by a longstanding tacit belief that replacing charge with spin may yield some advantages in terms of increased processing speed, lower power consumption, and/or increased device density on a chip. While this may not always be true, there are some scenarios where it *may* become true in the near future. In this textbook, we place particular emphasis on identifying situations where "spin" may have an advantage over "charge" and where it may not (see, in particular, Chapters 13–15).

The advent of quantum computing has added a new dimension to all this. The spin polarization of a single electron can exist in a coherent superposition of two orthogonal spin polarizations (i.e., mutually anti-parallel spin orientations) for a relatively long time without losing phase coherence. The charge degree of freedom, on the other hand, loses phase coherence much faster. Therefore, spin has become the preferred vehicle to host a quantum bit (or "qubit"), which is a coherent superposition of two orthogonal states of a quantum mechanical entity representing classical logic bits 0 and 1. The potential application of spin to scalable quantum logic processors has a short history, but has provided a tremendous boost to spintronics.

This textbook is expected to equip the reader with sufficient knowledge and understanding to conduct research in the field of spintronic devices, particularly semiconductor-based spintronic devices. We assume that readers have first-year graduate-level knowledge of device engineering, solid state physics, and quantum mechanics.

The first edition of this book was organized into fifteen chapters and the second edition contains eighteen chapters. The first chapter provides a historical perspective to those who have had little or no exposure to this field. It traces the early history of spin, the anomalous Zeeman effect, and ends with an account of the accidental discovery of "spin" by Stern and Gerlach in 1922.

Chapter 2 introduces the quantum mechanics machinery needed to under-

stand spin physics, as well as analyze spin transport and general spin dynamics in solid state structures. It also introduces the concept of Pauli spin matrices, the Pauli equation, and finally its relativistic refinement—the Dirac equation. Since, in this textbook, we will never encounter any situation where relativistic corrections become important, we will not have any occasion to use the Dirac equation. The Pauli equation will be sufficient for all scenarios. Nonetheless, it is important to gain an appreciation for the Dirac equation, since the quantum mechanical nuances associated with spin cannot be absorbed without an understanding of Dirac's seminal work.

Chapter 3 introduces the Bloch sphere concept, since it is a very useful tool to visualize the dynamics of a spin-1/2 particle (e.g., an electron), or qubit encoded in the spin of an electron, under the action of external magnetic fields. Applications of the Bloch sphere concept are elucidated with a number of examples. A spinor, representing an electron's spin, is viewed as a radial vector in the Bloch sphere, and this serves as a nice visualization tool for students interested in quantum computing and other applications of spintronics. All coherent motions of the spinor (where spin does not relax) are essentially excursions on the surface of the Bloch sphere.

Chapter 4 deals with an important application of the Bloch sphere concept, namely, the derivation of Rabi oscillation and the Rabi formula for coherent spin rotation or spin flip. These have important applications in many spin-related technologies such as electron spin resonance spectroscopy, nuclear magnetic resonance, and ultimately solid state versions of quantum computing. This chapter is somewhat mathematical and "seasons" the student to deal with the algebra (and recipes) necessary for calculating quantities that are important in spintronics. This chapter can be skipped at first and revisited later.

Chapter 5 introduces the concept of the "density matrix," pure and mixed states, Bloch equations (that describe the temporal relaxation of spin), the Bloch ball concept, and the notion of the longitudinal (T_1) and transverse (T_2) relaxation times. Several numerical examples are also presented to strengthen key concepts. Since here we allow the dynamics of spin to be incoherent, the motion of the spinor is no longer constrained to the Bloch sphere. The "Bloch sphere" actually refers only to the surface of the sphere and excludes the interior. If spin relaxes so that the norm of the sphere's radius is no longer conserved, then we have to allow excursions into the interior of the Bloch sphere. Therefore, we extend the Bloch sphere concept to the "Bloch ball" concept. This chapter contains advanced concepts and may also be skipped at first reading.

Chapter 6 introduces the rather important topic of spin–orbit interaction which is at the heart of many spintronic devices, since it offers a "handle" to manipulate spins. We introduce the general notion of spin–orbit interaction and then focus on the two special types of spin–orbit interactions that are predominant in the conduction band of most direct-gap semiconductors: the Rashba interaction arising from structural inversion asymmetry, and the

Dresselhaus interaction arising from bulk (crystallographic) inversion asymmetry. These two interactions form the basis of spintronic field effect transistors where the current flowing between two of the transistor's terminals is modulated by influencing the spin–orbit interaction in the device via a potential applied to the third terminal. Therefore, it is particularly important for applied physicists, materials scientists, and engineers to understand these interactions.

In Chapter 7, we derive the electron dispersion relations (energy versus wavevector) of electrons in quasi two- and one-dimensional structures (quantum wells and wires) in the presence of Rashba and Dresselhaus spin–orbit interactions, as well as an external magnetic field. We also derive the spin eigenstates, which allows us to deduce the spin polarization of carriers in any band. All this is accomplished by solving the Pauli equation. This is an example of how the Pauli equation is applied to solve a real life problem. We place special emphasis on how the dispersion relations are modified by an external magnetic field. This is important since it ultimately helps the student to appreciate how an external magnetic field can affect the performance of spin-based devices.

Chapter 8 discusses spin relaxation of conduction electrons in metals and semiconductors. We focus on four primary spin relaxation mechanisms: the D'yakonov–Perel, the Elliott–Yafet, the Bir–Aronov–Pikus and hyperfine interactions with nuclear spins, since these are dominant in the conduction band of semiconductors and therefore are most important in device contexts. Because spin relaxation limits the performance of most, if not all, spin-based devices, it is a vital issue. Ultimately, the aim of all device engineers and physicists is to reduce the rate of spin relaxation in spin devices, in order to make them more robust and useful. Spin relaxation also has peculiarities that are completely unexpected and without parallel in solid state physics. We present one example where spin can relax in *time* but not in *space*.

Chapter 9 is a new addition to the second edition and was not included in the first. Since the publication of the first edition in 2008, there has been an explosion in the study of spin-related physical phenomena, particularly those that may have applications in spintronic devices. In this chapter, we also discuss seven important physical phenomena that all have device applications: the extrinsic and intrinsic spin Hall effect along with the inverse spin Hall effect and the giant spin Hall effect, the spin Hanle effect, the spin capacitor effect, the spin-torque effect, the spin Galvanic effect, the spin Seebeck effect, and the inverse spin Seebeck effect or spin Peltier effect.

Chapter 10 introduces the more advanced concepts of exchange and spin–spin interaction. These form the basis of ferromagnetism and also the basis of single-spin computing schemes that are dealt with in Chapters 15 and 16.

Chapter 11 is an introduction to spin transport in solid state structures in the presence of spin relaxation. We focus on two basic models: the drift–diffusion model of spin transport and the semi-classical model that goes beyond the drift–diffusion model. The "spin" drift–diffusion model is very

similar to the "charge" drift–diffusion model applied to bipolar transport; the "up-spin" and "down-spin" carriers assume roles analogous to electrons and holes. However, it has limitations. One limitation that we emphasize with specific examples is that it fails to describe essential features of spin transport, even qualitatively, if electrons are traveling "upstream" against the force exerted on them by an electric field. In this chapter, we present many examples of how spin relaxes in time and space in quasi one-dimensional structures in the presence of the D'yakonov–Perel' spin relaxation mechanism, since it is usually the dominant mechanism for spin relaxation in semiconductor structures. These examples are based on the semi-classical model and therefore applicable to both low field transport and high field (hot electron) transport. The semiclassical model is based on combining the Liouville equation for the time evolution of the spin density matrix with the Boltzmann transport equation for time evolution of the carrier momentum in the presence of scattering and external electric fields.

In Chapter 12, we discuss *passive* spintronic devices such as spin valves and devices based on the giant magnetoresistance effect. Most commercial spintronic products that are currently available (magnetic read heads for reading data in computer hard disks or entertainment systems such Apple iPods, and magnetic random access memory) utilize these passive devices. Therefore, an adequate understanding of these devices is vital for engineers. We also discuss the important notions of spin injection efficiency, spin extraction, and the recently discussed spin blockade. This is a long chapter with many topics and it is intended to introduce the reader to important concepts encountered in the modern spintronics literature. We also discuss three very specific devices that are spintronic "sensors"; one is a magnetic field sensor, another is a light sensor (photodetector), and the third is a mechanical strain sensor. These are discussed to show the reader that spintronics has myriad applications in magnetics, mechanics and optics.

Chapter 13 introduces *active* spintronic devices, such as spin field effect transistors and spin bipolar transistors. We explain the physical basis of how these devices operate and what their shortcomings are. We make a simple estimate of their performance figures in order to project a realistic picture of whether they are or are not competitive with traditional electronic devices that are currently extant. Regardless of their actual device potential, these devices are standard-bearers that aroused early interest in the field among engineers and applied physicists. This chapter discusses only the early variants of spin transistors because of their pedagogical importance. New twists to spin transistors appear in the literature frequently and it was not possible to do justice to them within the limited space available. The only way the reader can keep pace with this field is to follow the literature closely.

Chapter 14 discusses the recently discovered field of "spintronics without magnetism," which allows one to manipulate spin currents by purely electrical means. The reader is introduced to lateral spin–orbit interaction, and its many nuances, and the possibility of implementing spin polarizers and analyzers

using quantum point contacts. This too is a new addition to the second edition.

Chapter 15 introduces more exotic concepts dealing with single-spin processors. Here, a single electron spin acts as the primitive bistable "switch" with two stable (mutually anti-parallel spin orientation) states that encode classical logic bits 0 and 1. Switching between the bits is accomplished by flipping the electron's spin without moving the electron in space and causing current flow. This chapter addresses fundamental notions like the ultimate limits of dissipation in performing Boolean logic operations and has relevance to the celebrated *Moore's law* scaling. Another distinguishing feature is that this chapter addresses spin-based architectures and not discrete devices like transistors. For example, it describes combinational logic circuits implemented with single-spin-switches that communicate with each other via exchange interaction and not physical wires. This is an area that has remained neglected, but is really no more challenging than spin-based quantum computing, since phase coherence of spin is not required. Being classical, it does not have the promise of quantum speedup of computation, or the ability to solve intractable problems, but it may provide valuable insights into the limits of classical computation.

Chapter 16 is an introduction to the field of spin-based reversible logic gates (that can, in principle, compute without dissipating energy) and spintronic embodiments of quantum computers. This is a rapidly advancing field, extremely popular among many spintronic researchers, and discoveries are made at a fast pace. This chapter is written mostly for engineers and applied physicists (not computer scientists or theoretical physicists), and should provide them with the preliminary knowledge required to delve further into this field. We have also focused on electrical manipulation of spin qubits rather than optical manipulation since this book is almost entirely devoted to electro-spintronics rather than opto-spintronics. Needless to say, because of the rapid advances in this area, it is impossible to address this field comprehensively. The reader is provided with a few examples to whet her/his appetite and is urged to follow the literature closely to keep abreast of the most recent developments.

Chapter 17 introduces the concept of "single-domain-nanomagnet" based computing and is a more practical rendition of the single-spin logic architecture ideas of Chapter 15. This is a new addition to the second edition. In a single domain ferromagnet, all the spins rotate in unison under an external influence because of strong exchange interaction between spins, making the entire ferromagnet behave like a giant classical spin. This chapter is focused primarily on logic architectures and discusses two main variants: dipole coupled nanomagnetic logic (also known as magnetic quantum cellular automata) and magneto-tunneling junction logic. Particular emphasis is placed on various magnet switching methodologies (magnetic field, spin-torque, spin-Hall effect, toplogical insulators, and magneto-elastic switching) since they determine the energy efficiency of nanomagnetic architectures. Much of the mate-

rial presented in this chapter, dealing with magneto-elastic devices, was the result of collaborative research with Prof. Jayasimha Atulasimha at Virginia Commonwealth University.

Chapter 18 is a stand-alone chapter that can be treated as an appendix. At first sight, it will appear unrelated to spintronics, which it is, but it has been included for a reason. There are many instances in this book when a student will have to recollect or refamiliarize herself/himself with some key results of quantum mechanics. Rather than making a trip to the nearest library, it would be more convenient to have a "quantum mechanics primer" handy where these key results have been re-derived. This chapter is included for completeness and comprehensiveness. The reader can refer to it if and when necessary.

At the time of writing the second edition, this book is still the only known "textbook" in spintronics written in English. By its very nature, it must be incomplete and omit many topics that are both important and interesting. We have focused mostly on electron spin, and, with the sole exception of discussing hyperfine nuclear interactions, we have ignored nuclear spin altogether. Hence, we do not discuss such well-known phenomena as the Overhauser effect, which is more relevant to nuclear spin. Another area that we have intentionally not covered in any detail is *organic spintronics*. We omitted any discussion of this field (it is still in its infancy) and do not discuss it primarily because we feel that this is very much in evolution. Organic semiconductors (mostly hydrocarbons) have weak spin–orbit interactions, so that spin relaxes slowly in these materials compared to inorganic semiconductors. Hence, they have a major advantage over inorganics when it comes to applications where spin relaxation must be suppressed, such as in spin-based classical or quantum computing. Some reviews have appeared in the literature covering organic spintronics and an edited book is available from this publisher.

This textbook also heavily emphasizes transport phenomena as opposed to optical phenomena dealing with the interaction of polarized photons with spin-polarized electrons and holes. Hence, we do not discuss such devices as spin-light-emitting diodes. Delving into "opto-spintronics" would have easily added a couple hundred pages to the 600-odd pages in this textbook. Our own expertise is more in transport phenomena, which has led us to focus more on transport. However, there are many excellent books (although not necessarily "textbooks") available that deal with opto-spintronics, and the interested reader can easily find an assortment of literature in that area.

Table of Universal Constants

Free electron mass (m_0)	9.1×10^{-31} kilograms
Dielectric constant of free space (ϵ_0)	8.854×10^{-12} Farads/meter
Electronic charge (e)	1.61×10^{-19} Coulombs
Reduced Planck constant ($\hbar$)	1.05×10^{-34} Joules-sec
Bohr radius of ground state in H atom (a_0)	0.529 Å$= 5.29 \times 10^{-11}$ meters
Bohr magneton (μ_B)	9.27×10^{-24} Joules/Tesla

Acknowledgments

Some acknowledgments are due. Many of our associates have contributed indirectly to this book. They are our students, laboratory interns and post-doctoral research associates, past and present. They include Prof. Sandipan Pramanik of the University of Alberta, Canada, who was a graduate student and then a post-doctoral researcher at Virginia Commonwealth University (VCU) at the time this book was composed; Dr. Bhargava Kanchibotla, an erstwhile graduate student at VCU who provided the T_2 data in cadmium sulfide nanostructures from his experiments (Figure 8.5); Dr. Sivakumar Ramanathan and Dr. Sridhar Patibandla, two graduate students at VCU working with S. B. who took the first spintronics graduate course offered at VCU by S. B. and provided valuable feedback; Harsh Agarwal, a summer undergraduate intern visiting VCU from the Banaras Hindu University Institute of Technology, Varanasi, India, who computed and generated some of the plots in Chapter 15; and Dr. Amit Trivedi, another summer undergraduate intern from the Indian Institute of Technology, Kanpur, who performed some of the calculations; and Saumil Bandyopadhyay, who performed some of the coherent room temperature spin transport experiments in single subband quantum wires.

At the University of Cincinnati, a former graduate student, Dr. Junjun Wan, and intern Lindsay Ficke, contributed immensely to the generation of data and plots. We also thank graduate student Nishant Vepachedu for proofreading the manuscript. We remain grateful to all of them.

In spite of all our best efforts, quite a few typographical errors made their way into the first edition. We corrected as many of them as we could catch, but some may have still eluded us. Matthew David Mower, a student from the University of Missouri, brought some typographical errors to our notice. We thank him. As always, we will remain grateful to any reader who points out such errors to us. Our e-mails are sbandy@vcu.edu and marc.cahay@uc.edu.

Welcome to the world of spintronics!

This page intentionally left blank

1

Early History of Spin

1.1 Spin

Most students of science and engineering know that every elementary particle, such as electrons, neutrons, photons, neutrinos, etc., has a quantum mechanical property called "spin" which can be measured (perhaps not easily, but at least in principle) and has a quantized value, including zero. The vast majority of these students mentally visualize spin as the angular momentum associated with the elementary particle spinning or rotating about its own axis (like a top or a planetary object). This mental picture, although convenient and comforting, is actually somewhat crude and certainly incomplete. Landau and Lifshitz, in their classic textbook on quantum mechanics [1], wrote "[the spin] property of elementary particles is peculiar to quantum theory. [It] has no classical interpretation... It would be wholly meaningless to imagine the 'intrinsic' angular momentum of an elementary particle as being the result of its rotation about its own axis."

The simplistic notion of self-rotation about an axis, shown in Fig. 1.1, cannot explain many features of spin, such as why its magnitude cannot assume continuous values and why it is quantized to certain specific values. It also causes serious problems if taken too literally. As we will see later (Problem 1.2), if we think of an electron as a solid sphere of radius equal to the Lorentz radius $e^2/(4\pi\epsilon_0 m_0 c^2)$ (where e is the electron's charge, m_0 is the mass, c is the speed of light in vacuum, and ϵ_0 is the dielectric constant of vacuum), then the velocity on the surface of a rotating electron would have to be many times the velocity of light in vacuum if such a rotation were to generate an angular momentum equal to the electron's spin. Obviously that would not be permitted by the theory of relativity. Indeed, a deep understanding of quantum mechanics is required to understand how the spin property comes about. Its origin is in relativistic quantum mechanics and really was first appreciated by Paul Andrew Maurice Dirac when he derived the Dirac equation which is the cornerstone of relativistic quantum mechanics. Richard Feynman, noted for his distaste for mysticism in physics, wrote about the notion of "spin": "It appears to be one of the few places in physics where there is a rule which can be stated very simply, but for which no one has found a simple and easy explanation. The explanation is down deep in relativistic quantum mechanics.

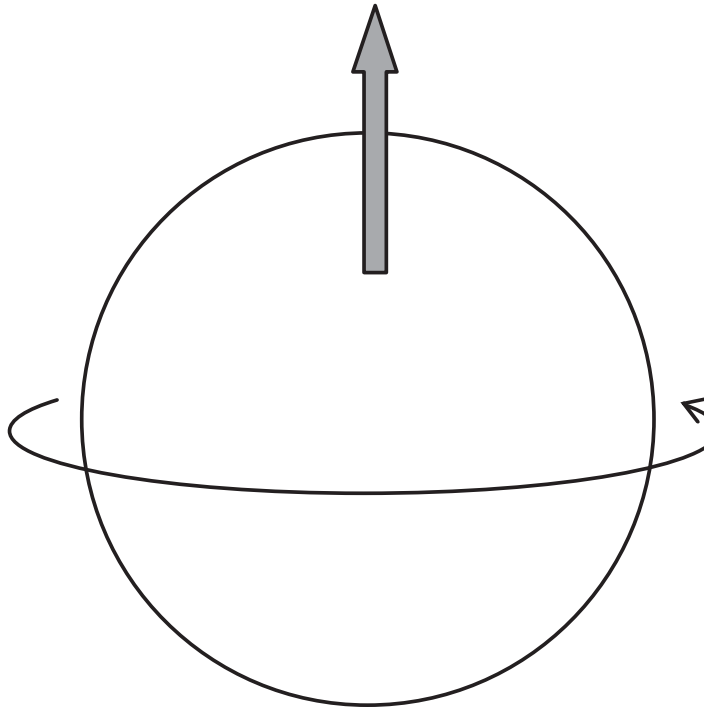


FIGURE 1.1
A mental picture to visualize spin as the angular momentum associated with self-rotation of a particle about an axis. This picture, although certainly incomplete and heuristic, is adequate for many situations that the reader will encounter in this book.

This probably means that we do not have a complete understanding of the fundamental principle involved" [2]. We do not intend to mystify "spin" any more than it already is, but rather we wish to underscore the fact that spin is *not* a classical property easily explained in terms of everyday experience. It is a property associated with relativistic quantum mechanics, for which there is no everyday experience. This is the only place in this textbook where we will mention this, since in the rest of this textbook we will have no occasion to reflect on this truism. We will be primarily involved with practical and applied aspects of spin, without relishing the fact that it is indeed an exotic property far outside what we normally deal with in the applied sciences and engineering.

1.2 Bohr planetary model and space quantization

The history of how the concept of spin was established is somewhat tortuous. In 1913, when Niels Bohr published his theory of the hydrogen atom, he had thought of a planetary model where the electron orbits around the nucleus as shown in Fig. 1.2. The radii of the allowed orbits are quantized. Each orbit has a radius $n^2 a_0$ where a_0 is the Bohr radius of the ground state in the hydrogen atom ($=4\pi\epsilon_0\hbar^2/m_0e^2$), which is 0.529 Å. Here $\hbar$ is the reduced Planck constant, m_0 is the electron's mass, ϵ_0 is the permittivity (dielectric constant) of free space, and e is the magnitude of its charge. The quantity n is an integer called the "principal quantum number," and it takes positive non-zero values 1, 2, 3, etc. Later Arnold Sommerfeld (in 1916), and independently Peter Debye, introduced two more quantum numbers l and m, which were called the "orbital" and "magnetic" quantum numbers. While the principal quantum number determines the radius (or size) of the orbit, the angular quantum number l determines its shape. It also determines the angular momentum associated with the orbital motion in units of $\hbar$. The integers n and l obey the relation $n \geq l$. If the atom is placed in a magnetic field, the component of its angular momentum along the field takes on quantized values of $m\hbar$. The number m is an integer and satisfies the inequality $-l \leq m \leq l$. This last inequality limits the number of m values to $2l + 1$, and accordingly, the number of allowed directions of the angular momentum vector in a magnetic field is $2l + 1$. This is known as *space quantization of angular momentum.*

The energy of an electron in an atom was thought to be determined by the three quantum numbers n, l and m. When an electron makes a transition from one energy state to another, the transition involves a change in one or more of these quantum numbers. In the process of transition, the electron absorbs or emits light of a particular frequency ν which is determined by the energy difference between the initial and final states in accordance with conservation of energy:

$$E_{final} - E_{initial} = 2\pi\hbar\nu = h\nu, \tag{1.1}$$

where $E_{initial}$ and E_{final} are the electron's energy in the initial and final states, respectively.

However, when an atom is placed in a magnetic field and the spectra of emitted and absorbed light are measured, it is found that the multiplicity of the spectra (meaning all the observed frequencies) cannot be explained by the space quantization rules (that means allowed values of n, l, and m) alone. In 1920, Sommerfeld tried to explain the multiplicity by invoking yet another quantum number j that he called the *inner* quantum number. However, this was not able to completely explain multiplicity. Additional frequencies (where each line split into two) were observed in a strong magnetic field; this was

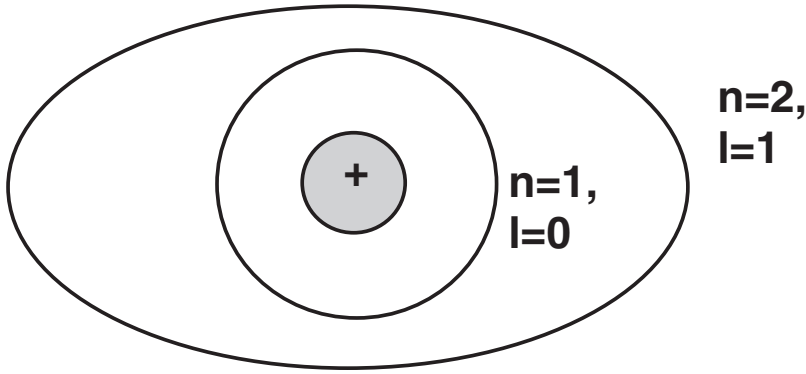


FIGURE 1.2
The Bohr planetary model of electron orbits around the nucleus of an atom (not to scale).

referred to as the anomalous Zeeman effect. This effect could not be explained by space quantization.

1.3 Birth of "spin"

In 1925, the anomalous Zeeman effect led a young American scientist, Ralph de Laer Kronig, to postulate that in addition to the orbital angular momentum, an electron has an additional angular momentum caused by spinning about its own axis. Kronig postulated that the angular momentum associated with this self-rotation has a fixed magnitude of $(1/2)\hbar$, although he obviously had no explanation as to why the angular momentum should have a fixed value, let alone why that value should be $(1/2)\hbar$ (such is the nature of a "postulate"). The correct explanation for these two features had to wait for Dirac's formulation, which is discussed later. However, such an angular momentum will cause multiplicity in atomic spectra due to a relativistic effect. The rotating electron in an atom's orbit experiences an electric field due to the positively charged nucleus, which will transform into a magnetic field $\vec{H}$ via Lorentz transformation in the rest frame of the electron. The angular

momentum will give rise to a magnetic moment $\vec{M}$ related to the angular momentum $\vec{W}$ by the Landé relation $\vec{M} = g_0\vec{W}$, where g_0 is the so-called Landé g-factor (Kronig assumed it to be 2, which is the correct value for vacuum). The magnetic moment will interact with the magnetic field and the energy of that interaction will be $E_{int} = -\vec{M} \cdot \vec{H}$. Since the angular momentum has two possible values $\pm(1/2)\hbar$, this will result in energy splitting of $2\left|\vec{M} \cdot \vec{H}\right|$. With this additional angular momentum, he was able to explain the multiplicity of spectra (within a factor of 2), and presented his idea to Wolfgang Pauli, who was not impressed by it (Pauli had his own ideas to explain the multiplicity of spectra, which later turned out to be wrong). Kronig himself was not very confident because his calculations based on this spinning electron model still did not completely explain every feature of the experimentally observed spectra. More important, the idea of a self-rotating electron presents a conundrum within the framework of classical theory. If one thinks of the electron as a sphere of radius $r_e = e^2/(4\pi\epsilon_0 mc^2)$, as considered by Lorentz, then the rotation rate required to produce an angular momentum of $(1/2)\hbar$ is so high that the electron's surface reaches a speed more than 130 times that of light in vacuum, in stark violation of Einstein's theory of relativity! Because of this apparent contradiction, Kronig never published his ideas. Six months later, Uhlenbeck and Goudsmit published essentially the same spinning electron idea that Kronig had come up with in the journal *Naturwissenschaften*. Actually, when they realized the problem with the surface speed (see Problem 1.2), they tried to hurriedly withdraw their paper, but it was too late. The paper appeared in print. Kronig sent a letter to the British journal *Nature* criticizing the idea of Uhlenbeck and Goudsmit, pointing out the problems with the spinning electron model. Meanwhile, in a second paper that appeared in *Nature*, Uhlenbeck and Goudsmit pointed out that their theory did not quite explain the experimental observations of atomic spectra. In fact, there was a discrepancy by a factor of 2 with the experimental results. Later L. H. Thomas showed that this discrepancy comes about because of an incorrect definition of the electron rest system. An electron in an atom moves in a closed orbit around the nucleus. Therefore, there is always a component of the field perpendicular to the instantaneous velocity which causes an additional acceleration in the direction perpendicular to the velocity. This is what causes the electron trajectory to be curved. As a result, the electron is moving in a rotating frame of reference and when this is ocrrectly taken into account, it leads to the "factor of 2 correction". Thomas published his findings in a letter to *Nature* in February 1926, following which, all discrepancies between theory and experiment could be resolved. This made a convert out of Pauli, who ultimately endorsed the spinning electron idea of Kronig, Uhlenbeck and Goudsmit (KUG).

The spinning electron picture remains in vogue, although it obviously raises unanswered questions, such as why the spin angular momentum is quantized, or why the idea cannot be reconciled with the postulates of the special theory

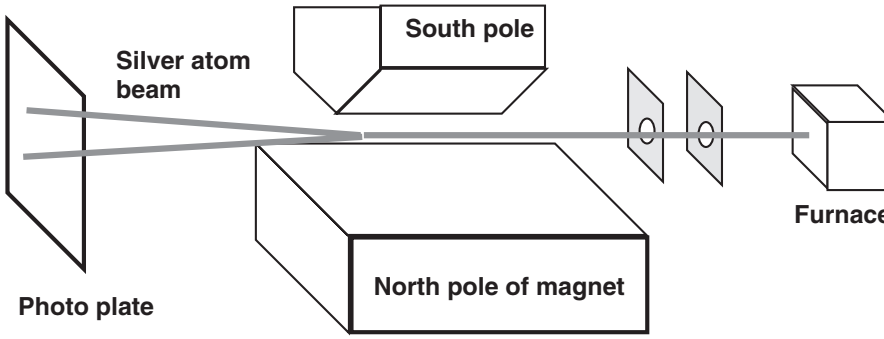


FIGURE 1.3
Experimental arrangement of the Stern-Gerlach apparatus (not to scale).

of relativity. These unanswered questions merely reveal the inadequacy of the model. The reader should be aware of this inadequacy, but that should not affect her/his understanding of any of the topics discussed in this book.

1.4 The Stern-Gerlach experiment

At least three years before the KUG ideas were published, the spin of an electron was already *unwittingly* measured in the famous Stern-Gerlach experiment, which remains a watershed event in the history of spin. Otto Stern and Walther Gerlach in Frankfurt, Germany, took the space quantization business literally and designed an experiment to verify it. To them, the space quantization business involved only the quantum numbers m, l, and n, but no spin, since the spin idea had not yet been thought of by KUG. Space quantization will dictate that the angular momentum of an electron in the Bohr atom will be quantized to $l\hbar$, and, because an orbiting electron will give rise to a magnetic moment proportional to the angular momentum, the magnetic moment will be quantized as well. Therefore, by measuring the magnetic moment of atoms and showing that it can assume only discrete values, it should be possible to demonstrate space quantization.

Consider a beam of atoms coming out of an effusion furnace, going through collimators and a magnetic field, and finally impinging on a detection plate, as shown in Fig. 1.3. Assuming that an atom possesses a non-zero magnetic moment $\vec{\mu}$ due to the oribital angular momenta of its electrons, the magnetic

field will produce two effects. First, the magnetic field will exert a torque on the magnetic dipole of the atom and make the magnetic moment vector precess about the magnetic field [3]. Second, the potential energy of the atom will be $U = -\vec{\mu} \cdot \vec{B} = -\mu_z B$, if we assume that the magnetic field is directed along the z-direction. Because of the spatial *non-uniformity* of the magnetic field in the z-direction, the atoms will experience a z-directed force given by [3]

$$F_z = -\frac{\partial U}{\partial z} = \mu_z \frac{\partial B(z)}{\partial z}. \tag{1.2}$$

Classically, one would expect that, owing to random thermal effects, the atoms in a beam will have their magnetic moment vectors pointing in different directions, resulting in a continuous spread in the value of μ_z from $-|\mu_z|$ to $+|\mu_z|$. This will cause a continuous spread in F_z and therefore a continuous spread in the deflection of the beam in the z-direction. If the beam then impinges on a detection plate, one should see a continuous line on the plate in the z-direction.

However, if space quantization holds sway, then the magnetic moment μ_z is quantized and cannot assume any arbitrary value between -$|\mu_z|$ and +$|\mu_z|$. It is not clear why he thought this, but Stern was convinced that the magnetic moment in the direction of the magnetic field (i.e., μ_z) will be quantized to two values of opposite sign and therefore every atom will experience the same magnitude of the deflecting force, except some will experience it in one direction and the rest in the opposite direction. Therefore, an atomic beam of hydrogen would split into two beams in a magnetic field, and, despite the smearing effect of the inter-atomic collisions that exert a random force on the atoms, they should split so far apart in a strong field that the oppositely directed components would be deflected outside the width of the original beam. If that happens, then one should observe *two* distinct beams that will not produce a line on the detecting plate, but just two spots. This would have been a remarkable experiment anyway, since classical theory predicted a line. Thus, any birefringence (and observation of just two distinct spots) would demonstrate that quantum physics supersedes classical physics, and it would be a *literal* demonstration of space quantization, which is a quantum mechanical phenomenon outside the realm of classical physics.

Bohr himself was not so convinced of his own model of the atom and did not think one could take quantum physics so literally. He thought that space quantization was a symbolic expression and convenient for calculating atomic spectra, but could not be taken literally. However, Stern was fortunately not dissuaded by Bohr and found a convert in Gerlach, who, until that time, had apparently not heard of space quantization!

The actual Stern-Gerlach experiment took place a year after it was conceived by Stern. It was not hydrogen atoms, but a beam of silver atoms produced by the effusion of the metal from an oven heated to 1000°C, that was collimated by two narrow slits 0.03 mm wide. It traversed a magnet 3.5

cm long that produced a magnetic flux density of 0.1 Tesla and a field gradient of 10 Tesla/cm. The calculated splitting of the beam was only 0.2 mm so that extreme alignment accuracy was required. When the experiment was completed, Stern and Gerlach could not even see the traces of the silver beam on the collector plate, so that it was impossible to determine if the splitting was there. In an episode that has now become famous, Stern's breath on the plate made history. He used to smoke cigars and the sulfur from his breath turned the silver into silver sulfide, which is jet black and is easily visible. The results of this experiment, however, were still inconclusive and would not have convinced a sceptic. After many efforts, Stern and Gerlach met in Göttingen in 1922 and decided to give up. But a railroad strike detained Gerlach and he attempted the experiment once again with improved accuracy of alignment. Finally a clear signature of beam splitting was observed [4].

We now know that the Stern-Gerlach experiment had nothing to do with space quantization, contrary to what they believed. In fact, Gerlach could never reproduce the experiment with sodium atoms instead of silver atoms, and Einstein questioned their interpretation (correctly). The agreement of the experiment with Stern's and Gerlach's picture of space quantization (or the Bohr model) was nothing but a coincidence. The net angular momentum of silver atoms is actually *zero*, contrary to what was presumed by Stern and Gerlach. The magnetic moment of Silver atoms is therefore due solely to the *spin angular momentum* (and has nothing to do with the orbital angular momentum), which accounts for the observed splitting of the beam into two. Stern and Gerlach had *unknowingly measured the spin angular momentum.* This is the first convincing experimental observation of "spin."

It is surprising that, even though the Stern-Gerlach experiment was reported in 1922 and was widely known among physicists, the postulation of the electron spin in 1925 did not immediately lead to a re-interpretation of the experiment as being a demonstration of the spin. The earliest attribution of the splitting to spin was reported in 1927, when Ronald Fraser noted that the ground state orbital angular momentum and associated magnetic moments of silver, hydrogen and sodium atoms are zero. Therefore, obviously the splitting could not be due to orbital angular momentum; it had to be due to "spin".

In February, 2002, a team of physicists and chemists tried to re-enact the Stern-Gerlach experiment with sulfur tainted breath and all [5]. For a hilarious account of this re-enactment, see the December 2003 issue of the magazine *Physics Today.* Mere sulfurous breath was not sufficient to reveal traces of silver on a detecting plate, but direct exposure to cigar smoke was enough to reveal the traces.

1.5 Advent of spintronics

Although "spin" plays a fundamental role in explaining the multiplicity of atomic spectra, this is not its most important role. It was realized in the middle of the 20th century that spin plays a fundamental role in magnetism. Every theoretical model postulated to explain the physical origin of magnetism invoked "spin" in some way or the other. This included the Bloch model, the Heisenberg model, the Stoner model, and every other model advanced. While magnetism remains the domain of spin, in the late 20th century, it was realized that spin, alone or in conjunction with charge, can be harnessed to process information, particularly digital information encoded with binary bits 0 and 1. This is the central theme in the field of *spintronics* today (see, for example, [6, 7]). Of course, spintronics is much more than just information rendition and includes the more traditional areas of *magnetoelectronics*, which deals with magnetic or magnetoresistive effects for sensing and storing information. Early successes in this area include the developments of read heads for sensing massively dense magnetic storage media (these read heads are now routinely used in laptop computers and entertainment systems such as Apple iPods), non-volatile magnetic random access memory (MRAM) [8], programmable spintronic logic devices based on magnetic tunnel junction elements [9], rotational speed control systems [10], positioning control devices in robotics and related systems (such as automobile braking systems) [11], perimeter defense systems, magnetometers, and high current monitoring devices for power systems [12], etc. Many of these developments were fueled by the investigation of how spin-polarized electric currents can be injected into ferromagnetic/paramagnetic multilayers which, in the 1980s, had led to the important discovery of the phenomenon of giant magnetoresistance (GMR) [13, 14]. Several books and review articles have been written on magnetoelectronics [15, 16, 17, 18, 19, 20, 21]. The reader is referred to these books to gain an understanding of magnetoelectronics. In this book, we will focus more on the "information technology" (IT) applications of spin, which are the basis of modern spintronics.

The application of spintronics in information technology, particularly for computing and signal processing, is a relatively new field. Early efforts in this area were concerned with developing spin-based analogs of conventional signal processing devices such as transistors - both the field effect type [22, 23] and the bipolar junction type [24, 25]. In Spin Field Effect Transistors (SPINFETs) or Spin Bipolar Junction Transistors (SBJTs), information is still processed by modulating the charge current flowing between two terminals via the application of either a voltage or a current at the third terminal. However, the process by which the voltage or current at the third terminal exercises control over device behavior is spin mediated. Therefore "spin" plays somewhat of a secondary role in these devices, while "charge" retains

the primary role. These devices are interesting and demonstrate how spin can play a role in information handling. They are discussed in Chapter 14.

The more radical branch of spintronics is what we call *single-spin spintronics*, where charge has no direct role and information is encoded entirely in the spin polarization of a single electron, which is made to have only two values, "up-spin" and "down-spin", by placing the electron in a magnetic field. The down-spin orientation could correspond to polarization parallel to the magnetic field and the up-spin orientation is anti-parallel to the magnetic field. These two values could represent binary bits 0 and 1. Boolean logic circuits can be fashioned from interacting single electron spins by properly engineering the interactions. An embodiment of this approach is the so-called "single spin logic" (SSL) idea [28, 29] discussed in Chapter 15. What makes this idea attractive is that bits can be flipped by simply toggling an electron's spin, without physically moving it in space and causing a current flow. This results in much reduced heat dissipation in the circuits.

The ultimate rendition of spintronic computing circuits is spin-based quantum computers, which dissipate no energy at all to complete a logic operation, since they operate on the basis of *reversible* quantum dynamics. Recently, there has been a great deal of interest in encoding a quantum bit (*qubit*) using the spin degree of freedom of a single electron confined in a quantum dot [30, 31, 32, 33] or bound to a donor atom [34] or housed in a nitrogen vacancy center in diamond [35, 36] to implement a quantum logic gate. The quantum mechanical phase coherence of "spin" is much longer lived than that of "charge"; consequently, spin is a natural choice for building solid state scalable quantum logic processors. Spin-based quantum computing is a rapidly expanding field of research endeavor and Chapter 16 in this textbook discusses some of the basic ideas in this field. Unfortunately, Chapter 16 will be almost certainly outdated very quickly, perhaps even by the time this book appears in print, since extremely rapid strides are being made in this field. Suffice it to say then that spintronics is now poised at a critical juncture where technological breakthroughs may be just around the corner. Therefore, understanding the science and technology of spintronics has become imperative for students of electrical engineering, physics, and materials science.

1.6 Problems

- **Problem 1.1**

  Using the Bohr model of the hydrogen atom, show that the magnetic moment associated with an electron moving in the lowest circular orbit

is the Bohr magneton given by

$$\mu_B = \frac{e\hbar}{2m_0} \ . \tag{1.3}$$

Show also that the magnetic flux density B_n associated with the n-th orbital is given by

$$\mu_B B_n = \frac{1}{4}(\frac{\alpha^4 Z^4}{n^5})m_o c^2 \ , \tag{1.4}$$

where m_o is the rest mass of the electron, c is the speed of light in vacuum, Z is the atomic number of the atom, n is the principal quantum number and α is the fine structure constant.

Calculate B_n in Tesla for the electron in the first orbit of the hydrogen atom, i.e.,, Z=1, n=1. Use $m_o c^2 \approx 0.5\texttt{MeV}$.

Solution

Consider a nucleus with Z protons. The total (kinetic + potential) energy of an electron moving around this nucleus is conserved and is given by:

$$E = \frac{1}{2}m_o v^2 - \frac{Ze^2}{4\pi\epsilon_o r} \ , \tag{1.5}$$

where e is the magnitude of the charge of the electron, v is the electron's orbital velocity, and ϵ_o is the dielectric constant of vacuum. The first term in the right hand side is the kinetic energy of the electron and the second term is the (electrostatic) potential energy. We ignore any gravitational potential energy since it is negligible.

According to Newton's Law, the centripetal force should be equal to the Coulomb force; hence, we have:

$$m_o \frac{v^2}{r} = \frac{Ze^2}{4\pi\epsilon_o r^2} \ . \tag{1.6}$$

Using Bohr's quantization condition for the angular momentum around the nth orbit:

$$L_n = m_o v_n r = n\hbar \ , \tag{1.7}$$

where the subscript n is the principal quantum number.

We get from Equations (1.6) and (1.7)

$$L_n v_n = \frac{Ze^2}{4\pi\epsilon_o} \ . \tag{1.8}$$

This yields

$$v_n = \frac{Ze^2}{4\pi\epsilon_o\hbar}\frac{1}{n} \ . \tag{1.9}$$

We can compare this velocity to that of light and get

$$\frac{v_n}{c} = \frac{Z}{n}\frac{e^2}{4\pi\epsilon_o\hbar c} \ , \tag{1.10}$$

where the constant $\frac{e^2}{4\pi\epsilon_o\hbar c} = \alpha$ and is referred to as the fine structure constant.

Equation (1.10) shows that an electron in the first orbit (ground state) of the hydrogen atom ($Z = n = 1$) has a velocity equal to $\alpha \approx \frac{1}{137}$ times the speed of light.

The radius of the various orbits allowed by Bohr's space quantization condition can then be obtained from Equation (1.7)

$$m_o v_n r_n = n\hbar \ . \tag{1.11}$$

We get

$$r_n = (\frac{4\pi\epsilon_o}{Ze^2})(\frac{n^2\hbar^2}{m_o}) \ . \tag{1.12}$$

This is also written as: $r_n = n^2 a_Z$, where a_Z is the effective Bohr radius given by

$$a_Z = \frac{4\pi\epsilon_o\hbar^2}{m_o Ze^2} \ . \tag{1.13}$$

The latter is equal to 0.529 $\mathring{A}$ for the hydrogen atom ($Z = 1$).

The total energy of an electron in the n-th orbit is found from Equation (1.5) and (1.9) to be

$$E_n = \frac{1}{2}m_o {v_n}^2 - \frac{Ze^2}{4\pi\epsilon_o r_n} = -\frac{m_o e^4 Z^2}{32{\epsilon_o}^2\hbar^2} \ . \tag{1.14}$$

This is the *binding energy* of the electron. It is also called the *ionization energy*, since this is the minimum amount of energy that will be required to liberate the electron from the nuclear attraction and make it free, thereby ionizing the parent atom. This ionization energy is 13.6 eV for the ground state of the hydrogen atom.

According to Biot-Savart's law, an electron moving around the nucleus in one of these orbits will feel a magnetic field given by

$$\vec{B} = \frac{\vec{\mathcal{E}} \times \vec{v}}{2c^2} \ , \tag{1.15}$$

where $\mathcal{E}$ is the electric field experienced by the moving electron and the factor 2 is due to Thomas's correction for a rotating frame of reference.

From Coulomb's law, the electric field in the nth orbital is given by

$$\mathcal{E}_n = \frac{Ze}{4\pi\epsilon_o {r_n}^2} \ . \tag{1.16}$$

Using Equations (1.9), (1.15), and (1.16), the magnitude of the magnetic field in the n-th orbit is

$$|\vec{B_n}| = \frac{\mathcal{E}_n v_n}{2c^2} = \frac{1}{2c^2}(\frac{Ze}{4\pi\epsilon_o {r_n}^2})v_n \tag{1.17}$$

and it is directed perpendicular to the orbital plane of the electron.

Finally, using Equations (1.12), (1.9) and (1.17) and the expression for the Bohr magneton given, one finds that

$$\mu_B B_n = \frac{1}{4}(\frac{\alpha^4 Z^4}{n^5})m_o c^2 \ . \tag{1.18}$$

From the above, the magnetic flux density associated with an electron orbiting the hydrogen atom in the first orbit is 6.16 Tesla.

- **Problem 1.2**

  Show that in the classical spinning electron model, the electron's surface speed must be more than 60 times the speed of light in order to produce an angular momentum of $(1/2)\hbar$. This is why, it is inappropriate to think in classical terms that the spin of an electron is associated with rotation about its own axis.

  Solution

  The angular momentum is

$$(1/2)m_0 v_s r_e = (1/2)\hbar \ , \tag{1.19}$$

  where v_s is the speed on the surface of the electron and r_e is the Lorentz radius of the electron. Solving the above equation with universal constants $m_0 = 9.1{\times}10^{-31}$ Kg, $\hbar = 1.05{\times}10^{-34}$ Joules-sec, and $r_e = e^2/(4\pi\epsilon_0 m_0 c^2) = 2.8$ femtometers,

$$v_s \approx 134c. \tag{1.20}$$

  Therefore, the speed of rotation on the surface of the electron is more than 130 times the speed of light.

What does this apparent fallacy imply? It tells us that the concept of spin is inherently quantum mechanical and cannot be described within the framework of classical mechanics as done here. Furthermore, the electron cannot be visualized as a nearly point charge with the Lorentz radius.

1.7 References

[1] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Non-relativistic Theory)*. 3rd. edition (Pergamon, New York), p. 198.

[2] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics*, Vol. 3, (Addison-Wesley, Reading, MA, 1965), p. 4-3.

[3] J. D. Cresser, *Quantum Physics Notes*, Macquarie University, Australia. Chapter 6.

[4] W. Gerlach and O. Stern, "Der experimentalle nachweis der richtung-squantelung in magnetic field", Z Phys., **9**, pp. 349-352 (1922).

[5] B. Friedrich and D. Hershbach, "Stern and Gerlach: How a bad cigar helped reorient atomic physics", *Physics Today*, **56**, pp. 53-59 (December 2003).

[6] D. D. Awschalom, M. E. Flatte and N. Samarth, "Spintronics", Scientific American, **286**, 66, June 2002.

[7] S.A. Wolf and D.M. Treger, "Special issue on spintronics", Proc. IEEE, **91**, 647 (2003).

[8] P.P. Freitas, F. Silva, N.J. Oliveira, L.V. Melo, L. Costa and N. Almeida, "Spin valve sensors", Sensors and Actuators A, **81**, 2 (2000).

[9] J. Wang, H. Meng and J.P. Wang, "Programmable spintronics logic device based on a magnetic tunnel junction element", J. Appl. Phys., **97**, 100509 (2005).

[10] P.P. Freitas, L. Costa, N. Almeida, L.V. Melo, F. Silva, J. Bernardo and C. Santos, "Giant magnetoresistive sensors for rotational speed control", J. Appl. Phys., **85**, 5459 (1999).

[11] W.J. Ku, P.P. Freitas, P. Compadrinho and J. Barata, "Precision X-Y robotic object handling using a dual GMR bridge sensor", IEEE Trans. Magn., **36**, 2782 (2000).

[12] J. Pelegri, J.B. Egea, D. Ramirez and P.P. Freitas, "Design, fabrication, and analysis of a spin-valve base current sensor", Sensors and Actuators A, **105**, 132 (2003).

[13] M.N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Eitenne, G. Creuzet, A. Friederich and J. Chazelas, "Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices", Phys. Rev. Lett., **61**, 2472 (1988).

[14] G. Binasch, P. Grünberg, F. Saurenbach and W. Zinn, "Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange", Phys. Rev. B, **39**, 4828 (1989).

[15] A detailed review of the work on giant magnetoresistance using magnetic multilayers is given by M.A.M. Gijs and G.E.W. Bauer, "Perpendicular giant magnetoresistance of magnetic multilayers", Adv. Phys., **46**, 285 (1997).

[16] S.S.P. Parkin, X. Jiang, C. Kaiser, A. Panchula, K. Roche and M. Samant, "Magnetically engineered spintronic sensors and memory", Proc. IEEE, **91**, 661 (2003).

[17] D.D. Awschalom and J.M. Kikkawa, "Electron spin and optical coherence in semiconductors", Physics Today, **52**, 33 (1999).

[18] *Magnetic Multilayers and Giant Magnetoresistance*, Ed. U. Hartman (Springer, Berlin, 2000).

[19] *Spin Electronics*, Eds. M. Ziese and M. J. Thornton, (Springer, New York, 2001).

[20] *Spin Dependent Transport in Magnetic Nanostructures*, Eds. S. Maekawa and T. Shinjo, (Taylor and Francis, New York, 2002).

[21] *Magnetic Interactions and Spin Transport*, Eds. A. Chtchelkanova, S. Wolf and Y. Idzerda, (Kluwer Academic Press, Dordrecht, 2003).

[22] S. Datta and B. Das, "Electronic analog of the electro-optic modulator", Appl. Phys. Lett., **56**, 665 (1990).

[23] S. Bandyopadhyay and M. Cahay, "Alternate spintronic analog of the electro-optic modulator", Appl. Phys. Lett., **85**, 1814 (2004).

[24] J. Fabian, I Zutic and S. Das Sarma, "Magnetic bipolar transistors", Appl. Phys. Lett., **84**, 85 (2004).

[25] M. E. Flatté and G. Vignale, "Unipolar spin diodes and transistors", Applied Physics Letters, Vol. 78, 1273 (2001); M. E. Flatté, Z. G. Yu, E. Johnston-Halperin and D. D. Awschalom, "Theory of semiconductor magnetic bipolar transistors", Appl. Phys. Lett., **82**, 4740 (2003); M. E. Flatté and G. Vignale, "Heterostructure unipolar spin transistors", J. Appl. Phys., **97**, 104508 (2005).

[26] S. Bandyopadhyay and M. Cahay, "Re-examination of some spintronic field effect device concepts", Appl. Phys. Lett., **85**, 1433 (2004).

[27] S. Bandyopadhyay and M. Cahay, "Are spin junction transistors useful for signal processing?", Appl. Phys. Lett., **86**, 133502 (2005).

[28] S. Bandyopadhyay, B. Das and A. E. Miller, "Supercomputing with spin polarized single electrons in a quantum coupled architecture", Nanotechnology, **5**, 113 (1994).

[29] S. Bandyopadhyay, "Computing with spins: from classical to quantum computing", Superlat. Microstruct., **37**, 77 (2005).

[30] S. Bandyopadhyay and V. P. Roychowdhury, "Switching in a reversible spin logic gate", Superlat. Microstruct., **22**, 411 (1997).

[31] A. M. Bychkov, L. A. Openov and I. A. Semenihin, "Single electron computing without dissipation", JETP Lett., **66**, 298 (1997).

[32] D. Loss and D.P. DiVincenzo, "Quantum computation with quantum dots", Phys. Rev. A, **57**, 120 (1998).

[33] S. Bandyopadhyay, "Self assembled nanoelectronic quantum computer based on the Rashba effect in quantum dots", Phys. Rev. B, **61**, 13813 (2000).

[34] B. E. Kane, "A silicon-based nuclear spin quantum computer", Nature (London), **393**, 133 (1998).

[35] M. V. G. Dutt, et al., "Quantum register basd on individual electronic and nuclear spin qubits in diamond", Science, **316**, 1312 (2007).

[36] J. R. Weber, W. F. Koehl, J. B. Varley, A. Janotti, B. B. Buckley, C. G. Van de Walle and D. D. Awschalom, "Quantum computing with defects", Proc. National Academy of Sciences of the United States of America, **107**, 8513 (2010).

2

Quantum Mechanics of Spin

In the 1920s, the old quantum theory was gradually being superseded by the new quantum theory. The cornerstone of the old theory was Bohr's model of the hydrogen atom, which predicted that an electron cannot orbit the proton in the hydrogen atom in any arbitrary fashion. Orbits are "quantized," meaning that only certain sizes, shapes, and magnetic properties are allowed. The principal quantum number n determined the allowed radii of the orbits, the orbital quantum number l determined the allowed shapes, and the magnetic quantum number m determined the magnetic behavior. Additionally, there is a fourth quantum number s which denotes the fact that the electron has an additional angular momentum, loosely associated with self rotation about its own axis, and that is quantized in units of $(1/2)\hbar$. The old quantum theory was useful to infer the existence of discrete energy levels in atoms, calculate energy spacings between these levels, and therefore allowed one to interpret atomic spectra.

The new quantum theory appeared to be more revolutionary and more powerful. It was triggered by Heisenberg's discovery of *matrix mechanics* and Schrödinger's discovery of *wave mechanics*. These two formalisms would not only predict the quantization of energy and provide a prescription to determine the energy difference between the levels (and thus explain the multiplicity of atomic spectra), but also allow one to calculate easily *probabilities of transitions* between different quantized energy states. At first, matrix mechanics and wave mechanics looked entirely different in their mathematical appearance and physical meaning. However, Schrödinger and Eckart [1] independently showed that the two theories are mathematically equivalent. Toward the end of 1926, Dirac unified the two theories using the concept of state vector and thus established the *transformation theory* of quantum mechanics. This ultimately had a profound implication for the quantum mechanical (mathematical) recipe to treat spin, as we will show in this chapter.

The transformation theory is the mathematical recipe to handle modern quantum mechanics. In Heisenberg's matrix mechanics, a physical quantity is expressed by a matrix, whereas in Schrödinger's wave mechanics, a physical quantity is expressed by a linear operator. In the unified transformation theory, physical quantities are represented by abstract linear operators called Dirac's *q-numbers*, which are linear operators in an infinite-dimensional linear space. Depending upon which types of orthogonal coordinate systems are used in this linear space, either matrix mechanics or wave mechanics emerges.

In other words, by using coordinate transformation in this linear space, we can derive matrix mechanics from wave mechanics and vice versa. Therefore, this unified theory was named transformation theory. The state of a quantum mechanical object is represented by a so-called *state vector*, which is an abstract vector in this linear space (the "wavefunction" in wave mechanics is an example of this) and the linear space is called the *state space*.

Earlier D. Hilbert and J. von Neumann had introduced the notion of a linear space that could absorb the mathematics of matrices and vectors, as well as the mathematics of linear operators and functions. This so-called *Hilbert space* admitted a finite or denumerably (countably) infinite number of co-ordinate axes. Therefore, a state vector in the Hilbert space could have at most a denumerably infinite number of mutually orthogonal components. Dirac extended this concept to a non-denumerably infinite number of coordinate axes in his linear space via the introduction of his famous δ-function [4]. The state vector therefore could have a non-denumerably infinite number of mutually orthogonal components and could be expressed as

$$\psi(q), \quad q \in [q_1, q_2] \ , \tag{2.1}$$

where the variable q is a continuous variable in the domain $[q_1, q_2]$. On the other hand, if the coordinate axes were countable, then the state vector would be expressed as

$$\psi_n, \quad n = 1, 2, 3, ... \tag{2.2}$$

where the variable n is an integer.

According to Dirac's transformation theory, the state vector (i) evolves in time according to a unitary transformation, and (ii) satisfies a *first order* differential equation with respect to time. This second property is very important, as we shall see later. Depending on whether the physical quantity* represented by the state vector will yield discrete or continuous values upon measurement, the eigenvalues of the linear operator describing this physical quantity will have discrete or continuous values. Accordingly, the coordinate axes in the linear space will be discrete or continuous, and the state vector will be ψ_n or $\psi(q)$. The magnitude squared of the component of the state vector, i.e.,, $|\psi_n|^2$ or $|\psi(q)|^2$ gives the probability of the physical quantity taking on the n-th (or q-th) value when the quantity is measured. This is the physical interpretation (or significance) of the state vector. Therefore, each component of the state vector is called a *probability amplitude*. The familiar "wavefunction" in the Schrödinger formalism of wave mechanics is the probability amplitude where the physical quantities corresponding to the coordinate axes in linear space are the position coordinates and time, i.e.,,

$$\psi(\vec{r}) = \psi(x, y, z, t) \ . \tag{2.3}$$

*A physical quantity, by definition, is anything that can be measured, even if by a gedanken experiment only.

In wave mechanics, the Schrödinger equation for a single particle tells us how the wavefunction evolves in time and space:

$$i\hbar\frac{\partial\psi(\vec{r})}{\partial t} = H_0\psi(\vec{r}) \ . \tag{2.4}$$

If we neglect spin, then

$$\begin{aligned} H_0 &= \frac{|\vec{p}|^2}{2m} + V(\vec{r}) \\ \vec{p} &= p_x\hat{\mathbf{x}} + p_y\hat{\mathbf{y}} + p_z\hat{\mathbf{z}} = -i\hbar\frac{\partial}{\partial x}\hat{\mathbf{x}} - i\hbar\frac{\partial}{\partial y}\hat{\mathbf{y}} - i\hbar\frac{\partial}{\partial z}\hat{\mathbf{z}}, \\ \vec{r} &= \left[x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}\right], t \end{aligned} \tag{2.5}$$

where the quantities with "hats" are unit vectors along the coordinate axes.

Solution of Equation (2.4) yields the wavefunction $\psi(\vec{r})$. The quantity H_0 is the so-called Hamiltonian whose first term is the kinetic energy and second term is the potential energy. The only restriction is that the potential energy term should be a real quantity so that the Hamiltonian remains a Hermitian operator, which guarantees that its eigenvalue (which is its expected value and therefore the expected value of the energy) remains a real quantity.

The question now is how to include "spin" in Equation (2.4)?

This was investigated by Wolfgang Pauli. He derived an equation to replace Equation (2.4) which bears his name and is known as the *Pauli Equation*. But before we discuss this equation, we need to understand an important concept, namely, *Pauli spin matrices*, since they appear in the Pauli equation.

2.1 Pauli spin matrices

In quantum mechanics, any physical observable is associated with an operator (which would be a linear operator in the Schrödinger formalism, or a matrix in the Heisenberg formalism). The eigenvalues of the linear operator, or the eigenvalues of the matrix, are the expectation values of the physical quantity, i.e.,, the values we expect to find if we measure the physical quantity in an experiment[†]. Spin is a physical observable since the associated angular momentum can be measured, as was done unwittingly by Stern and Gerlach. Consequently, there must be a quantum mechanical operator associated with spin. Pauli derived the quantum mechanical operators for the spin components along three orthogonal axes – S_x, S_y and S_z. They are 2×2 complex

[†]Repeated measurements of a physical observable will produce a distribution of values whose average will be the expectation value.

matrices that came to be known as the Pauli spin matrices. Pauli's approach was based on the premise that: (1) the measurement of the spin angular momentum component along any coordinate axis for an electron should give the results $+\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$, and (2) the operators for spin components along three mutually orthogonal axes should obey commutation rules similar to those obeyed by the operators associated with components of the orbital angular momentum. This would put spin angular momentum and orbital angular momentum on the same footing.

The operators (matrices) for the orbital angular momentum are known to satisfy the commutation relations

$$\begin{aligned} L_y L_z - L_z L_y &= i\hbar L_x, \\ L_z L_x - L_x L_z &= i\hbar L_y, \\ L_x L_y - L_y L_x &= i\hbar L_z, \end{aligned} \tag{2.6}$$

which merely reflect the fact that the orbital angular momenta along any two mutually orthogonal axes cannot be simultaneously measured with absolute precision unless the orbital angular momentum along a third axis, perpendicular to both the other two axes, vanishes.

Pauli adopted similar commutation relations for the spin angular momentum operators S_x, S_y and S_z:

$$\begin{aligned} S_y S_z - S_z S_y &= i\hbar S_x, \\ S_z S_x - S_x S_z &= i\hbar S_y, \\ S_x S_y - S_y S_x &= i\hbar S_z. \end{aligned} \tag{2.7}$$

Now, in the Stern-Gerlach experiment, assuming that the z-axis is the axis joining the south to north pole of the magnet, the observation of two traces on the photographic plate was interpreted as being caused by a spin angular momentum $\vec{S}$ whose z-component has two values $\pm\frac{\hbar}{2}$. Therefore, the matrix operator S_z must be (i) a 2×2 matrix (because such a matrix has *two* eigenvalues), and (ii) these eigenvalues must be $\pm\frac{\hbar}{2}$.

A 2×2 matrix that has eigenvalues of $\pm\frac{\hbar}{2}$ is the matrix

$$M_{2\times 2} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.8}$$

This is not the only 2×2 matrix with eigenvalues $\pm\frac{\hbar}{2}$ – there could be many others – but this is the matrix that Pauli chose as a start for the operator S_z.

Next, he had to find appropriate matrices to serve as operators S_x and S_y. Pauli realized that since the choice of the z-axis as the axis joining the north and south poles of the magnet is completely arbitrary, the result of the Stern-Gerlach measurement should not be affected if he had chosen this axis to be the x- or y-axis, instead. This means that the expectation values of S_x and

S_y, i.e.,, their eigenvalues, should also be $\pm\frac{\hbar}{2}$. Moreover, all three matrices – S_x, S_y and S_z – must satisfy the commutation relations in Equation (2.7).

Pauli first defined three dimensionless matrices σ_x, σ_y and σ_z such that

$$\begin{aligned} S_x &= \frac{\hbar}{2}\sigma_x, \\ S_y &= \frac{\hbar}{2}\sigma_y, \\ S_z &= \frac{\hbar}{2}\sigma_z. \end{aligned} \tag{2.9}$$

Since S_x, S_y and S_z must have eigenvalues of $\pm\frac{\hbar}{2}$, it is obvious that the σ-matrices must have eigenvalues of ± 1. Furthermore, Equation (2.7) mandates that

$$\begin{aligned} \sigma_y\sigma_z - \sigma_z\sigma_y &= 2i\sigma_x, \\ \sigma_z\sigma_x - \sigma_x\sigma_z &= 2i\sigma_y, \\ \sigma_x\sigma_y - \sigma_y\sigma_x &= 2i\sigma_z. \end{aligned} \tag{2.10}$$

According to Equations (2.8) and (2.9),

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.11}$$

So now Pauli needed to pick two matrices σ_x and σ_y such that they have eigenvalues of ± 1 and obey Equation (2.10). Since these matrices will be operators for physical observables (spin components), they must be Hermitian as well. It is easy to verify that σ_z is Hermitian.

We can start our search for σ_x and σ_y with Hermitian matrices that have off-diagonal elements only, i.e.,

$$\sigma_x = \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}, \tag{2.12}$$

and

$$\sigma_y = \begin{pmatrix} 0 & b \\ b^* & 0 \end{pmatrix}. \tag{2.13}$$

Since the eigenvalues of these matrices are ± 1, we must have $|a|^2 = |b|^2 = 1$, which leads to the possible choices for a and $b = \pm 1$ or $\pm i$.

Next, we must satisfy Equation (2.10) and that mandates

$$Im(ab^*) = 1, \tag{2.14}$$

where *Im* stands for imaginary part.

Therefore, if we select $a = +1$, then we must choose $b =$ –i, and this yields

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{2.15}$$

and

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{2.16}$$

This is how Pauli came up with expressions for σ_x, σ_y and σ_z. These matrices are called *Pauli spin matrices* and serve as operators for the spin components according to Equation (2.9).

It is obvious that Pauli's choice was by no means unique. There are other legitimate choices (e.g., we could have chosen $a = -\mathrm{i}$ and $b = +1$), but Pauli's choice is now history and universally adopted.

From the expressions for the Pauli spin matrices, we notice that the square of each of the Pauli matrices is the 2×2 unit matrix $[I]$. Hence

$$|S|^2 = S_x^2 + S_y^2 + S_z^2 = \frac{3}{4}\hbar^2[I] = \overline{s}(\overline{s}+1)\hbar^2[I], \tag{2.17}$$

with $\overline{s} = 1/2$. This should be compared with the equivalent relation for the orbital angular momentum operator

$$|L|^2 = m(m+1)\hbar^2[I], \quad m = 1, 2, 3... \tag{2.18}$$

2.1.1 Eigenvectors of the Pauli matrices: Spinors

The eigenvalues of the Pauli spin matrices are ±1. We now evaluate the corresponding eigenvectors that we denote as $|\pm>$.

Matrix σ_z: The eigenvectors of σ_z must satisfy

$$\sigma_z|\pm>_z = \pm 1|\pm>_z . \tag{2.19}$$

These eigenvectors (with unit norm) will be

$$|+>_z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{2.20}$$

and

$$|->_z = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.21}$$

It is easy to verify that these two eigenvectors are orthonormal, as they must be since they are eigenvectors of a Hermitian matrix corresponding to distinct (non-degenerate) eigenvalues.

Matrix σ_x: The eigenvectors of σ_x must satisfy

$$\sigma_x|\pm>_x = \pm 1|\pm>_x . \tag{2.22}$$

Starting with Equation (2.15), these eigenvectors are found to be

$$|+>_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{2.23}$$

and

$$|- >_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{2.24}$$

Once again, the two eigenvectors are orthonormal. As can be easily checked, these eigenvectors can also be expressed as

$$|\pm >_x = \frac{1}{\sqrt{2}}[|+ >_z \pm |- >_z]. \tag{2.25}$$

Matrix σ_y: The eigenvectors of σ_y must satisfy

$$\sigma_y |\pm >_y = \pm 1 |\pm >_y . \tag{2.26}$$

Using Equation (2.16), these eigenvectors are found to be

$$|+ >_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \tag{2.27}$$

and

$$|- >_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \tag{2.28}$$

These eigenvectors are orthonormal and can be expressed as

$$|\pm >_y = \frac{1}{\sqrt{2}}[|+ >_z \pm i |- >_z]. \tag{2.29}$$

The eigenvectors of the Pauli spin matrices are examples of "spinors" which are 2×1 column vectors that represent the spin state of an electron. If we know the spinor associated with an electron in a given state, we can deduce the electron's spin orientation, i.e., find the quantities $< S_x >$, $< S_y >$ and $< S_z >$, where the angular brackets $< ... >$ denote expectation values. We will see this later.

2.2 The Pauli equation and spinors

We can absorb the space and time dependent part of an electron's wavefunction in the spinor, so that the general form of a spinor will be

$$[\psi(\mathbf{x})] = \begin{bmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \end{bmatrix} , \tag{2.30}$$

where $\mathbf{x} \equiv (x, y, z, t)$, and ϕ_1 and ϕ_2 are the two components of the spinor wavefunction (assumed to be properly normalized).

With a 2-component wavefunction, the Schrödinger equation must be recast as

$$\left\{[H] + \frac{\hbar}{i}\frac{\partial}{\partial t}[I]\right\}[\psi(\mathbf{x})] = [0]\ , \tag{2.31}$$

where the Hamiltonian is a 2×2 matrix (since it may contain the 2×2 Pauli spin matrices), $[I]$ is the 2×2 identity matrix, and $[0]$ is the 2×1 null vector. Equation (2.31) is a set of two simultaneous differential equations for the two components of the spinor wavefunction – ϕ_1 and ϕ_2. Equation (2.31) is referred to as the *Pauli equation* [2].

Solution of the Pauli equation yields the two-component spinor wavefunction $[\psi(\mathbf{x})]$. Its practical use is in calculating the expected value of the spin angular momentum of an electron along any coordinate axis. The expected value along the n-th coordinate axis at location ($\vec{r}_0 = x_0, y_0, z_0$) at an instant of time t will be $[\psi(\vec{r}_0, t)]^\dagger[S_n][\psi(\vec{r}_0, t)]$, where $S_n = (\hbar/2)\sigma_n$ and the superscript † (dagger) represents the Hermitian conjugate. Using Equation (2.9), we get

$$\begin{aligned}
S_x(\vec{r}_0, t) &= (\hbar/2)[\psi(\vec{r}_0, t)]^\dagger\,[\sigma_x]\,[\psi(x_0, y_0, z_0, t)] \\
&= (\hbar/2)\,[\phi_1^*(\vec{r}_0, t)\ \phi_2^*(\vec{r}_0, t)] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \phi_1(\vec{r}_0, t) \\ \phi_2(\vec{r}_0, t) \end{bmatrix} \\
&= \hbar Re\,[\phi_1^*(\vec{r}_0, t)\phi_2(\vec{r}_0, t)]\,, \\
S_y(\vec{r}_0, t) &= (\hbar/2)\,[\psi(\vec{r}_0, t)]^\dagger\,[\sigma_y]\,[\psi(\vec{r}_0, t)] \\
&= (\hbar/2)\,[\phi_1^*(\vec{r}_0, t)\ \phi_2^*(\vec{r}_0, t)] \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} \phi_1(\vec{r}_0, t) \\ \phi_2(\vec{r}_0, t) \end{bmatrix} \\
&= \hbar Im\,[\phi_1^*(\vec{r}_0, t)\phi_2(\vec{r}_0, t)]\,, \\
S_z(\vec{r}_0, t) &= (\hbar/2)\,[\psi(\vec{r}_0, t)]^\dagger\,[\sigma_z]\,[\psi(\vec{r}_0, t)] \\
&= (\hbar/2)\,[\phi_1^*(\vec{r}_0, t)\ \phi_2^*(\vec{r}_0, t)] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \phi_1(\vec{r}_0, t) \\ \phi_2(\vec{r}_0, t) \end{bmatrix} \\
&= (\hbar/2)\{|\phi_1(\vec{r}_0, t)|^2 - |\phi_2(\vec{r}_0, t)|^2\}\ ,
\end{aligned} \tag{2.32}$$

where Re stands for the real part, Im stands for the imaginary part and the superscript * (asterisk) represents complex conjugate.

Therefore, if we can find the 2-component wavefunction in Equation (2.30) by solving the Pauli equation (2.31), then we can find the three components of the expected value of the spin angular momentum at any location at any instant of time. This is why the Pauli equation and the spinor concept are useful and important. In Chapter 7, we will also show how the Pauli equation can be used to derive the energy dispersion relations (relation between the energy and the wavevector) of an electron in a solid in the presence of spin-dependent effects.

2.3 More on the Pauli equation

Referring to Equation (2.31), we ask what terms will the 2×2 Hamiltonian $[H]$ contain. Normally, it will consist of three types of terms:

$$[H] = H_0[I] + [H_B] + [H_{SO}], \tag{2.33}$$

where H_0 is the spin-independent Hamiltonian, and $[H_B]$, $[H_{SO}]$ are 2×2 matrices that depend on spin and will therefore involve the Pauli spin matrices.

To understand where $[H_B]$ comes from, consider the fact that if we view "spin" as being associated with self-rotation of an electron about its axis, then the self-rotation of the charged entity will give rise to a magnetic moment $\vec{\mu}_e$. This magnetic moment will interact with any externally applied magnetic field, if such a field is present. Let us say that the flux density associated with the external field is $\vec{B}$. Then the energy of interaction of $\vec{\mu}_e$ with $\vec{B}$ is

$$E_{int} = -\vec{\mu}_e \cdot \vec{B}. \tag{2.34}$$

Landé had shown that the ratio of the magnetic moment $\vec{\mu}_e$ (in units of the Bohr magneton μ_B) to the angular momentum of self-rotation $\vec{S}$ (in units of $\hbar$) is the so-called gyromagnetic factor g [3]. Therefore, the operator associated with E_{int} is

$$[H_B] = -(g/2)\mu_B \vec{B} \cdot \vec{\sigma}, \tag{2.35}$$

since $\vec{S} = (\hbar/2)\vec{\sigma}$, where $\vec{\sigma} = \sigma_x \hat{x} + \sigma_y \hat{y} + \sigma_z \hat{z}$.

Obviously, the two eigenvalues of the matrix $[H_B]$ will not be the same, meaning that the eigenenergies associated with this Hamiltonian will *not* be degenerate. Therefore, this term will cause spin-splitting, or lift the degeneracy between the two spin states. This splitting is the *Zeeman splitting*. The Hamiltonian $[H_B]$ is called the *Zeeman Hamiltonian* or the *Zeeman interaction term*.

The Hamiltonian $[H_{SO}]$ is associated with spin-orbit interaction which also lifts the spin degeneracy. This interaction is discussed in Chapter 6.

Finally, the general form of the Pauli equation is

$$\left\{ H_0[I] + [H_B] + [H_{SO}] + \frac{\hbar}{i}\frac{\partial}{\partial t}[I] \right\} [\psi(\mathbf{x})] = [0]. \tag{2.36}$$

Its solution yields the 2-component wavefunction $[\psi(\mathbf{x})]$, which then yields the spin components of an electron at any position and at any instant of time from Equation (2.32). The Pauli Equation of course has many other uses, as we will see later in Chapter 7.

2.4 Extending the Pauli equation - the Dirac equation

The Pauli equation is completely non-relativistic and Pauli never found the avenue to reconcile it with relativity. That task was completed by Paul Andrew Maurice Dirac.

Both Schrödinger and two physicists, O. Klein and W. Gordon, had independently derived a relativistic equivalent of the Schrödinger equation. This is known as the *Klein-Gordon equation* [5]. A free particle not subjected to any force has a constant potential energy which can be taken to be zero (since potential is always undefined to the extent of an arbitrary constant). According to Einstein's special theory of relativity, such a particle obeys the relation

$$\overline{E}^2 = p^2c^2 + m_0^2c^4 \ , \tag{2.37}$$

where $\overline{E}$ is the total energy and p is the momentum.

According to De Broglie[‡],

$$\begin{aligned} \overline{E} &= h\nu, \\ p &= h/\lambda \ , \end{aligned} \tag{2.38}$$

where ν is the frequency and λ is the wavelength of the De Broglie wave associated with the particle ($\lambda = c\nu$). Therefore, the last equation can be re-written as

$$\nu^2 - \left(\frac{c}{\lambda}\right)^2 = \left(\frac{m_0c^2}{h}\right)^2 \ . \tag{2.39}$$

The above equation is known as the *Einstein-De Broglie equation.*

In the operator version of quantum mechanics, the energy operator is $i\hbar(\partial/\partial t)$, and the momentum operator (describing momentum along the x_r axis) is $-i\hbar(\partial/\partial x_r)$. Therefore, the quantum mechanical representation of Equation (2.37) is

$$\left[\left(i\hbar\frac{\partial}{c\partial t}\right)^2 - \sum_{r=1}^{3}\left(-i\hbar\frac{\partial}{\partial x_r}\right)^2 - m_0^2c^2\right]\psi(x,y,z,t) = 0 \ . \tag{2.40}$$

The above equation has the solution of a plane wave

$$\begin{aligned} \psi(x,y,z,t) &= e^{i(\vec{k}\cdot\vec{r}-\omega t)}, \\ \omega &= 2\pi\nu, \\ \vec{k} &= (2\pi)/\vec{\lambda}, \\ \vec{r} &= x\hat{x} + y\hat{y} + z\hat{z} \ . \end{aligned} \tag{2.41}$$

[‡]The De Broglie relation is reviewed in Chapter 18.

Substituting this solution into Equation (2.40) immediately yields Equation (2.39).

Equation (2.40) is valid for a free particle. Klein and Gordon extended it to a particle subjected to a (time-dependent) force field. Let the time dependent vector potential associated with the force field be $\vec{A} = (A_0, A_x, A_y, A_z)$, where we have treated space and time on the same footing as mandated by the theory of relativity. Klein and Gordon modified Equation (2.40) as

$$\left[\left(i\hbar\frac{\partial}{c\partial t} + eA_0\right)^2 - \sum_{r=1}^{3}\left(-i\hbar\frac{\partial}{\partial x_r} + eA_r\right)^2 - m_0^2c^2\right]\psi(x,y,z,t) = 0. \tag{2.42}$$

The above equation is the *Klein-Gordon Equation*. For a time, it was thought to be the fundamental equation of relativistic quantum mechanics.

Dirac, however, questioned the Klein-Gordon construct. This equation is a second order differential equation with respect to time, and, according to Dirac's transformation theory, all meaningful equations of quantum mechanics must be first order with respect to time (the Schrödinger and Pauli equations are).

Dirac insisted on an equation that will be first order with respect to time. Now, in relativity, space and time are treated as equivalent and therefore the desired equation must also be first order with respect to space. Accordingly, the sought after equation needs to have a form

$$\left[\left(i\hbar\frac{\partial}{c\partial t} + eA_0\right) - \sum_{r=1}^{3}\alpha_r\left(-i\hbar\frac{\partial}{\partial x_r} + eA_r\right) - \alpha_0 m_0 c\right]\psi(x,y,z,t) = 0. \tag{2.43}$$

When Dirac postulated the above equation, he simultaneously came up with four new quantities α_0 and α_r (r = 1,2,3), and he also told us how to determine them. He realized that a free particle without the vector potentials A_0 and A_r has to satisfy the Einstein-De Broglie Equation (Equation(2.39)) and therefore must satisfy Equation (2.40). In other words, the wavefunction $\psi(x,y,z,t)$ must be a solution of Equation (2.40). The latter equation is second order in space and time, whereas Dirac's Equation (Equation(2.43)) for a free particle,

$$\left[\left(i\hbar\frac{\partial}{c\partial t}\right) - \sum_{r=1}^{3}\alpha_r\left(-i\hbar\frac{\partial}{\partial x_r}\right) - \alpha_0 m_0 c\right]\psi(x,y,z,t) = 0, \tag{2.44}$$

is first order.

In order to take the above equation for a free particle and make it second order to match Equation (2.40), we apply it to the operator

$$\left[\left(i\hbar\frac{\partial}{c\partial t}\right) + \sum_{r=1}^{3}\alpha_r\left(-i\hbar\frac{\partial}{\partial x_r}\right) + \alpha_0 m_0 c\right], \tag{2.45}$$

to yield

$$\left[\left(i\hbar\frac{\partial}{c\partial t}\right)+\sum_{r=1}^{3}\alpha_r\left(-i\hbar\frac{\partial}{\partial x_r}\right)+\alpha_0 m_0 c\right]$$
$$\cdot\left[\left(i\hbar\frac{\partial}{c\partial t}\right)-\sum_{r=1}^{3}\alpha_r\left(-i\hbar\frac{\partial}{\partial x_r}\right)-\alpha_0 m_0 c\right]\psi(x,y,z,t)=0. \quad (2.46)$$

Dirac insisted that the above equation be Equation (2.40). This can only happen if the quantities α_0 and α_r are not ordinary numbers, but *matrices*. In that case, Equation (2.46) will become

$$[\left(i\hbar\frac{\partial}{c\partial t}\right)^2-\sum_{r=1}^{3}\{\alpha_r\}^2\left(-i\hbar\frac{\partial}{\partial x_r}\right)^2-$$
$$\sum_{m<n}(\{\alpha_m\}\{\alpha_n\}-\{\alpha_n\}\{\alpha_m\})(i\hbar)^2\frac{\partial^2}{\partial x_m x_n}-\{\alpha_0\}^2 m_0^2 c^2]\psi(x,y,z,t)=0, \quad (2.47)$$

where the matrices $\{\alpha_0\}$ and $\{\alpha_r\}$ $(r = 1, 2, 3)$ have the properties

$$\{\alpha_m\}^2=[I] \qquad (m=0,1,2,3),$$
$$\{\alpha_m\}\{\alpha_n\}+\{\alpha_n\}\{\alpha_m\}=[0]\ (m\neq n; m,n=0,1,2,3). \quad (2.48)$$

The simplest matrices that possess the properties listed in Equation (2.48) are

$$\{\alpha_0\}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix},$$
$$\{\alpha_1\}=\begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix},$$
$$\{\alpha_2\}=\begin{bmatrix}0&0&0&-i\\0&0&i&0\\0&-i&0&0\\i&0&0&0\end{bmatrix},$$
$$\{\alpha_3\}=\begin{bmatrix}0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0\end{bmatrix}. \quad (2.49)$$

When these 4×4 matrices are introduced into Equation (2.47), the wavefunction $\psi(x, y, z, t)$ becomes a 4×1 column vector

$$\psi(x, y, z, t) = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} , \tag{2.50}$$

so that Equation (2.47) becomes a set of four coupled differential equations with respect to the four variables ψ_1, ψ_2, ψ_3 and ψ_4.

A little bit of inspection will quickly reveal that the Dirac matrices in Equation (2.49) can be written in terms of the Pauli spin matrices as

$$\begin{aligned} \{\alpha\}_0 &= \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix}, \\ \{\alpha\}_1 &= \begin{bmatrix} \mathbf{0} & \sigma_1 \\ \sigma_1 & \mathbf{0} \end{bmatrix}, \\ \{\alpha\}_2 &= \begin{bmatrix} \mathbf{0} & \sigma_2 \\ \sigma_2 & \mathbf{0} \end{bmatrix}, \\ \{\alpha\}_3 &= \begin{bmatrix} \mathbf{0} & \sigma_3 \\ \sigma_3 & \mathbf{0} \end{bmatrix} , \end{aligned} \tag{2.51}$$

where $\mathbf{I}$ is a 2×2 identity matrix and $\mathbf{0}$ is a 2×2 null matrix.

Dirac applied his equation to the hydrogen atom and showed that the correct atomic level spacings can be obtained. More importantly, he showed that the orbital angular momentum *alone* is not a conserved quantity, but when the quantity represented by the operator

$$\frac{1}{2} \begin{bmatrix} \mathbf{0} & \sigma \\ \sigma & \mathbf{0} \end{bmatrix} \tag{2.52}$$

is added to it, the total quantity is conserved. Viewed from the perspective of conservation of total angular momentum, this shows that an electron has spin angular momentum given by the operator in Equation (2.52). *This is the first convincing theoretical demonstration of the existence of spin* and hence Dirac is credited with establishing the concept of spin rigorously. In the process, he also demonstrated that spin angular momentum must be quantized to two distinct values since the matrix in Equation (2.52) has two distinct and discrete eigenvalues. Therefore, Dirac was able to explain spin quantization, which the self-rotation model of the electron could never explain by itself.

Dirac also found that when an external force field is present, the procedure used to obtain Equation (2.47) from Equation (2.44) does not yield the Klein-Gordon equation (Equation (2.42)). The discrepancy can be explained by taking into account the interaction of the spin magnetic moment with the external field (recall the Zeeman interaction). Once this was demonstrated, the concept of spin was established on a firm footing.

2.4.1 Connection to Einstein's relativistic equation

If we substitute the results of Equation (2.4) in Equation (2.43), the Dirac equation reduces to

$$i\hbar\frac{\partial}{\partial t}[\psi(x,y,z,t)] = \begin{bmatrix} \left(m_0c^2+V\right)\mathbf{I} & c\vec{\sigma}\cdot\left(\vec{p}+e\vec{A}\right) \\ c\vec{\sigma}\cdot\left(\vec{p}+e\vec{A}\right) & \left(-m_0c^2+V\right)\mathbf{I} \end{bmatrix}[\psi(x,y,z,t)]. \quad (2.53)$$

where $\mathbf{I}$ is, once again, the 2×2 identity matrix.

Interpreting the operator $i\hbar\partial/\partial t$ as the energy operator $\overline{E}_{op}$, the above equation can be written as

$$\overline{E}_{op} = \begin{bmatrix} \left(m_0c^2+V\right)\mathbf{I} & c\vec{\sigma}\cdot\left(\vec{p}+e\vec{A}\right) \\ c\vec{\sigma}\cdot\left(\vec{p}+e\vec{A}\right) & \left(-m_0c^2+V\right)\mathbf{I} \end{bmatrix}. \quad (2.54)$$

When both vector and scalar potentials are absent, i.e., $V = A = 0$, we can square both sides of the last equation to obtain

$$\begin{aligned} \overline{E}_{op}^2 &= \begin{bmatrix} m_0c^2\mathbf{I} & c\vec{\sigma}\cdot\vec{p} \\ c\vec{\sigma}\cdot\vec{\sigma} & -m_0c^2\mathbf{I} \end{bmatrix}^2 \\ &= \left(p^2c^2+m_0^2c^4\right)\begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{bmatrix}, \end{aligned} \quad (2.55)$$

which matches exactly Equation (2.37). Therefore, indeed the Dirac equation is the relativistic version of the Schrödinger equation.

2.5 Time-independent Dirac equation

From Equation (2.53), it is straightforward to show that the *time independent* Dirac equation will be

$$\begin{bmatrix} A & 0 & C & D^* \\ 0 & A & D & -C \\ C & D^* & B & 0 \\ D & -C & 0 & B \end{bmatrix}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} = \overline{E}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix}, \quad (2.56)$$

where the asterisk denotes complex conjugate and

$$\begin{aligned} A &= m_0c^2 + V \\ B &= -m_0c^2 + V \\ C &= c(p_z + eA_z) \\ D &= c[(p_x + eA_x) + i(p_y + eA_y)]. \end{aligned} \tag{2.57}$$

Equation (2.56) can be written more compactly as

$$\begin{bmatrix} (m_0c^2 + V)[I] & c\vec{\sigma} \cdot \left[\vec{p} + e\vec{A}\right] \\ c\vec{\sigma} \cdot \left[\vec{p} + e\vec{A}\right] & (-m_0c^2 + V)[I] \end{bmatrix} \begin{bmatrix} \{\psi\}(x,y,z) \\ \{\phi\}(x,y,z) \end{bmatrix} = \overline{E} \begin{bmatrix} \{\psi\}(x,y,z) \\ \{\phi\}(x,y,z) \end{bmatrix}, \tag{2.58}$$

where V is the scalar potential energy and

$$\{\psi\}(x,y,z) = \begin{pmatrix} \psi_1(x,y,z) \\ \psi_2(x,y,z) \end{pmatrix}, \tag{2.59}$$

and

$$\{\phi\}(x,y,z) = \begin{pmatrix} \psi_3(x,y,z) \\ \psi_4(x,y,z) \end{pmatrix}. \tag{2.60}$$

From Equation (2.58), we can show that

$$\begin{aligned} &\left\{ \left(m_0c^2 + V\right)[I] + \left[c\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right] \frac{1}{\overline{E} + m_0c^2 - V} [I] \left[c\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right] \right\} [\psi] = \overline{E}[\psi], \\ &\left\{ \left(-m_0c^2 + V\right)[I] + \left[c\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right] \frac{1}{\overline{E} - m_0c^2 - V} [I] \left[c\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right] \right\} [\phi] = \overline{E}[\phi]. \end{aligned} \tag{2.61}$$

2.5.1 Non-relativistic approximation to the Dirac equation

Consider a non-relativistic electron moving at speeds much less than the speed of light in vacuum. For such a particle, Equation (2.37) yields $\overline{E} \approx m_0c^2$. Using that result in the first of the two equations above yields

$$\left\{ \left(m_0c^2 + V\right)[I] + \frac{\left[\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right]^2}{2m_0} \right\} [\psi] = \overline{E}[\psi], \tag{2.62}$$

which reduces to

$$\begin{aligned} \left(\overline{E} - m_0c^2\right)[\psi] = E[\psi] &= \left(\frac{(\vec{p} + e\vec{A})^2}{2m_0} [I] + \mu_B \vec{B} \cdot \vec{\sigma} + V[I] \right) [\psi], \\ &= \{[H_0] + [H_B]\} [\psi], \end{aligned} \tag{2.63}$$

since $[\vec{\sigma}\cdot(\vec{p}+e\vec{A})]^2 = (\vec{p}+e\vec{A})^2[I] + 2m_0\mu_B\vec{B}\cdot\vec{\sigma}$ (see Problem 2.5). Here E is the total energy *minus* the rest energy m_0c^2.

The second of the two equations in Equation (2.61) will yield the same result as Equation (2.63) provided we make the transformation $m_0 \rightarrow -m_0$. This shows that the second equation applies to particles with negative mass, namely, *anti-matter*. The existence of anti-matter is a foregone conclusion from Equation (2.37). Noting that the De Broglie relation relates momentum p to wavevector k of a particle as $p = \hbar k$, Equation (2.37) gives *two* dispersion relations E versus k. They are shown in Fig. 2.1. One branch has a positive curvature and therefore positive mass. This corresponds to "matter." The other has a negative curvature and therefore negative mass. That corresponds to "anti-matter." The energy separation between the two curves is $2m_0c^2$, which is ~ 1 MeV for a free electron. These energy scales are seldom encountered in solid state physics, which is why anti-matter is usually of concern only in high energy physics. We will not have any occasion to worry about anti-matter anywhere in this textbook.

2.5.2 Relationship between the non-relativistic approximation to the Dirac equation and the Pauli equation

The reader will immediately recognize Equation (2.63) as the Pauli equation (Equation 2.33) without the spin-orbit interaction term. The latter term does not arise here since, strictly speaking, spin-orbit interaction is a relativistic effect and therefore cannot be captured within a non-relativistic picture. However, what is amazing is that the Zeeman interaction term appears automatically, without having to introduce it separately. Therefore, Dirac was able to explain the Zeeman interaction directly from his equation!

The spin-orbit interaction term is not beyond the Dirac equation, but it is beyond the non-relativistic approximation since spin-orbit interaction has a relativistic origin. If we make a binomial expansion of Equation (2.61) and retain only the lowest order terms, then we get an equation [6]

$$\left[\frac{|\vec{p}+e\vec{A}|^2}{2m_0} + V - \frac{|\vec{p}+e\vec{A}|^4}{8m_0^3c^2} + \frac{\hbar}{4im_0^2c^2}\left(\vec{\nabla}V\cdot(\vec{p}+e\vec{A})\right) + \frac{\hbar}{4m_0^2c^2}\left(\vec{\nabla}V\times(\vec{p}+e\vec{A})\right)\cdot\vec{\sigma}\right][\psi] = E[\psi], \quad (2.64)$$

where the last term in the left hand side represents the spin-orbit interaction term. Therefore, the Dirac equation incorporates the spin-orbit interaction physics as well.

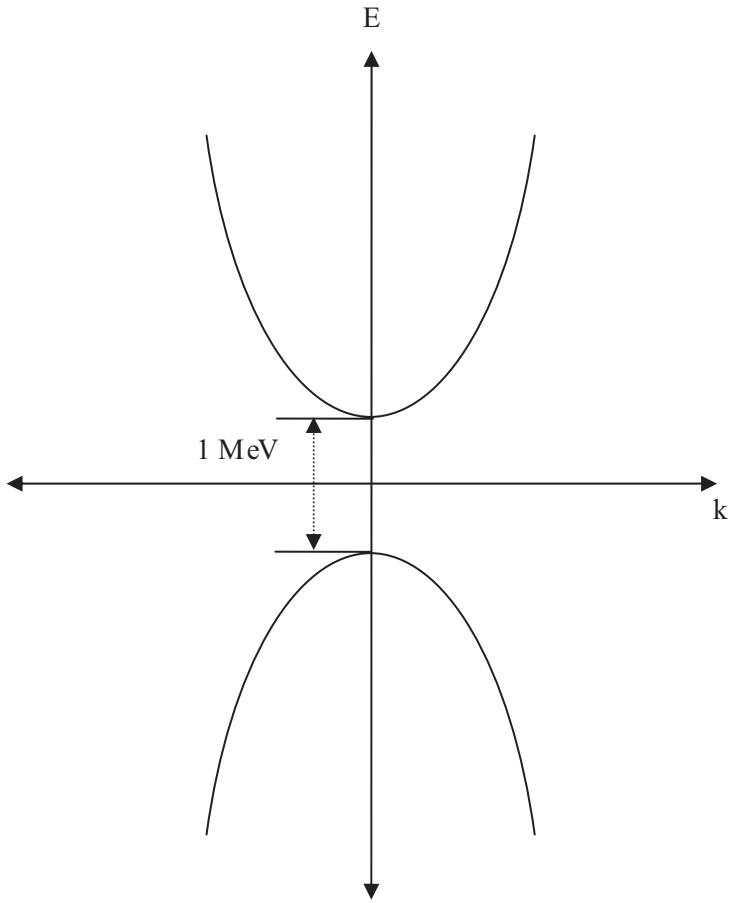


FIGURE 2.1
Dispersion relations for matter and anti-matter branches.

2.6 Problems

- **Problem 2.1**

  Show that the operators for S_x, S_y and S_z satisfy Equations (2.7) and (2.17).

  Solution

$$S_xS_y - S_yS_x = (\hbar^2/4)\left\{\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}0&-i\\i&0\end{bmatrix} - \begin{bmatrix}0&-i\\i&0\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}\right\},$$
$$= \left(\frac{\hbar}{2}\right)^2\begin{bmatrix}2i&0\\0&-2i\end{bmatrix} = i\hbar\left[\frac{\hbar}{2}\sigma_z\right] = i\hbar S_z. \qquad (2.65)$$

  Furthermore,

$$|S|^2 = S_x^2 + S_y^2 + S_z^2,$$
$$= (\hbar^2/4)\left\{\begin{bmatrix}0&1\\1&0\end{bmatrix}^2 + \begin{bmatrix}0&-i\\i&0\end{bmatrix}^2 + \begin{bmatrix}1&0\\0&-1\end{bmatrix}^2\right\},$$
$$= (3/4)\hbar^2\begin{bmatrix}1&0\\0&1\end{bmatrix} = (3/4)\hbar^2[I]. \qquad (2.66)$$

- **Problem 2.2**

  Using Equation (2.32), show that if the 2-component spinor is the state $|+>_z$ given in Equation (2.20), then $< S_x > = < S_y > = 0$ and $< S_z > = \hbar/2$. Also show that if the 2-component spinor is the state $|->_z$ given in Equation (2.21), then $< S_x > = < S_y > = 0$ and $< S_z > =$ -$\hbar/2$. Hence the state $|+>_z$ is referred to as the *+z-polarized state* and the state $|->_z$ is referred to as the *–z-polarized state*. This means that an electron in these states has its spin polarized along the +z and –z axes, respectively.

Solution

For the state $|+>_z$, $\phi_1 = 1$ and $\phi_2 = 0$ in Equation (2.30).

From Equation (2.32)

$$\begin{aligned} < S_x >= \hbar Re(\phi_1^* \phi_2) = 0, \\ < S_y >= \hbar Im(\phi_1^* \phi_2) = 0, \\ < S_z >= (\hbar/2)[|\phi_1|^2 - |\phi_2|^2] = \hbar/2. \end{aligned} \tag{2.67}$$

For the state $|->_z$, $\phi_1 = 0$ and $\phi_2 = 1$, hence

$$\begin{aligned} < S_x >= \hbar Re(\phi_1^* \phi_2) = 0, \\ < S_y >= \hbar Im(\phi_1^* \phi_2) = 0, \\ < S_z >= (\hbar/2)[|\phi_1|^2 - |\phi_2|^2] = -\hbar/2. \end{aligned} \tag{2.68}$$

Similarly, for the states $|\pm>_x$, $< S_x >= \pm\frac{\hbar}{2}$ and $< S_y >=< S_z >= 0$, so that these states are $\pm$x-polarized states. One can show that the states $|\pm>_y$ are $\pm$y-polarized states.

- **Problem 2.3**

Show that if any 2×2 Hermitian matrix $[H]$ is used to perform a unitary transformation of the Pauli matrices defined above, i.e., if matrices $\sigma_x{}'$, $\sigma_y{}'$, $\sigma_z{}'$ are defined such that

$$\sigma_n{}' = S\sigma_n S^{-1}, \tag{2.69}$$

where $S = e^{iH}$, for $n = x$, y, or z, the commutation rules in Equation (2.7) are satisfied by the matrices $\sigma_n{}'$.

Furthermore, the square of each of the matrices $\sigma_n{}'$ is equal to the 2×2 identity matrix.

Solution

The solution of this problem is left to the reader. Use the fact that the matrix S is

$$[S] = [M]^{-1} \begin{bmatrix} e^{i\lambda_1} & 0 \\ 0 & e^{i\lambda_2} \end{bmatrix} [M], \tag{2.70}$$

where $\lambda_{1,2}$ are the eigenvalues of the matrix $[H]$, and $[M]$ is a 2×2 matrix whose two columns are the eigenfunctions of $[H]$ corresponding to eigenvalues λ_1 and λ_2.

- **Problem 2.4**

Show the following properties of the Pauli spin matrices by actual computation.

1. $\det(\sigma_j) = -1$; for j = x, y, or z.
2. $\mathrm{Tr}(\sigma_j) = 0$.
3. $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$.
4. $\sigma_x\sigma_y\sigma_z = iI$.
5. $\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z$.
6. $\sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x$.
7. $\sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y$.
8. $\sigma_p\sigma_q + \sigma_q\sigma_p = 0, \quad (p \neq q; p, q = x, y, z)$.

- **Problem 2.5**

  Show that $\left[\vec{\sigma}\cdot(\vec{p}+e\vec{A})\right]^2 = (\vec{p}+e\vec{A})^2[I] + 2m_0\mu_B\vec{B}\cdot\vec{\sigma}$.

  Solution

$$\begin{aligned}\left[\vec{\sigma}\cdot(\vec{p}+e\vec{A})\right]^2 &= [\sigma_x(p_x+eA_x)+\sigma_y(p_y+eA_y)+\sigma_z(p_z+eA_z)]\\ &\quad\times[\sigma_x(p_x+eA_x)+\sigma_y(p_y+eA_y)+\sigma_z(p_z+eA_z)],\\ &= \sigma_x^2(p_x+eA_x)^2+\sigma_y^2(p_y+eA_y)^2+\sigma_z^2(p_z+eA_z)^2\\ &\quad+\sigma_x\sigma_y(p_x+eA_x)(p_y+eA_y)+\sigma_y\sigma_x(p_y+eA_y)(p_x+eA_x)\\ &\quad\sigma_y\sigma_z(p_y+eA_y)(p_z+eA_z)+\sigma_z\sigma_y(p_z+eA_z)(p_y+eA_y)\\ &\quad\sigma_z\sigma_x(p_z+eA_z)(p_x+eA_x)+\sigma_x\sigma_z(p_x+eA_x)(p_z+eA_z).\end{aligned} \tag{2.71}$$

Now the results of Problem 2.4 lead to

$$\begin{aligned}\left[\vec{\sigma}\cdot(\vec{p}+e\vec{A})\right]^2 &= (p_x+eA_x)^2+(p_y+eA_y)^2+(p_z+eA_z)^2\\ &\quad+i\sigma_z\left(-ie\hbar\frac{\partial A_y}{\partial x}+ie\hbar\frac{\partial A_x}{\partial y}\right)\\ &\quad+i\sigma_y\left(-ie\hbar\frac{\partial A_x}{\partial z}+ie\hbar\frac{\partial A_z}{\partial x}\right)\\ &\quad+i\sigma_x\left(-ie\hbar\frac{\partial A_z}{\partial y}+ie\hbar\frac{\partial A_y}{\partial z}\right)\\ &= (p_x+eA_x)^2+(p_y+eA_y)^2+(p_z+eA_z)^2\\ &\quad+e\hbar\sigma_z\left(\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right)\\ &\quad+e\hbar\sigma_y\left(\frac{\partial A_x}{\partial z}-\frac{\partial A_z}{\partial x}\right)\\ &\quad+e\hbar\sigma_x\left(\frac{\partial A_z}{\partial y}-\frac{\partial A_y}{\partial z}\right)\\ &= \left[\vec{p}+e\vec{A}\right]^2+e\hbar\left(\vec{\nabla}\times\vec{A}\right)\cdot\vec{\sigma},\end{aligned} \tag{2.72}$$

where we have used the fact that $\vec{p}_n = -i\hbar(\partial/\partial\vec{x}_n)$.

Since $\mu_B = e\hbar/2m_0$ and $\vec{B} = \vec{\nabla} \times \vec{A}$, this proves the equality stated in the problem.

- **Problem 2.6**

  Show that the 4×4 Dirac matrices are related to the 2×2 Pauli matrices as follows:

$$\begin{aligned} -i\{\alpha\}_0\{\alpha\}_2\{\alpha\}_3 &= \begin{pmatrix} \sigma_x & 0 \\ 0 & -\sigma_x \end{pmatrix} \\ -i\{\alpha\}_0\{\alpha\}_3\{\alpha\}_1 &= \begin{pmatrix} \sigma_y & 0 \\ 0 & -\sigma_y \end{pmatrix} \\ -i\{\alpha\}_0\{\alpha\}_1\{\alpha\}_2 &= \begin{pmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{pmatrix} \end{aligned} \tag{2.73}$$

  Show also that

$$\begin{aligned} \{\alpha\}_0\{\alpha\}_1 &= \begin{pmatrix} 0 & \sigma_x \\ -\sigma_x & 0 \end{pmatrix} \\ \{\alpha\}_0\{\alpha\}_2 &= \begin{pmatrix} 0 & \sigma_y \\ -\sigma_y & 0 \end{pmatrix} \\ \{\alpha\}_0\{\alpha\}_3 &= \begin{pmatrix} 0 & \sigma_z \\ -\sigma_z & 0 \end{pmatrix} \end{aligned} \tag{2.74}$$

- **Problem 2.7**

  Show that when an electron's or positron's momentum $\hbar k \ll m_0 c$, the dispersion relations in Fig. 2.1 are approximately parabolic.

2.7 Appendix

2.7.1 Working with spin operators

Applying the operators σ_x, σ_y and σ_z to the 2-component wavefunction in Equation (2.30) yields

$$\begin{aligned} \sigma_x\psi(\mathbf{x}) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \end{bmatrix}, \\ \sigma_y\psi(\mathbf{x}) &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \end{bmatrix}, \\ \sigma_z\psi(\mathbf{x}) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \end{bmatrix}, \end{aligned} \tag{2.75}$$

and therefore

$$S_x\psi = (\hbar/2)\sigma_x\psi = \frac{\hbar}{2}\begin{bmatrix}\phi_2(\mathbf{x})\\ \phi_1(\mathbf{x})\end{bmatrix},$$
$$S_y\psi = (\hbar/2)\sigma_x\psi = \frac{\hbar}{2}\begin{bmatrix}-i\phi_2(\mathbf{x})\\ i\phi_1(\mathbf{x})\end{bmatrix},$$
$$S_z\psi = (\hbar/2)\sigma_x\psi = \frac{\hbar}{2}\begin{bmatrix}\phi_1(\mathbf{x})\\ -\phi_2(\mathbf{x})\end{bmatrix}. \tag{2.76}$$

Thus, operating with S_x interchanges the two components of the spinor, operating with S_y interchanges the two components of the spinor while causing a phase shift of –90° to the second component and a phase shift of 90° to the first component, and operating with S_z introduces a phase shift of 180° to the second component.

2.7.2 Two useful theorems

A trivial decomposition of any 2×2 matrix M is obviously

$$M = m_{11}\begin{pmatrix}1&0\\0&0\end{pmatrix} + m_{12}\begin{pmatrix}0&1\\0&0\end{pmatrix} + m_{21}\begin{pmatrix}0&0\\1&0\end{pmatrix} + m_{22}\begin{pmatrix}0&0\\0&1\end{pmatrix}, \tag{2.77}$$

since the four matrices on the right hand side form a complete basis for all 2×2 matrices. Now that we have introduced the Pauli spin matrices, a more subtle decomposition of any 2×2 complex matrix can be found, as discussed next.

Theorem I: Any 2×2 matrix M

$$M = \begin{pmatrix}m_{11}&m_{12}\\m_{21}&m_{22}\end{pmatrix} \tag{2.78}$$

can be decomposed as

$$M = \frac{m_{11}+m_{22}}{2}I + \frac{m_{11}-m_{22}}{2}\sigma_z + \frac{m_{12}+m_{21}}{2}\sigma_x + i\frac{m_{12}-m_{21}}{2}\sigma_y. \tag{2.79}$$

The proof is left as an exercise. In other words, the 4 matrices $(I, \sigma_x, \sigma_y, \sigma_z)$ form a complete set of bases in the space of 2×2 complex matrices.

The last equation can be written in the more condensed form

$$M = a_o I + \vec{a}\cdot\vec{\sigma}, \tag{2.80}$$

where

$$a_o = \frac{1}{2}Tr(M), \tag{2.81}$$

and

$$\vec{a} = \frac{1}{2}Tr(M\vec{\sigma}), \tag{2.82}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, and Tr stands for the trace of the matrix.

A comparison of Equations (2.79) and (2.80) shows that M is Hermitian if a_0 and the three components of the vector $\vec{a}$ are real.

Exercise: Calculate the values of $(a_0, \vec{a})$ for the matrix M given by

$$M = \begin{pmatrix} 2 & \frac{i\sqrt{3}}{3} \\ \frac{-i\sqrt{3}}{3} & 4 \end{pmatrix}. \tag{2.83}$$

Next, we prove an identity which will be used in the next chapter to interpret geometrically the Pauli matrices after the introduction of the Bloch sphere concept.

Theorem II: If θ is real and if the matrix A is such that $A^2 = I$, the following identity holds:

$$e^{i\theta A} = cos\theta I + isin\theta A. \tag{2.84}$$

This is the generalization to operators of the well-known Euler relation for complex numbers, i.e., $e^{iz} = cosz + isinz$.

From the Taylor series expansion

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \tag{2.85}$$

and the definition of the function of an operator, we get

$$e^{i\theta A} = I + (i\theta)A + \frac{(i\theta)^2 A^2}{2!} + \frac{(i\theta)^3 A^3}{3!} + \frac{(i\theta)^4 A^4}{4!} + ... \tag{2.86}$$

or

$$\begin{aligned} e^{i\theta A} &= (1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - ... + (-1)^k \frac{\theta^{2k}}{(2k)!})I \\ &+ i(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - ... + (-1)^k \frac{i\theta^{2k+1}}{(2k+1)!})A, \end{aligned} \tag{2.87}$$

which is indeed Equation (2.84) if we use the the Taylor expansions

$$sinx = \sum_{k=0}^{\infty} (-1)^k x^{2k+1}/(2k+1)!, \tag{2.88}$$

and

$$cosx = \sum_{k=0}^{\infty} (-1)^k x^{2k}/(2k)!. \tag{2.89}$$

2.7.3 Applications of the *Postulates of Quantum Mechanics* to a few spin problems

The Postulates of Quantum Mechanics are briefly reviewed in Chapter 18.

Example 1: If we measure the z-component of an electron's spin, apply the postulate of quantum projective measurement (Postulate 3) discussed in Chapter 18 to calculate the probability of the measurement to give the result $\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$, respectively, if prior to the measurement, the state of the system $|\psi>$ is either

(1) $|0>$,
(2) $|1>$, or
(3) $\frac{1}{\sqrt{2}}(|0>+|1>)$,

where $|0>$ is the +z-polarized state $|+>_z$ and $|1>$ is the -z-polarized state $|->_z$.

Solution

(1) The operator $S_z = \frac{\hbar}{2}\sigma_z$ has the spectral decomposition

$$S_z = \frac{\hbar}{2}|0><0| + (-\frac{\hbar}{2})|1><1| = \sum_m mP_m, \tag{2.90}$$

where $P_m = |m><m|$, with $|m>$ being the eigenvectors of S_z.

Hence, if $|\psi>=|0>$, the probability of measuring $+\frac{\hbar}{2}$ is equal to

$$p(+\frac{\hbar}{2}) =<0|(|0><0|)|0>= 1, \tag{2.91}$$

and the probability of measuring $-\frac{\hbar}{2}$ is equal to

$$p(-\frac{\hbar}{2}) =<0|(|1><1|)|0>= 0 \tag{2.92}$$

and the sum of the probabilities is indeed equal to unity.

(2) If $|\psi>=|1>$, we get

$$p(\frac{\hbar}{2}) =<1|(|0><0|)|1>= 0, \tag{2.93}$$

and

$$p(-\frac{\hbar}{2}) =<1|(|1><1|)|1>= 1. \tag{2.94}$$

(3) Finally, if $|\psi>=\frac{1}{\sqrt{2}}(|0>+|1>)$, we find

$$p(+\frac{\hbar}{2}) = \frac{1}{\sqrt{2}}\{|0>+<1|\}(|0><0|)\frac{1}{\sqrt{2}}\{|0>+<1|\} = \frac{1}{2}\{1+0\}\cdot\{0+1\} = \frac{1}{2}, \tag{2.95}$$

and similarly, $p(-\frac{\hbar}{2}) = \frac{1}{2}$.

According to *Postulate 3* discussed in Chapter 18, right after the measurement, the spinor collapses into the state

$$|\psi > \rightarrow |\psi^{new} >= \frac{(|0><0|)}{\sqrt{p(+\frac{\hbar}{2})}}|\psi >= (\frac{|0><0|}{\sqrt{\frac{1}{2}}})(\frac{(|0>+|1>)}{\sqrt{2}}) = |0> . \tag{2.96}$$

Exercise: Repeat the previous exercise if the component S_x is measured instead.

Example 2: Suppose an electron is prepared in the spinor $|0>$ eigenstate of S_z with eigenvalue $+\frac{\hbar}{2}$ and repeated measurements are made of the x-component of its intrinsic angular momentum; calculate the average value $< S_x >$ and the standard deviation $\Delta(S_x)$ of these measurements.

Solution

The spectral decomposition of the operator S_x is given by

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (+\frac{\hbar}{2})|+>_{x,x}<+| + (-\frac{\hbar}{2})|->_{x,x}<-|, \tag{2.97}$$

where $|+>_x$ is the $+x$-polarized state and $|->_x$ is the $-x$-polarized state.

Hence,

$$< S_x >= \frac{\hbar}{2}\left[<0|S_x|0>= (1\ 0)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right] = 0. \tag{2.98}$$

Furthermore,

$$\begin{aligned} \Delta(S_x) &= \sqrt{< {S_x}^2 > - < S_x >^2} = \sqrt{< {S_x}^2 >}, \\ &= \frac{\hbar}{2}\left[<0|\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}|0>\right]^{1/2}, \\ &= \frac{\hbar}{2}\left[(1\ 0)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]^{1/2} = \frac{\hbar}{2}. \end{aligned} \tag{2.99}$$

Example 3: Suppose an electron is characterized by the spinor

$$|\psi >= \frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ -1 \end{pmatrix}, \tag{2.100}$$

which is properly normalized, as easily checked. If we measure the y-component of the spin, what is the probability of finding that its value is $\frac{\hbar}{2}$?

Solution

The eigenvector of σ_y are $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}$ with eigenvalue +1 and $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix}$ with eigenvalue –1

Hence, if we write the spinor $|\psi>$ as a linear combination of $|+>_y$ and $|->_y$,

$$|\psi>=\frac{1}{\sqrt{10}}\begin{pmatrix}3\\ -1\end{pmatrix}=\alpha\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}+\beta\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix}. \tag{2.101}$$

The probability of finding $+\frac{\hbar}{2}$ when measuring S_y is given by $|\alpha|^2$ or

$$p(+\frac{\hbar}{2})=|_y<+|\psi>|^2=\left|\frac{1}{\sqrt{2}}(1,\ -i)\frac{1}{\sqrt{10}}\begin{pmatrix}3\\ -1\end{pmatrix}\right|^2=\frac{1}{20}\left|(3+i)\right|^2=\frac{1}{2}. \tag{2.102}$$

Example 4: Suppose an electron is characterized by the spinor

$$|\psi>=\frac{4}{5}|0>+\frac{3}{5}|1>. \tag{2.103}$$

(1) What is the probability that a measurement of z-component of the spin will be $+\frac{\hbar}{2}$? and $-\frac{\hbar}{2}$?

Solution

$$p(+\frac{\hbar}{2})=\left|<0|(\frac{4}{5}|0>+\frac{3}{5}|1>)\right|^2=\frac{16}{25}, \tag{2.104}$$

and

$$p(\frac{-\hbar}{2})=\left|<1|(\frac{4}{5}|0>+\frac{3}{5}|1>)\right|^2=\frac{9}{25}. \tag{2.105}$$

(2) What is the expectation value of $S_z=\frac{\hbar}{2}\sigma_z$?

Solution

$$\begin{aligned}<\psi|S_z|\psi>&=\left(\frac{4}{5}<0|+\frac{3}{5}<1|\right)\frac{\hbar}{2}\sigma_z\left(\frac{4}{5}|0>+\frac{3}{5}|1>\right),\\ &=\frac{16}{25}<0|\sigma_z|0>\frac{\hbar}{2}+\left(\frac{3}{5}\right)^2\frac{\hbar}{2}<1|\sigma_z|1>=\frac{16}{25}\left(\frac{\hbar}{2}\right)-\frac{9}{25}\left(\frac{\hbar}{2}\right),\\ &=\frac{7}{25}\left(\frac{\hbar}{2}\right).\end{aligned} \tag{2.106}$$

(3) What is the standard deviation of S_z if measurements are made on many electrons prepared in the state $|\psi>$ above?