A guide to quantum computation

Teo Banica

Department of Mathematics, University of Cergy-Pontoise, F-95000 Cergy-Pontoise, France. teo.banica@gmail.com

2010 *Mathematics Subject Classification.* 81R50

Key words and phrases. Quantum computation, Quantum information

Abstract. This is an introduction to quantum computation and quantum information, starting from the very basics, with all the needed quantum physics preliminaries included. We insist on both the physical and engineering aspects, in relation with the architecture of the quantum computers, and on the mathematics of the computing part.

Preface

What is the best type of computer? Believe me or not, that is your brain. It's powerful, lightweight and very efficient. Assuming for instance that you know some calculus, just feed it in the morning with some bacon, eggs and coffee, and then all day long that brain will produce a lot of interesting formulae, that heavier machinery like mechanical or electric computers, running on a lot of power, will have troubles in matching.

This type of computer, like your brain, is called chemical. These are the best computers, and despite their functioning being relatively known, and the demand for such things being enormous, no one knows how to implement them. The only known exploitation scheme so far is the good old one, pay a human money, so that human can afford some housing, and buy bacon, eggs and coffee for the morning, and then work for you.

At the other extreme of the spectrum, also good and solid machinery, but this time easy to implement, while not very powerful, are the mechanical computers. There are all sorts of possible devices here, known to mankind since centuries, and you have surely already met many of them. In case you are a happy owner of a garage, you can invent some more, like so many people do. The car can of course sleep outside in the rain, and in the garage, whenever you have a problem say in computing this or that length, for this or that piece of material that you work on, you can always invent some clever combination of pieces of wood, bits of metal, strings and so on, computing that length for you.

The story is not over here, because in between the mechanical and chemical computers, we have a sort of good compromise, the electric computers. These need no presentation, and with the remark that in recent years these have become more popular than both mechanical and chemical computers, heavily used for instance by people doing nothing all day long. Also, as an interesting historical remark, the electric computers were originally concieved as weapons, on one hand by John von Neumann and others in the US, for helping in computations for the atomic bomb, matter of destroying Germany, and on the other hand by Alan Turing and others in the UK, for intercepting the war communications of the same Germans. Good old stories, followed by many more electric computer stories in relation with weaponry, during the Cold War. And there is still a flavor of this nowadays, with the best supercomputers being typically reserved for defense matters.

So, this was the story of computers and computing, and it has been like this for ages, and more specifially since WW2. But, can we come someday with something new? We would need here a new type of computer, easy to implement and of widespread use, complementing the known mechanical, electric and chemical ones.

Quantum computing says that yes, it is possible to build quantum computers, appearing as a quantum version of the electric computers. And with "quantum" being something quite advanced, we can only expect such computers to be much harder to build, but also be far more powerful in computing, once built, than the electric computers.

Such ideas have been around for a while, especially among physicists. One of the main principles in modern physics is that electromagnetism refines into quantum mechanics, so the idea that electric computers will one day refine into quantum computers too is something very natural. In practice, the buzz however came only in the early 90s, and from computer scientists, notably due to two famous papers, of Shor [**79**] and Grover [**46**], alerting mankind that the one who builds a quantum computer, wins.

These two papers had a tremendous influence. Foremost among people doing politics, military and finance, not to forget the big electric computer makers, who started pouring money into the thing. But also among scientists of all kinds, mathematicians, physicists of all orientations and skills combined, and of course, computer scientists. And things have been constantly advancing since then, both at the practical level, that of constructing such computers, and at the theoretical level, on how to effectively use such a beast, once constructed. Overall, this is something doing very well, and one day, soon, we will all have a quantum computer in our pocket. Or at least that's what theory says.

The present book is an introduction to all this, quantum computation and quantum information, starting from the very basics, with all needed quantum physics preliminaries included. We will insist on both the physical and engineering aspects, in relation with the architecture of the quantum computers, and on the mathematics of the computing part. There are of course many other books on the subject, for all tastes and of all flavors, and in comparison for instance with Nielsen-Chuang [**69**], which is the overall classic, we will be doing here the same sort of thing, namely global introduction to the subject, by being however a bit stronger of mathematics, and also on some recent aspects.

I would like to thank my mathematics and physics colleagues for countless coffee room discussions about quantum computing, and this since Shor [**79**] and Grover [**46**], papers which came out right at the time when I was doing my PhD studies. It's been some time already, we know each other well, and in the hope that quantum computing will grow younger and fresher than me. Many thanks go as well to my cats, for some help with the mathematics. In the lack of a quantum computer, I usually ask them.

Contents

Part I

Computer science

Sweet child in time
You'll see the line
The line that's drawn between
Good and bad

CHAPTER 1

Games and strategy

1a. John von Neumann

Welcome to computing. Computing is not exactly about who does what, and how, but who does it faster. As an example, when you reach for your calculator, in order to avoid doing some complicated fractions in your mind, that is because you consider that, at least for that precise fraction task, the calculator will beat your brain.

In view of this, even before discussing what computers are, how they can be constructed, using various methods, and how they run, we must talk games and strategy. Computers can wait a bit, let us first understand, in the context of basic games, who wins. And then we will talk computers, with the idea in mind that these computers do not come alone, but rather in pairs, fighting over common computing questions, with the faster computer being the good one, and the slower computer being the bad one.

There is also some history, in relation with this. A lot of concepts that we will be talking about in this book, namely games and strategy, electric computers, abstract and applied quantum mechanics, and in a certain sense, the possibility of quantum computers too, are very related to John von Neumann, who worked on all these, among the pioneers. So, in order to understand all this, which looks like some sort of puzzle, bringing together seemingly unrelated branches of science, we must get to know a bit John von Neumann, who he was, what science he did, why he did that science, and so on.

There are of course many sources of information about John von Neumann, including books, videos, and the usual internet websites. So, in parallel with reading the present book, have a look at all that too, that would be my advice. Again, quantum computing that we're trying to understand will turn to be a complex and complicated business, and poetically speaking, somehow, what we want to understand is what exactly was on John von Neumann's mind. By the way, as usual with learning about someone, that is best done from the source. So, have a look as well at von Neumann's writings, and notably at his two well-known books, the one on the mathematics of quantum mechanics [85], and the one with Morgenstern on games and economy [86].

Getting started now, we need a first game to play, and we will choose, of course, something en hommage to von Neumann. The story with him was that he was not exactly your average tenebrous scientist, but rather the opposite. Loving life, people, food, drinks, parties, sharp dressing and so on, and fun in general. He used to work a bit everywhere, seemingly without much need for a regular, burial-type academic environement, such as at home, with the TV on, screaming news, or, interestingly, when driving his car.

So, inspired by this, here will be our first game, and I bet that von Neumann won't turn in his grave if I refer to this as the "von Neumann game":

Game 1.1 (von Neumann). *How to drive your car from A to B, as to:*

(1) *Get there as fast as possible.*
(2) *Not bother fellow motorists.*
(3) *Not disturb either the highway police.*
(4) *Have a pleasant ride, allowing you to do some math.*

Normally solving this type of game requires some modelling and mathematics, and we will learn how to do such things, in this chapter. However, coming a bit in advance, let me tell you the solution, because that is interesting, and good to know in the real life. There are two possible strategies, and ask anyone around having a bureau job, and commuting, they all know about it, and will confirm. These strategies are as follows:

(1) Drive slow. This will basically put you in the right lane, and you will surely miss (1), perform average with respect to (2), certainly have a big win at (3), and also perform quite well at (4), provided that you know how to absorb the inevitable honks.

(2) Drive fast. This will do the opposite, basically putting you in the left lane, with big win at (1), very good performance at (2), perhaps some bugs with respect to (3), and then big win at (4), what's more pleasant than a car running basically by itself.

So, this will be the type of game that we will discuss in this chapter, obviously needing some mathematical modelling and work, and with in the end several possible strategies, basically appearing as the local extrema of a certain function.

Of course, this is not the end of the story, because you might ask, what then about the best strategy, appearing by comparing the various local extrema, such as the above strategies (1) and (2), in relation with the von Neumann game. This is something that can be discussed too, of course. More on this later, and in the meantime just believe me, once you will read this book, and become a big fan of von Neumann, you will certainly get to know how to deal with Game 1.1, for the best of yourself, and of mankind.

1b. Games, saddle points

Games, saddle points.

1c. Minimax theorem

Minimax theorem.

1d. Life, economy

We discuss here some applications of the above to various life questions, including economy. Speaking economy, before everything, read Marx. You will be surprised. His main book, Capital, is something truly enjoyable, and very interesting, having not much to do with what is called today marxism, or with politics of any kind. Marx simply does, for the most of his book, a brilliant scientific study of various things which happened to mankind, from the perspective of money and capital, with the conclusion, which was totally new at that time, that money and capital are far more interesting, and involved in the nature of society, and in the things happening, than previously thought.

Which is absolutely brilliant. First-class scientific discovery. There are in fact 3 main modern discoveries, which provide the key to everything social science, namely:

(1) Darwin: when it comes to analyzing life, always have evolution in mind.

(2) Freud: when it comes to analyzing humans, always have sex in mind.

(3) Marx: when it comes to analyzing history, always have money in mind.

So, read what these 3 gentlemen had to say, that would be my advice, all this is first class reading, that you will certainly not regret. And in what regards Marx, yes of course at the end of his book he argues that, in view of his findings, a strong regulation of money and capital, socialist style, could probably solve some problems, and bring equilibrium to society. But that final part rather comes as an abstract conclusion, and is of course not the end of the story, because according to Darwin, and to Freud too, for various other reasons, human society is not meant to function on an equilibrium basis. So, to finish, definitely read Marx, and in what regards the end, things there are certainly more complicated than what he says, and you can even skip that part, if you disagree.

By the way, speaking social science reading, let me recommend as well Plato, Kant, Nietzsche, Eliade, Foucault, all sorts of interesting ideas inside. Also, do not forget to read the Bible too. You will be surprised, that is a very reasonable and interesting text, having nothing much to do with most of what modern Christian religion says.

Getting to math now, we will be interested in understanding some simple life and economy phenomena, by using our game theory knowledge. And, although there will be some money in what we will discuss, please be sure that there will be no evolution, sex and religion in all this, that is promised. This is rather advanced theory, coming after.

1e. Exercises

Exercises.

CHAPTER 2

Computers, war

CHAPTER 3

Coding, cryptography

CHAPTER 4

Networks, neurons

Part II

Quantum physics

Hey girls
Hey boys
Superstar DJ's
Here we go

CHAPTER 5

Quantum mechanics

5a. Advertisement

Welcome to quantum mechanics, that will help us build better computers. But what is quantum mechanics? The answer is very simple, as follows:

Definition 5.1. *Quantum mechanics is the science of everything, starting from the nanoscale level* (1 nm $= 10^{-9}$ m)*, and all the way below.*

This certainly deserves some explanations. Really everything? Why quantum? Why mechanics? And also, what does 10^{-9} have to do with all this?

In what regards "everything", yes the claim is that quantum mechanics has answers to everything happening at $d < 10^{-9}$, including mathematics, physics, chemistry, biology, engineering, and even things like philosophy, literature and sports. To be more precise, in what regards sports for instance, if one day we'll learn that everything is made of tiny little particles, far smaller than those presently known, who spend their time playing soccer and making fun of us, believe me, it is physicists doing quantum mechanics that will discover them, and not the sports anchors from your local TV station.

In what regards "mechanics", no surprise here, in view of what we've seen so far in this book. Everything in life is some sort of mechanics, with forces acting, experiments needed for writing equations, and then math needed for solving these equations. So, business as usual, with mechanics meaning more or less "physics". In fact, if you hear people talking about "quantum physics", that is exactly the same thing as quantum mechanics.

In what regards "quantum", things here are trickier. To start with, we have:

Terminology 5.2. *Quantum, plural quanta, comes from the Latin quantus, meaning "how much". In the context of physics, a quantum is the minimum possible amount of any physical quantity. And with quantity coming by the way from quantus, too.*

And isn't this confusing. Leaving now aside the Latin quantus, and the word quantity naturally derived from it, the key words in the above are "minimum possible amount". So, that is the precise definition of quantum, in physics parlance.

As an example here, money is quantized, that is, made of quanta, a quantum of money being 1 cent, if you live in the US. Sugar in a sugar box is quantized too, a quantum being here a cube of sugar. And so on. Note in passing that quantum has not necessarily something to do with "small", in the usual sense. For instance a herd of elephants is quantized too, a quantum here being 1 elephant. But things are relative, and assuming that you are interested in elephants only, 1 elephant is certainly something small.

Getting back now to physics, what is quantized, and what not? Common sense would suggest that all the basic physical quantities, such as distance, mass, energy and so on, vary continuously, and so are not quantized. And here comes the whole point, with quantum mechanics making the following bold statement:

Claim 5.3. *Quantum mechanics claims that all the basic physical quantities are in fact quantized, and that below the* 10^{-9} m *range, nothing cannot be really understood, if not taking into account the quantized nature of things.*

Summarizing, in regards with the questions raised after Definition 5.1, we have solved all of them, and by killing two rabbits with one shot, Claim 5.3 explaining both "quantum" and 10^{-9}. So, it simply remains to justify a bit this claim, and then get to work.

So, why should be things quantized in physics, and in life in general. Good question, going back to the cavern men, thinking about it. In the lack of anything spectacular, let us start with some philosophy. Our first piece of support for Claim 5.3 comes from:

Fact 5.4. *Life is quantized, the quanta being the cells, and not even need for a miscroscope for that, the orange cells for instance being big enough. That's how this world is made, quantized, or at least the fancy, living part of it. So if you look long enough at a rock, as to fall in love with that rock, that rock will become quantized too.*

It is possible to further build along these lines, with this being the occupation of mankind and philosophers for long millennia in a row. In order to reach however to something more precise, as in Claim 5.3, some modern physics is needed.

Summarizing, with some modern physics knowledge, we have some serious evidence for Claim 5.3. Or at least for the quantization claim there, with the precise figure 10^{-9} still needing to be discussed. And so, to end this discussion, Definition 5.1 seems to be justified, we know what quantum mechanics should be, and where its name comes from, and all that is left now is to find this quantum mechanics, its laws and everything.

The first thought goes to experiments, but here we stumble upon:

Fact 5.5. *You cannot really measure tiny little things, smaller than the resolution of your machinery. In addition, for the same reasons, measuring might perturb them.*

This is of course something as old as engineering, and the solution is always to wait for long years, for technology to evolve. But in our case, we are really looking for tiny little things, well beyond the range of usual scientific machinery, so don't really count on that machinery, and better try to develop some theory first.

But what to start with? Fortunately, there is an answer to this, as follows:

IDEA 5.6 (Let there be light). *Burning matter, and observing the color of the flame, gives you information about the intimate, infinitesimal structure of that matter.*

And isn't this amazing. This says more or less that when burning some gas, or wood, or salt, or whatever other substance, you're doing first-class quantum mechanics there, observing things so small that you never dreamed of coming upon.

So, this will be our starting point, in order to get into quantum mechanics, burning matter and recording the color of the light. We can in fact do even better, by avoiding the chemical reactions associated with burning, which will affect the matter that we are observing, and proceeding in a perhaps less glamorous way, as follows:

IDEA 5.7 (Let there be heat). *Heating matter, and observing the changing colors, gives you information about the intimate, infinitesimal structure of that matter.*

The above two ideas are of course as old as physics, or perhaps as fire and metallurgy, and the whole human civilization, in what concerns their everyday, macroscopic uses. Try cooking some food, or a blade, and you'll naturally get into them. In what concerns however their microscopic use, as suggested above, things are more recent, and the discipline of modern physics based on them is called spectroscopy.

5b. Light and heat

Getting started now, we first need to talk about light. This is actually far less obvious than it seems, involving some advanced physics. The idea is that light is a wave, and so, we first need to talk about waves. The result that we will need is as follows:

THEOREM 5.8. *The wave equation is*

$$\ddot{\varphi} = v^2 \Delta\varphi$$

where $\Delta = \sum_i d^2/dx_i^2$ *is the Laplace operator.*

PROOF. There are several proofs here, a nice one, by discretizing, being as follows:

(1) Let us first consider the 1D case. In order to understand the propagation of waves, we will model $\mathbb{R}$ as a network of balls, with springs between them, as follows:

$$\cdots \times\!\!\times\!\!\times \bullet \times\!\!\times\!\!\times \bullet \times\!\!\times\!\!\times \bullet \times\!\!\times\!\!\times \bullet \times\!\!\times\!\!\times \bullet \times\!\!\times\!\!\times \cdots$$

Now let us send an impulse, and see how balls will be moving. For this purpose, we zoom on one ball. The situation here is as follows, l being the spring length:

$$\cdots\cdots \bullet_{\varphi(x-l)} \times\!\!\times\!\!\times \bullet_{\varphi(x)} \times\!\!\times\!\!\times \bullet_{\varphi(x+l)} \cdots\cdots$$

We have two forces acting at x. First is the Newton motion force, mass times acceleration, which is as follows, with m being the mass of each ball:

$$F_n = m \cdot \ddot{\varphi}(x)$$

And second is the Hooke force, displacement of the spring, times spring constant. Since we have two springs at x, this is as follows, k being the spring constant:

$$\begin{aligned} F_h &= F_h^r - F_h^l \\ &= k(\varphi(x+l) - \varphi(x)) - k(\varphi(x) - \varphi(x-l)) \\ &= k(\varphi(x+l) - 2\varphi(x) + \varphi(x-l)) \end{aligned}$$

We conclude that the equation of motion, in our model, is as follows:

$$m \cdot \ddot{\varphi}(x) = k(\varphi(x+l) - 2\varphi(x) + \varphi(x-l))$$

(2) Now let us take the limit of our model, as to reach to continuum. For this purpose we will assume that our system consists of $N >> 0$ balls, having a total mass M, and spanning a total distance L. Thus, our previous infinitesimal parameters are as follows, with K being the spring constant of the total system, which is of course lower than k:

$$m = \frac{M}{N} \quad , \quad k = KN \quad , \quad l = \frac{L}{N}$$

With these changes, our equation of motion found in (1) reads:

$$\ddot{\varphi}(x) = \frac{KN^2}{M}(\varphi(x+l) - 2\varphi(x) + \varphi(x-l))$$

Now observe that this equation can be written, more conveniently, as follows:

$$\ddot{\varphi}(x) = \frac{KL^2}{M} \cdot \frac{\varphi(x+l) - 2\varphi(x) + \varphi(x-l)}{l^2}$$

With $N \to \infty$, and therefore $l \to 0$, we obtain in this way:

$$\ddot{\varphi}(x) = \frac{KL^2}{M} \cdot \frac{d^2\varphi}{dx^2}(x)$$

(3) In arbitrary N dimensions now, the same argument carries on, and we are led to the following equation, with $v = \sqrt{K/M} \cdot L$ being the propagation speed:

$$\ddot{\varphi}(x) = v^2 \sum_i \frac{d^2\varphi}{dx_i^2}(x)$$

But we recognize at right the Laplace operator, and we are done. There is of course some more discussion to be made here, arguing that our spring model in (1) is indeed the correct one, for modelling such wave propagation questions. But hey, we're doing theoretical physics here. And don't worry, experiments confirm our findings. $\square$

Going now towards light, we will need as well a second piece of physics, coming as a main course, if thinking that Theorem 5.8 was a cute appetizer:

THEOREM 5.9 (Maxwell theory). *In regions of space where there is no charge or current present the Maxwell equations for electrodynamics read*

$$< \nabla, E >=< \nabla, B >= 0$$

$$\nabla \times E = -\dot{B} \quad , \quad \nabla \times B = \dot{E}/c^2$$

and both the electric field E and magnetic field B are subject to the wave equation

$$\ddot{\varphi} = c^2 \Delta \varphi$$

where $\Delta = \sum_i d^2/dx_i^2$ is the Laplace operator, and $c = 299, 792, 458$.

PROOF. This is something fundamental, appearing as a tricky mixture of physics facts and mathematical results, the idea being as follows:

(1) To start with, electrodynamics is the science of moving electrical charges. And this is something quite complicated, because unlike in classical mechanics, where the Newton law is good for both the static and the dynamic setting, the Coulomb law, which is actually very similar to the Newton law, does the job when the charges are static, but no longer describes well the situation when the charges are moving.

(2) The problem comes from the fact that moving charges produce magnetism, and with this being visible when putting together two electric wires, which will attract or repel, depending on orientation. Thus, in contrast with classical mechanics, where static or dynamic problems are described by a unique field, the gravitational one, in electrodynamics we have two fields, namely the electric field E, and the magnetic field B.

(3) Fortunately, there is a full set of equations relating the electric field E and the magnetic field B. These are the Maxwell equations, which look as follows:

$$< \nabla, E >= \frac{\rho}{\varepsilon_0} \quad , \quad < \nabla, B >= 0$$

$$\nabla \times E = -\dot{B} \quad , \quad \nabla \times B = \mu_0 J + \mu_0 \varepsilon_0 \dot{E}$$

(4) To be more precise, regarding first the math, the dots denote derivatives with respect to time, and ∇ is the gradient operator, or space derivative, given by:

$$\nabla = \begin{pmatrix} \frac{d}{dx} \\ \frac{d}{dy} \\ \frac{d}{dz} \end{pmatrix}$$

(5) Regarding the physics, the first formula is the Gauss law, ρ being the charge, and ε_0 being a constant, and with this Gauss law more or less replacing the Coulomb law from electrostatics. The second formula is something basic, and anonymous. The third

formula is the Faraday law. As for the fourth formula, this is the Ampère law, as modified by Maxwell, with J being the volume current density, and μ_0 being a constant.

(6) Without bothering too much about the precise meaning of all this, we can see right away that under the circumstances in the statement, namely in the regions of space where there is no charge or current present, the Maxwell equations take a simple and comprehensible form, readable even without much physics background, namely:

$$< \nabla, E >=< \nabla, B >= 0$$

$$\nabla \times E = -\dot{B} \quad , \quad \nabla \times B = \mu_0\varepsilon_0\dot{E}$$

(7) Thus, we have reached to the equations in the statement, modulo a discussion about the constant $\mu_0\varepsilon_0$. And the point here is that, according to a remarkable discovery of Biot and Savart, the main electrodynamics constants μ_0, ε_0 are magically related to the observed speed of light in vacuum $c = 299, 792, 458$ by the following formula:

$$\mu_0\varepsilon_0 = \frac{1}{c^2}$$

(8) Summarizing, we have our equations. Leaving aside the first two equations, by applying the curl operator to the last two equations, we obtain:

$$\nabla \times (\nabla \times E) = -\nabla \times \dot{B} = -(\nabla \times B)' = -\ddot{E}/c^2$$

$$\nabla \times (\nabla \times B) = \nabla \times \dot{E}/c^2 = (\nabla \times E)'/c^2 = -\ddot{B}/c^2$$

But the double curl operator is subject to the following formula:

$$\nabla \times (\nabla \times \varphi) = \nabla < \nabla, \varphi > -\Delta\varphi$$

Now by using the first two equations, we are led to the conclusion in the statement. □

So, what is light? Light is the wave predicted by Theorem 5.9, travelling at speed c, and with an important extra property being that it depends on a real positive parameter, that can be called, upon taste, frequency, wavelength, or color.

In what regards the creation of light, the mechanism here is as follows:

Fact 5.10. *An accelerating or decelerating charge produces electromagnetic radiation, called light, whose frequency and wavelength can be explicitely computed.*

This phenomenon can be observed is a variety of situations, such as the usual light bulbs, where electrons get decelerated by the filament, acting as a resistor, or in usual fire, which is a chemical reaction, with the electrons moving around, as they do in any chemical reaction, or in more complicated machinery like nuclear plants, particle accelerators, and so on, leading there to all sorts of eerie glows, of various colors.

In view of the above, and especially of the light bulb example, a natural question appears: what about a resistor which is not a light bulb filament, what happens to the light produced there? This is a good question, and we already know a part of answer to it, from the Joule law, saying that the resistor will start heating. The second part of it, that we will discuss in a moment, states that responsible for heat is, guess who, light again, but this time in non-visible wavelenghts, typically IR.

Getting back now to Fact 5.10, in general, as stated above, this is something which can be deduced via some math, based on the Maxwell equations. However, all this math is not exactly trivial, and for details here, you can see for instance Griffiths [**42**].

Let us go back now to the wave equation $\ddot{\varphi} = v^2\Delta\varphi$ from Theorem 5.9, and try to understand its simplest solutions. In 1D, the situation is as follows:

THEOREM 5.11. *The 1D wave equation, with speed v, namely*

$$\ddot{\varphi} = v^2 \frac{d^2\varphi}{dx^2}$$

has as basic solutions the following functions,

$$\varphi(x) = A\cos(kx - wt + \delta)$$

with A being called amplitude, $kx-wt+\delta$ being called the phase, k being the wave number, w being the angular frequency, and δ being the phase constant. We have

$$\lambda = \frac{2\pi}{k} \quad , \quad T = \frac{2\pi}{kv} \quad , \quad \nu = \frac{1}{T} \quad , \quad w = 2\pi\nu$$

relating the wavelength λ, period T, frequency ν, and angular frequency w. Moreover, any solution of the wave equation appears as a linear combination of such basic solutions.

PROOF. There are several things going on here, the idea being as follows:

(1) Our first claim is that the function φ in the statement satisfies indeed the wave equation, with speed $v = w/k$. For this purpose, observe that we have:

$$\ddot{\varphi} = -w^2\varphi \quad , \quad \frac{d^2\varphi}{dx^2} = -k^2\varphi$$

Thus, the wave equation is indeed satisfied, with speed $v = w/k$:

$$\ddot{\varphi} = \left(\frac{w}{k}\right)^2 \frac{d^2\varphi}{dx^2} = v^2 \frac{d^2\varphi}{dx^2}$$

(2) Regarding now the other things in the statement, all this is basically terminology, which is very natural, when thinking how $\varphi(x) = A\cos(kx - wt + \delta)$ propagates.

(3) Finally, the last assertion is something standard, coming from Fourier analysis, that we will not really need, in what follows. □

As a first observation, the above result invites the use of complex numbers. Indeed, we can write the solutions that we found in a more convenient way, as follows:

$$\varphi(x) = Re\left[A\, e^{i(kx-wt+\delta)}\right]$$

And we can in fact do even better, by absorbing the quantity $e^{i\delta}$ into the amplitude A, which becomes now a complex number, and writing our formula as:

$$\varphi = Re(\widetilde{\varphi}) \quad , \quad \widetilde{\varphi} = \widetilde{A}e^{i(kx-wt)}$$

Moving ahead now towards electromagnetism and 3D, let us formulate:

DEFINITION 5.12. *A monochromatic plane wave is a solution of the 3D wave equation which moves in only 1 direction, making it in practice a solution of the 1D wave equation, and which is of the special from found in Theorem 5.11, with no frequencies mixed.*

In other words, we are making here two assumptions on our wave. First is the 1-dimensionality assumption, which gets us into the framework of Theorem 5.11. And second is the assumption, in connection with the Fourier decomposition result from the end of Theorem 5.11, that our solution is of "pure" type, meaning a wave having a well-defined wavelenght and frequency, instead of being a "packet" of such pure waves.

All this is still mathematics, and making now the connection with physics and electromagnetism, and more specifically with Theorem 5.9 and Fact 5.10, we have:

FACT 5.13. *Physically speaking, a monochromatic plane wave is the electromagnetic radiation appearing as in Theorem 5.9 and Fact 5.10, via equations of type*

$$E = Re(\widetilde{E}) \quad : \quad \widetilde{E} = \widetilde{E}_0\, e^{i(<k,x>-wt)}$$

$$B = Re(\widetilde{B}) \quad : \quad \widetilde{B} = \widetilde{B}_0\, e^{i(<k,x>-wt)}$$

with the wave number being now a vector, $k \in \mathbb{R}^3$. Moreover, it is possible to add to this an extra parameter, accounting for the possible polarization of the wave.

To be more precise, what we are doing here is to import the conclusions of our mathematical discussion so far, from Theorem 5.11 and Definition 5.12, into the context of our original physics discussion, from Theorem 5.9 and Fact 5.10. And also to add an extra twist coming from physics, and more specifically from the notion of polarization.

In any case, we have now a decent intuition about what light is, and more on this later, and let us discuss now the examples. The idea is that we have various types of light, depending on frequency and wavelength. These are normally referred to as "electromagnetic waves", but for keeping things simple and luminous, we will keep using the familiar term "light". The classification, in a rough form, is as follows:

Frequency	Type	Wavelength
	–	
$10^{18} - 10^{20}$	γ rays	$10^{-12} - 10^{-10}$
$10^{16} - 10^{18}$	X – rays	$10^{-10} - 10^{-8}$
$10^{15} - 10^{16}$	UV	$10^{-8} - 10^{-7}$
	–	
$10^{14} - 10^{15}$	blue	$10^{-7} - 10^{-6}$
$10^{14} - 10^{15}$	yellow	$10^{-7} - 10^{-6}$
$10^{14} - 10^{15}$	red	$10^{-7} - 10^{-6}$
	–	
$10^{11} - 10^{14}$	IR	$10^{-6} - 10^{-3}$
$10^{9} - 10^{11}$	microwave	$10^{-3} - 10^{-1}$
$1 - 10^{9}$	radio	$10^{-1} - 10^{8}$

Observe the tiny space occupied by the visible light, all colors there, and the many more missing, being squeezed under the $10^{14} - 10^{15}$ frequency banner. Here is a zoom on that part, with of course the remark that all this, colors, is something subjective:

Frequency THz $= 10^{12}$ Hz	Color	Wavelength nm $= 10^{-9}$ m
	–	
670 – 790	violet	380 – 450
620 – 670	blue	450 – 485
600 – 620	cyan	485 – 500
530 – 600	green	500 – 565
510 – 530	yellow	565 – 590
480 – 510	orange	590 – 625
400 – 480	red	625 – 750

Outside visible light we have, as you probably know it, UV on higher frequencies, and IR on lower frequencies. At the high frequency end we have X-rays, that you surely know about too, and γ rays, which are usually associated with various bad things, such as thunderstorms, solar flares, and small bugs with our nuclear energy technology.

As for the lower frequency end of the scale, first we have microwaves, but if you love physics and chemistry you should learn some cooking, that's first-class chemistry, that you can practice every day. And then we have all sorts of radio wavelenghts, including FM, followed by AM, and then by several more obscure low-frequency waves.

Back now to our business, with all the above in hand, we can do some optics. Light usually comes in "bundles", with waves of several wavelenghts coming at the same time, from the same source, and the first challenge is that of separating these wavelenghts. In order to discuss this, let us start with the following fact:

FACT 5.14. *Inside a linear, homogeneous medium, where there is no free charge or current present, the Maxwell equations for electrodynamics read*

$$<\nabla, E>=<\nabla, B>=0$$

$$\nabla \times E = -\dot{B} \quad , \quad \nabla \times B = \varepsilon\mu\dot{E}$$

with E, B being as before the electric and the magnetic field, and with $\varepsilon > \varepsilon_0$ and $\mu > \mu_0$ being the electric permittivity and magnetic permeability of the medium.

Observe that this is precisely the first part of Theorem 5.9, with the vacuum constants ε_0, μ_0 being replaced by their versions ε, μ, concerning the medium in question. In what regards now the second part of Theorem 5.9, which was a theorem, we have:

THEOREM 5.15. *Inside a linear, homogeneous medium, where there is no free charge or free current present, both E and B are subject to the wave equation*

$$\ddot{\varphi} = v^2 \Delta\varphi$$

with v being the speed of light inside the medium, given by

$$v = \frac{c}{n} \quad : \quad n = \sqrt{\frac{\varepsilon\mu}{\varepsilon_0\mu_0}}$$

with the quantity on the right $n > 1$ being called refraction index of the medium.

PROOF. This is something that we know well in vacuum, and the proof in general is identical, with the resulting speed being:

$$v = \frac{1}{\sqrt{\varepsilon\mu}}$$

But this formula can be written is a more familiar from, as above. □

As a first observation here, while the above is something quite trivial, mathematically speaking, from the physical viewpoint we are here into complicated things. Materials can be transparent or opaque, with the distinction between them being something very subtle, and advanced, and Theorem 5.15 obviously deals with the transparent case.

Next in line, and for interest for us, we have:

FACT 5.16. *When travelling through a material, and hitting a new material, some of the light gets reflected, at the same angle, and some of it gets refracted, at a different angle, depending both on the old and the new material, and on the wavelength.*

Again, this is something deep, and very old as well, and there are many things that can be said here, ranging from various computations based on the Maxwell equations, to all sorts of considerations belonging to advanced materials theory.

As a basic formula here, we have the famous Snell law, which relates the incidence angle θ_1 to the refraction angle θ_2, via the following simple formula:

$$\frac{\sin\theta_2}{\sin\theta_1}=\frac{n_1(\lambda)}{n_2(\lambda)}$$

Here $n_i(\lambda)$ are the refraction indices of the two materials, adjusted for the wavelength, and with this adjustment for wavelength being the whole point, which is something quite complicated. For an introduction to all this, we refer for instance to Griffiths [**42**].

As a simple consequence of the above, we have:

Theorem 5.17. *Light can be decomposed, by using a prism.*

Proof. This follows from Fact 5.16. Indeed, when hitting a piece of glass, provided that the hitting angle is not 90°, the light will decompose over the wavelenghts present, with the corresponding refraction angles depending on these wavelengths. And we can capture these split components at the exit from the piece of glass, again deviated a bit, provided that the exit surface is not parallel to the entry surface. And the simplest device doing the job, that is, having two non-parallel faces, is a prism. □

With this in hand, we can now talk about spectroscopy:

Fact 5.18. *We can study events via spectroscopy, by capturing the light the event has produced, decomposing it with a prism, carefully recording its "spectral signature", consisting of the wavelenghts present, and their density, and then doing some reverse engineering, consisting in reconstructing the event out of its spectral signature.*

This is the main principle of spectroscopy, and applications, of all kinds, abound. In practice, the mathematical tool needed for doing the "reverse engineering" mentioned above is the Fourier transform, which allows the decomposition of packets of waves, into monochromatic components. Finally, let us mention too that, needless to say, the event can be reconstructed only partially out of its spectral signature.

5c. Spectrum, theory

Getting now back to atoms, there is a long story here, involving many discoveries of many people, around 1890-1900, focusing on hydrogen H. We will present here things a bit retrospectively, as to bet fit with science as we know it now, and with the present book. First on our list is the following discovery, by Lyman in 1906:

FACT 5.19 (Lyman). *The hydrogen atom has spectral lines given by the formula*

$$\frac{1}{\lambda} = R\left(1 - \frac{1}{n^2}\right)$$

where $R \simeq 1.097 \times 10^7$ and $n \geq 2$, which are as follows,

n	Name	Wavelength	Color
	–	–	
2	α	121.567	UV
3	β	102.572	UV
4	γ	97.254	UV
$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	limit	91.175	UV

called Lyman series of the hydrogen atom.

Observe that all the Lyman series lies in UV. Due to this fact, namely the invisibility of UV to the human eye, this series, while theoretically being the most important, for certain reasons to be explained later, was discovered only second.

The first discovery, which was the big one, and the breakthrough, was by Balmer, the founding father of all this, back in 1885, in the visible range, as follows:

FACT 5.20 (Balmer). *The hydrogen atom has spectral lines given by the formula*

$$\frac{1}{\lambda} = R\left(\frac{1}{4} - \frac{1}{n^2}\right)$$

where $R \simeq 1.097 \times 10^7$ and $n \geq 3$, which are as follows,

n	Name	Wavelength	Color
	–	–	
3	α	656.279	red
4	β	486.135	aqua
5	γ	434.047	blue
6	δ	410.173	violet
7	ε	397.007	UV
$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	limit	346.600	UV

called Balmer series of the hydrogen atom.

So, this was Balmer's original result, which started everything, and with his original wavelength formula being in fact something equivalent to the above formula, but a bit more complicated, as follows, with $B \simeq 3.645 \times 10^{-7}$ being the Balmer constant:

$$\lambda = \frac{Bn^2}{n^2 - 4}$$

As a third main result now, this time in IR, due to Paschen in 1908, we have:

FACT 5.21 (Paschen). *The hydrogen atom has spectral lines given by the formula*

$$\frac{1}{\lambda} = R\left(\frac{1}{9} - \frac{1}{n^2}\right)$$

where $R \simeq 1.097 \times 10^7$ *and* $n \geq 4$, *which are as follows,*

n	Name	Wavelength	Color
	–	–	
4	α	1875	IR
5	β	1282	IR
6	γ	1094	IR
⋮	⋮	⋮	⋮
∞	limit	820.4	IR

called Paschen series of the hydrogen atom.

Observe the striking similarity between the above three results. In fact, we have here the following fundamental, grand result, due to Rydberg in 1888, based on the Balmer series, and with later contributions by Ritz in 1908, using the Lyman series as well:

CONCLUSION 5.22 (Rydberg, Ritz). *The spectral lines of the hydrogen atom are given by the Rydberg formula, depending on integer parameters* $n_1 \leq n_2$,

$$\frac{1}{\lambda_{n_1 n_2}} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$

with R *being the Rydberg constant for hydrogen, which is as follows:*

$$R \simeq 1.096\ 775\ 83 \times 10^7$$

These spectral lines combine according to the Ritz-Rydberg principle, as follows:

$$\frac{1}{\lambda_{n_1 n_2}} + \frac{1}{\lambda_{n_2 n_3}} = \frac{1}{\lambda_{n_1 n_3}}$$

Similar formulae hold for other atoms, with suitable fine-tunings of R.

Here the first part, the Rydberg formula, generalizes the results of Lyman, Balmer, Paschen, which appear at $n_1 = 1, 2, 3$, at least retrospectively. The Rydberg formula predicts further spectral lines, appearing at $n_1 = 4, 5, 6, \ldots$, and these were discovered later, by Brackett in 1922, Pfund in 1924, Humphreys in 1953, and others aftwerwards, with all these extra lines being in far IR. The simplified complete table is as follows:

n_1	n_2	Series name	Wavelength $n_2 = \infty$	Color $n_2 = \infty$
		–	–	
1	$2 - \infty$	Lyman	91.13 nm	UV
2	$3 - \infty$	Balmer	364.51 nm	UV
3	$4 - \infty$	Paschen	820.14 nm	IR
		–	–	
4	$5 - \infty$	Brackett	1458.03 nm	far IR
5	$6 - \infty$	Pfund	2278.17 nm	far IR
6	$7 - \infty$	Humphreys	3280.56 nm	far IR
...	...	...	...	...

Regarding the last assertion, concerning other elements, this is something conjectured and partly verified by Ritz, and fully verified and clarified later, via many experiments, the fine-tuning of R being basically $R \to RZ^2$, where Z is the atomic number.

But from a theoretical physics viewpoint, the main result remains the middle assertion, called Ritz-Rydberg combination principle. This is something at the same time extremely simple, and completely puzzling, the informal conclusion being as follows:

Thought 5.23. *The simplest observables of the hydrogen atom, combining via*

$$\frac{1}{\lambda_{n_1n_2}} + \frac{1}{\lambda_{n_2n_3}} = \frac{1}{\lambda_{n_1n_3}}$$

look like quite weird quantities. Why wouldn't they just sum normally.

Getting now to quantum mechanics, and to our dreams about it, formulated before, well, good news, we have some serious data here. These spectral lines are basic and beautiful, obviously of quantized type, and in order to get started with our theory, we first need to solve the puzzle of the Ritz-Rydberg combination principle.

But, how to do this? Fortunately, matrix theory comes to the rescue, as follows:

Thought 5.24. *The Ritz-Rydberg combination principle reminds the formula*

$$e_{n_1n_2}e_{n_2n_3} = e_{n_1n_3}$$

for the usual matrix units, which are the elementary matrices given by

$$e_{ij} : e_j \to e_i$$

perhaps taken in infinite dimensions, as to allow infinite-ranging indices.

In short, we are in familiar territory here, and we can start dreaming of:

Thought 5.25. *Observables in quantum mechanics should be some sort of infinite matrices, generalizing the Lyman, Balmer, Paschen lines of the hydrogen atom, and multiplying between them as the matrices do, as to produce further observables.*

And probably enough for now, this is the kind of discovery that should be celebrated with slaughtering 50 sheep and inviting your friends over, for a banquet, as the legend goes that Pythagoras did, after he discovered his $a^2 + b^2 = c^2$ theorem.

5d. Atomic structure

Time now to put everything together. As a main problem that we would like to solve, we have the understanding the intimate structure of matter, at the atomic level. There is of course a long story here, regarding the intimate structure of matter, going back centuries and even millennia ago, and our presentation here will be quite simplified. As a starting point, since we need a starting point, let us agree on:

Claim 5.26. *Ordinary matter is made of small particles called atoms, with each atom appearing as a mix of even smaller particles, namely protons* $+$*, neutrons* 0 *and electrons* $-$*, with the same number of protons* $+$ *and electrons* $-$*.*

As a first observation, this is something which does not look obvious at all, with probably lots of work, by many people, being involved, as to lead to this claim. And so it is. The story goes back to the discovery of charges and electricity, which were attributed to a small particle, the electron $-$. Now since matter is by default neutral, this naturally leads to the consideration to the proton $+$, having the same charge as the electron.

But why should be these electrons $-$ and protons $+$ that small? And also, what about the neutron 0? These are not easy questions, and the fact that indeed it is so came from several clever experiments. Let us first recall from Fact 5.5 that careful experiments with tiny particles are practically impossible. However, all sorts of brutal experiments, such as bombarding matter with other pieces of matter, accelerated to the extremes, or submitting it to huge electric and magnetic fields, do work. And it is such kind of experiments, due to Thomson, Rutherford and others, "peeling off" protons $+$, neutrons 0 and electrons $-$ from matter, and observing them, that led to the conclusion that these small beasts $+, 0, -$ exist indeed, in agreement with Claim 5.26.

Of particular importance here was as well the radioactivity theory of Becquerel and Pierre and Marie Curie, involving this time such small beasts, or perhaps some related radiation, peeling off by themselves, in heavy elements such as uranium $_{92}U$, polonium $_{84}Po$ and radium $_{88}Ra$. And there was also Einstein's work on the photoelectric effect, light interacting with matter, suggesting that even light itself might have associated to it some kind of particle, called photon. All this goes of course beyond Claim 5.26, with further particles involved, and more on this later, but as a general idea, all this deluge of small particle findings, all coming around 1900-1910, further solidified Claim 5.26.

So, taking now Claim 5.26 for granted, how are then the atoms organized, as mixtures of protons $+$, neutrons 0 and electrons $-$? The answer here lies again in the above-mentioned "brutal" experiments of Thomson, Rutherford and others, which not only proved Claim 5.26, but led to an improved version of it, as follows:

Claim 5.27. *The atoms are formed by a core of protons* $+$ *and neutrons* 0*, surrounded by a cloud of electrons* $-$*, gravitating around the core.*

This is a considerable advance, because we are now into familiar territory, namely some kind of mechanics. Remember from the beginning of this book the planets orbiting around the Sun, on ellipses, as found by Kepler? Well, the same should happen with electrons orbiting around the core, but this time due to the Coulomb force. And with this in mind, all the pieces of our puzzle start fitting together, and lead to:

Claim 5.28 (Bohr and others). *The atoms are formed by a core of protons and neutrons, surrounded by a cloud of electrons, basically obeying to a modified version of electromagnetism. And with a fine mechanism involved, as follows:*

(1) *The electrons are free to move only on certain specified elliptic orbits, labelled* $1, 2, 3, \ldots$*, situated at certain specific heights.*
(2) *The electrons can jump or fall between orbits* $n_1 < n_2$*, absorbing or emitting light and heat, that is, electromagnetic waves, as accelerating charges.*
(3) *The energy of such a wave, coming from* $n_1 \to n_2$ *or* $n_2 \to n_1$*, is given, via the Planck viewpoint, by the Rydberg formula, applied with* $n_1 < n_2$*.*
(4) *The simplest such jumps are those observed by Lyman, Balmer, Paschen. And multiple jumps explain the Ritz-Rydberg formula.*

And isn't this beautiful. Moreover, some further claims, also by Bohr and others, are that the theory can be further extended and fine-tuned as to explain many other phenomena, such as the above-mentioned findings of Einstein, and of Becquerel and Pierre and Marie Curie, and generally speaking, all the physics and chemistry known.

And the story is not over here. Following now Heisenberg, the next claim is that the underlying math in all the above can lead to a beautiful axiomatization of quantum mechanics, as a "matrix mechanics", along the lines of Thought 5.25.

We will be back to this, with more details, in the next chapter.

5e. Exercises

Exercises.

CHAPTER 6

Particle spin

6a. Stern-Gerlach

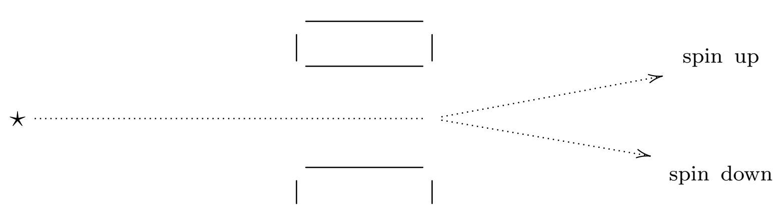
We have seen that the fine structure correction to the Bohr energy formula involved two things. First we had a relativistic correction, that we have not worked out in detail, but which looks like something within our range, that can only be understood and computed. But then we had as well as spin-related correction, involving the notion of spin, which is totally new to us. So, as a first question that we would like to solve, we have:

QUESTION 6.1. *What is the electron spin? That is, what experiments prove that the electron spins? And then, importantly, what is the mathematics of the spin?*

Talking mathematics first, the spin, if that beast exists indeed, is certainly not visible on the wave function ψ, because this wave function deals with position only. Thus, at least we know one thing, once the spin observed, we will most likely have to incorporate it into our theory by using the matrix mechanics formalism of Heisenberg.

Talking physics now, the main experiment leading to spin is as follows:

FACT 6.2 (Stern-Gerlach experiment). *When passing a beam of electrons through an inhomogeneous magnetic field, these electrons get deflected* $50-50$ *up or down,*



with the only possible explanation being that the electrons have a spin, which is $50-50$ *up or down. The same happens with a beam of neutral atoms, and a magnetic field strong enough, to be put at blame being the statistics of the spins of the constituents.*

So, this was the experiment, and what we call here "up" and "down" is of course the binary choice of the spin orientation, a bit as for usual, round objects in $\mathbb{R}^3$. That is, our Earth turns to the right, and in physics we would say that is has "spin up". Was the Earth turning to the left, we would say in physics that is has "spin down".

Of course, our presentation above is over-simplified. The original experiment was with neutral particles, namely silver atoms, and this in order to avoid the Lorentz force, which will curve the trajectory of any charged particle, to a much greater extent than the spin up/down deviation to be observed. Later experiments, with charged particles, used some extra apparatus, namely a suitable electric field, positioned after the electromagnet in the above diagram, designed to cancel the effects of the Lorentz force.

As an important observation, the Stern-Gerlach experiment does not observe the absolute, 3D spin up/down feature of the particles, but just a 1D component of it. However, it is possible to cascade experiments, by sending each of the output beams into separate Stern-Gerlach devices, and with these devices having various 3D orientations, and deduce some further conclusions from this. We refer here to Feynman [**35**].

So long for the Stern-Gerlach experiment. Getting back now to theory and speculations, as a first, innocent observation based on the above, we have:

Observation 6.3. *A single electron has an interesting life even when fixed, because it spins. Thus, no need for Heisenberg or Schrödinger for getting introduced to quantum mechanics, you can just try to understand the mathematics of a fixed electron.*

Moreover, as a cherry on the cake, as we will soon discover, the above-mentioned mathematics is that of the 2×2 complex matrices, which is at the same time something elementary, and fascinating. Which, getting us now into philosophy, leads us into the temptation of burying the physics, and talking right away about 2×2 matrices.

And shall we do this or not. Looking at the physics literature, there is a fair mess in the treatment of spin. At one end, we have books like Townsend [**83**], taking Observation 6.3 literally, and starting the book with a long, not to say never-ending, discussion about spin. Then we have quantum information related books, such as Bengtsson-Życzkowski [**7**], Nielsen-Chuang [**69**], Peres [**70**], which by a certain desire of brevity and efficiency, rapidly bury the physics of spin, and talk instead about 2×2 matrices. And then we have well-known and loved books such as Feynman [**35**], Griffiths [**43**], Weinberg [**90**], presenting all sorts of rather incomprehensible explanations regarding the spin, which vary with authors' taste, for eventually ending, of course, with 2×2 matrices.

And so again, what shall we do, talk about 2×2 matrices or not. Not clear. But, as usual in such difficult situations, we can always ask the cat. And cat answers:

Advice 6.4. *Be honest, and say what you have to say. And don't worry about your young readers, they will survive.*

This sounds wise as usual, thanks cat. So, we will follow this advice. But let me get first a huge mug of coffee, or rather huge mug of expresso, because fighting with the physics of the spin with bare hands is something which is reputed impossible.

To start with, and as a matter of reframing our discussion, and having something fresh to rely upon, let us demolish Observation 6.3 above with:

FACT 6.5. *Observation 6.3 is something toxic. You can't really measure spin, and build a serious theory on that alone. What you need to do is to observe spin in context, via its tiny corrections to quantum mechanics. More specifically, spin is an order*

$$\alpha^2 \simeq \frac{1}{10,000}$$

correction to quantum mechanics, and more precisely to the Bohr energy formula, with the spin correction there appearing as a complement to the relativistic correction. And with this being the correct, healthy and constructive definition of spin.

In short, we are getting here back to the beginning, general quantum mechanics, with the main conclusion of the Stern-Gerlach experiment, namely "spin exists", recorded. Of course it is possible to say a bit more from Stern-Gerlach, namely recording the scattering angle, and doing some math there, but this basically does not advance us much. So better forget about Stern-Gerlach, and get back to general quantum mechanics.

The point now is that, with the above fact in hand, not only we are into truth, as we should be, but also we start getting an idea on how to reach to the mathematics of the spin. To be more precise, we should just think relativity, in the context of quantum mechanics, and with a bit of luck, all this thinking will lead us into spin.

In practice now, all this is doable, but a bit complicated, and was done by Klein, Gordon, Dirac a few years after Uhlenbeck, Goudsmit, Pauli came up with their theory of spin. So, let us briefly explain this idea, which is very beautiful, and we'll come later to Uhlenbeck, Goudsmit, Pauli. Consider the Schrödinger equation for a free electron:

$$ih\dot{\psi} = -\frac{h}{2m}\Delta\psi$$

Relativity theory dictates that the 3 space coordinates and the 1 time coordinate should be on the same footing, and so that we should replace $\dot{\psi}$ by something of type $\ddot{\psi}$. But this can be done by replacing the kinetic energy operator $T = \Delta/2m$ by its relativistic analogue, and also by invoking the invariance under Lorentz transformations, and we are led in this way to the following equation, called Klein-Gordon equation:

$$\left(\Delta - \frac{1}{c^2}\cdot\frac{d^2}{dt^2}\right)\psi = \frac{m^2c^2}{h^2}\psi$$

The point now, which is the key one, discovered by Dirac short after Klein and Gordon, is that it is possible to extract the square root of the Klein-Gordon operator:

$$\Delta - \frac{1}{c^2}\cdot\frac{d^2}{dt^2} = \left(\frac{Pd}{dx} + \frac{Qd}{dy} + \frac{Rd}{dz} + \frac{i}{c}\cdot\frac{Sd}{dt}\right)^2$$

Indeed, we need for this purpose matrices P, Q, R, S which anticommute, $AB = -BA$, and whose squares equal one, $A^2 = 1$. But such beasts can be found in $M_4(\mathbb{C})$, and then we can take the formal square root of the Klein-Gordon equation:

$$\left(\frac{Pd}{dx} + \frac{Qd}{dy} + \frac{Rd}{dz} + \frac{i}{c} \cdot \frac{Sd}{dt}\right)\psi = \frac{mc}{h}\psi$$

And the thing now, which is truly remarkable, is that this latter equation, called Dirac equation, does work indeed, in the sense that it is a true equation of physics, improving the Schrödinger equation. And a closer look at all this reveals afterwards that the fine structure of hydrogen, comprising the relativistic correction and the spin correction, can be understood in this way, leading to a clear mathematics of the spin.

All this is very beautiful, and leads us into:

THOUGHT 6.6. *Our criticism from Fact 6.5 was probably too harsh, relativity and spin alike being probably more than a mere*

$$\alpha^2 \simeq \frac{1}{10,000}$$

order correction to quantum mechanics. And this is because the Dirac equation, which is of first order, is something simpler than the Schrödinger equation.

In fact, we're now again into Observation 6.3, and this time armed with some solid math, and more specifically with a first-grade weapon, the Dirac equation. Which starts to be a bit tiring, yes I know, looks like we're changing our opinion about spin faster than Madonna is changing her shoes. But blame the cat, he came with Advice 6.4.

Moving ahead now, and still following Advice 6.4, after some more thinking, the Dirac equation remains however something a bit speculative, or perhaps something too advanced, and it would be much better, at least to start with, to forget about relativity and abstractions, and have something more solid, regarding the spin.

And fortunately, there is a second way of viewing things, very elementary, inspired from our study of classical mechanics, or even from the movement of our good old Earth, which rotates and spins at the same time, which is as follows:

PHILOSOPHY 6.7. *In analogy with classical mechanics, spin should be something of same nature as angular momentum, coming on top of it.*

And good news, this will be our final, stable philosophy. Eventually.

To be more precise, following Uhlenbeck, Goudsmit, Pauli, we will first talk angular momentum, then we will axiomatize spin as being the quantity which naturally "complements" the angular momentum. Then we will talk about 2×2 matrices, and more.

6b. Math, operators

So, let us first talk about angular momentum, and we'll get to spin later. We will need the following basic result, for doing computations:

PROPOSITION 6.8. *The components of the position operator $x = (x_1, x_2, x_3)$ and momentum operator $p = -ih\nabla$ satisfy the following relations,*

$$[x_i, x_j] = [p_i, p_j] = 0$$

$$[x_i, p_j] = ih\delta_{ij}$$

where $[a, b] = ab - ba$, called canonical commutation relations.

PROOF. All the above formulae are elementary, as follows:

(1) The components of the position operator $x = (x_1, x_2, x_3)$ obviously commute with each other, $x_i x_j = x_j x_i$, which makes their commutators vanish, $[x_i, x_j] = 0$.

(2) Regarding the momentum operator $p = -ih\nabla$, its components are as follows:

$$p_1 = -ih \cdot \frac{d}{dx_1} \quad , \quad p_2 = -ih \cdot \frac{d}{dx_2} \quad , \quad p_3 = -ih \cdot \frac{d}{dx_3}$$

Since partial derivatives commute with each other, we obtain $[p_i, p_j] = 0$.

(3) It remains to prove the last formula, and we have here:

$$\begin{aligned}
[x_i, p_j]f &= (x_i p_j - p_j x_i)f \\
&= -ih\left(x_i \cdot \frac{df}{dx_j} - \frac{d}{dx_j}(x_i f)\right) \\
&= -ih\left(x_i \cdot \frac{df}{dx_j} - \frac{dx_i}{dx_j} \cdot f - x_i \cdot \frac{df}{dx_j}\right) \\
&= ih \cdot \frac{dx_i}{dx_j} \cdot f \\
&= ih\delta_{ij} \cdot f
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

The above might look a bit complicated, and the simplest way to remember it is that "everyhing commutes", that is, $ab = ba$, except for the coordinates and momenta coordinates taken in the same direction, which are subject to the following rule:

$$x_i p_i = p_i x_i + ih$$

Getting now to angular momentum, it is convenient to change notation, with (x, y, z) instead of (x_1, x_2, x_3), due to the vector product involved, which will break the symmetry between coordinates. We have the following result, to start with:

THEOREM 6.9. *The components of the angular momentum operator*

$$L = x \times (-ih\nabla)$$

satisfy the following equations,

$$[L_x, L_y] = ihL_z$$
$$[L_y, L_z] = ihL_x$$
$$[L_z, L_x] = ihL_y$$

called commutation relations for the angular momentum.

PROOF. With the more familiar notation $p = -ih\nabla$ for momentum, or rather for the associated operator, the components of the angular momentum operator are:

$$L_x = yp_z - zp_y$$
$$L_y = zp_x - xp_z$$
$$L_z = xp_y - yp_x$$

Let us prove the first commutation relation. We have:

$$\begin{aligned}[L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] - [yp_z, xp_z] - [zp_y, zp_x] + [zp_y, xp_z]\end{aligned}$$

By heavily using the commutation relations from Proposition 6.8, we have:

$$[yp_z, zp_x] = yp_z zp_x - zp_x yp_z = y(zp_z - ih)p_x - zyp_x p_z = -ihyp_x$$
$$[yp_z, xp_z] = yp_z xp_z - xp_z yp_z = 0$$
$$[zp_y, zp_x] = zp_y zp_x - zp_x zp_y = 0$$
$$[zp_y, xp_z] = zp_y xp_z - xp_z zp_y = zxp_y p_z - x(zp_z - ih)p_y = ihxp_y$$

We conclude that the commutator that we were computing is given by the following formula, which is precisely the one in the statement:

$$\begin{aligned}[L_x, L_y] &= -ihyp_x + ihxp_y \\ &= ih(xp_y - yp_x) \\ &= ihL_z\end{aligned}$$

The proof of the other two commutation relations is similar, or can be simply obtained by invoking the cyclic invariance $x \to y \to z \to x$ of our problem, which cyclic invariance is not broken by the vector product $\times$ used, and so can indeed be invoked. □

As an interesting consequence of Theorem 6.9, we have:

PROPOSITION 6.10. *The following operator, called square of angular momentum*

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

commutes with all 3 *operators* L_x, L_y, L_z.

PROOF. We have the following computation, to start with:

$$\begin{aligned}
[L^2, L_x] &= (L_x^2 + L_y^2 + L_z^2)L_x - L_x(L_x^2 + L_y^2 + L_z^2) \\
&= L_y^2 L_x + L_z^2 L_x - L_x L_y^2 - L_x L_z^2 \\
&= [L_y^2, L_x] + [L_z^2, L_x]
\end{aligned}$$

The first commutator can be computed with a trick, as follows:

$$\begin{aligned}
[L_y^2, L_x] &= L_y L_y L_x - L_x L_y L_y \\
&= L_y L_y L_x - L_y L_x L_y + L_y L_x L_y - L_x L_y L_y \\
&= L_y [L_y, L_x] + [L_y, L_x] L_y \\
&= L_y(-ihL_z) + (-ihL_z)L_y \\
&= -ih(L_y L_z + L_z L_y)
\end{aligned}$$

The second commutator can be computed with the same trick, as follows:

$$\begin{aligned}
[L_z^2, L_x] &= L_z L_z L_x - L_x L_z L_z \\
&= L_z L_z L_x - L_z L_x L_z + L_z L_x L_z - L_x L_z L_z \\
&= L_z [L_z, L_x] + [L_z, L_x] L_z \\
&= L_z(ihL_y) + (ihL_y)L_z \\
&= ih(L_z L_y + L_y L_z)
\end{aligned}$$

Now by summing we obtain the following commutation relation, as desired:

$$[L^2, L_x] = 0$$

The proof of the other two commutation relations is similar, or we can simply invoke here the cyclic symmetry argument from the end of the proof of Theorem 6.9. ☐

Let us discuss now the diagonalization of the momentum operators L_x, L_y, L_z. Since these operators do not commute, we cannot hope for a joint diagonalization for them. Thus, we must choose one of them, and for reasons that will become clear later, when writing things in spherical coordinates, we will choose L_x.

In view of Proposition 6.10, this operator L_x does commute with L^2, and so we can hope for a joint diagonalization of L^2, L_x. And, so is what happens:

THEOREM 6.11. *The operators L^2, L_x diagonalize as*

$$L^2 f_l^m = h^2 l(l+1) f_l^m$$

$$L_x f_l^m = hm f_l^m$$

where $l \in \mathbb{N}/2$ and $m = -l, -l+1, \ldots, l-1, l$.

PROOF. This is something quite long, the idea being as follows:

(1) For reasons that will become clear later on, let us introduce two operators as follows, called raising and lowering operators:

$$L_+ = L_y + iL_z$$

$$L_- = L_y - iL_z$$

We will often deal with these operators at the same time, using the following notation:

$$L_\pm = L_y \pm iL_z$$

(2) We have the following computation:

$$\begin{aligned}
[L_x, L_\pm] &= [L_x, L_y] \pm i[L_x, L_z] \\
&= ihL_z \pm i(-ihL_y) \\
&= h(iL_z \pm L_y) \\
&= \pm h(\pm iL_z + L_y) \\
&= \pm hL_\pm
\end{aligned}$$

(3) Our claim now is that $L^2 f = \lambda f$, $L_x f = \mu f$ imply:

$$L^2(L_\pm f) = \lambda(L_\pm f)$$

$$L_x(L_\pm f) = (\mu \pm h)(L_\pm f)$$

Indeed, the first formula follows from:

$$\begin{aligned}
L^2(L_\pm f) &= L_\pm(L^2 f) \\
&= L_\pm(\lambda f) \\
&= \lambda(L_\pm f)
\end{aligned}$$

As for the second formula, this follows from:

$$\begin{aligned}
L_x(L_\pm f) &= L_x L_\pm f \\
&= (L_x L_\pm - L_\pm L_x)f + L_\pm L_x f \\
&= \pm hL_\pm f + L_\pm(\mu f) \\
&= (\mu \pm h)(L_\pm f)
\end{aligned}$$

(4) Now in view of the formulae found in (3), the raising and lowering operators act on the joint eigenfunctions of L^2, L_x, by leaving the L^2 eigenvalue unchanged, and by raising and lowering the eigenvalue of L_x. But both this raising process and lowering process for

the eigenvalue of L_x cannot go on forever, because of the following estimate:

$$\begin{aligned}\lambda &= < L^2 > \\ &= < L_x^2 > + < L_y^2 > + < L_z^2 > \\ &= \mu^2 + < L_y^2 > + < L_z^2 > \\ &\geq \mu^2\end{aligned}$$

(5) In order to see exactly how the raising and lowering processes terminate, we will need some more computations. We have:

$$\begin{aligned}L_\pm L_\mp &= (L_y \pm iL_z)(L_y \mp iL_z) \\ &= L_y^2 + L_z^2 \mp i(L_yL_z - L_zL_y) \\ &= L_y^2 + L_z^2 \mp i(ihL_x) \\ &= L_y^2 + L_z^2 \pm hL_x \\ &= L^2 - L_x^2 \pm hL_x\end{aligned}$$

Thus, we have the following formula:

$$L^2 = L_\pm L_\mp + L_x^2 \mp hL_x$$

Now assuming $L_x f = hlf$, at termination of the raising process, we have:

$$\begin{aligned}L^2(f) &= (L_-L_+ + L_x^2 + hL_x)f \\ &= (0 + h^2l^2 + h^2l)f \\ &= h^2l(l+1)f\end{aligned}$$

Similarly, assuming $L_x f = hl'f$, at termination of the lowering process, we have:

$$\begin{aligned}L^2(f) &= (L_+ - L_- + L_x^2 - hL_x)f \\ &= (0 + h^2l'^2 - h^2l')f \\ &= h^2l'(l'-1)f\end{aligned}$$

We conclude from this that we have the following formula:

$$l(l+1) = l'(l'-1)$$

But since $l' = l + 1$ is impossible, due to raising vs lowering, we must have $l' = -l$, and this leads to the conclusion in the statement.

(6) Finally, for being complete, the full and conceptual understanding of all the above imperatively requires a certain cat climbing a certain ladder, and for full details here, and for other things missing from the above proof, we refer to Griffiths [**43**]. □

Moving ahead now, let us write everything in spherical coordinates, and find the eigenfunctions. We have here the following remarkable result:

THEOREM 6.12. *In spherical coordinates r, s, t we have*

$$L_x = -\frac{ih}{dt}$$

$$L_y = ih\left(\frac{\sin t}{ds} + \frac{\cos s\cos t}{\sin s}\cdot\frac{1}{dt}\right)$$

$$L_z = -ih\left(\frac{\cos t}{ds} - \frac{\cos s\sin t}{\sin s}\cdot\frac{1}{dt}\right)$$

and the spherical harmonics are joint eigenfunctions of L^2, L_x.

PROOF. We recall that, according to our usual, N-dimensional looking conventions, the spherical coordinates are as follows, with $r \in [0,\infty)$ being the radius, $s \in [0,\pi]$ being the polar angle, and $t \in [0, 2\pi]$ being the azimuthal angle:

$$\begin{cases} x &= r\cos s \\ y &= r\sin s\cos t \\ z &= r\sin s\sin t \end{cases}$$

(1) We know that we have $L = -ihx \times \nabla$, so let us first compute ∇ in spherical coordinates. We have here, according to the chain rule for derivatives:

$$\begin{aligned}\nabla &= \begin{pmatrix} dr/dx & ds/dx & dt/dx \\ dr/dy & ds/dy & dt/dy \\ dr/dz & ds/dz & dt/dz \end{pmatrix}\begin{pmatrix} d/dr \\ d/ds \\ d/dt \end{pmatrix} \\ &= \begin{pmatrix} dx/dr & dy/dr & dz/dr \\ dx/ds & dy/ds & dz/ds \\ dx/dt & dy/dt & dz/dt \end{pmatrix}^{-1}\begin{pmatrix} d/dr \\ d/ds \\ d/dt \end{pmatrix}\end{aligned}$$

(2) On the other hand, it is routine to check that we have:

$$\begin{pmatrix} dx/dr & dx/ds & dx/dt \\ dy/dr & dy/ds & dy/dt \\ dz/dr & dz/ds & dz/dt \end{pmatrix} = \begin{pmatrix} \cos s & -r\sin s & 0 \\ \sin s\cos t & r\cos s\cos t & -r\sin s\sin t \\ \sin s\sin t & r\cos s\sin t & r\sin s\cos t \end{pmatrix}$$

It is also routine to see that this latter matrix, say A, satisfies:

$$A^tA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2 s \end{pmatrix}$$

Now if we call D the diagonal matrix on the right, we conclude that the matrix, say B, appearing in the above formula of ∇ is given by:

$$\begin{aligned}
B &= (A^t)^{-1} \\
&= AD^{-1} \\
&= \begin{pmatrix} \cos s & -r\sin s & 0 \\ \sin s\cos t & r\cos s\cos t & -r\sin s\sin t \\ \sin s\sin t & r\cos s\sin t & r\sin s\cos t \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & 1/(r^2\sin^2 s) \end{pmatrix} \\
&= \begin{pmatrix} \cos s & -\sin s/r & 0 \\ \sin s\cos t & \cos s\cos t/r & -\sin t/(r\sin s) \\ \sin s\sin t & \cos s\sin t/r & \cos t/(r\sin s) \end{pmatrix}
\end{aligned}$$

(3) Thus, the angular momentum operator that we are looking for, $L = -ihx \times \nabla$, written more conveniently as $L = -ihx/r \times r\nabla$, is given by:

$$L = -ih \begin{pmatrix} \cos s \\ \sin s\cos t \\ \sin s\sin t \end{pmatrix} \times \begin{pmatrix} r\cos s & -\sin s & 0 \\ r\sin s\cos t & \cos s\cos t & -\sin t/\sin s \\ r\sin s\sin t & \cos s\sin t & \cos t/\sin s \end{pmatrix} \begin{pmatrix} d/dr \\ d/ds \\ d/dt \end{pmatrix}$$

And computing now the vector product gives the formula for L in the statement.

(4) Now with our explicit formula for L in hand, we next find that the raising and lowering operators are given by:

$$L_\pm = \pm he^{\pm it}\left(\frac{d}{ds} \pm i\,\frac{\cos s}{\sin s} \cdot \frac{1}{dt}\right)$$

Next, we find that these two operators satisfy the following formula:

$$L_+L_- = -h^2\left(\frac{d^2}{ds^2} + \frac{\cos s}{\sin s} \cdot \frac{d}{ds} + \frac{\cos^2 s}{\sin^2 s} \cdot \frac{d^2}{dt^2} + i\frac{d}{dt}\right)$$

And finally, by using this latter formula, we find that L^2 is given by:

$$L^2 = -h^2\left(\frac{1}{\sin s} \cdot \frac{d}{ds}\left(\sin s \cdot \frac{d}{ds}\right) + \frac{1}{\sin^2 s} \cdot \frac{d^2}{dt^2}\right)$$

(5) With all these formulae in hand, we can now finish. The eigenfunction equation for the above operator L^2, with eigenvalue $h^2 l(l+1)$, is as follows:

$$-h^2\left(\frac{1}{\sin s} \cdot \frac{d}{ds}\left(\sin s \cdot \frac{d}{ds}\right) + \frac{1}{\sin^2 s} \cdot \frac{d^2}{dt^2}\right) f = h^2 l(l+1) f$$

But this is precisely the angular equation for the hydrogen atom. As for the eigenfunction equation for the operator L_x, with eigenvalue hm, this is as follows:

$$-\frac{ih}{dt} f = hmf$$

But this latter equation is equivalent to the azimuthal equation, also for the hydrogen atom. Thus, we are dealing here with equations that we already know, and the solutions are the spherical harmonics for the hydrogen atom, as claimed. □

6c. Pauli matrices

In order to talk now about spin, we will regard, a bit as in the classical mechanics case, the spin and the angular momentum as being similar quantities. Thus, in analogy with the basic equations for angular momentum, we should have:

Definition 6.13. *The components of the spin operator are subject to*

$$[S_x, S_y] = ihS_z$$

$$[S_y, S_z] = ihS_x$$

$$[S_z, S_x] = ihS_y$$

called commutation relations for the spin operator.

The point now is that, with the above relations in hand, which are identical to the commutation relations for the angular momentum, all the general results from the previous section, based on that commutation relations, extend to our present setting, simply by changing L into S everywhere. And in particular, we are led in this way to:

Theorem 6.14. *We have the following diagonalization formulae*

$$S^2 f_s^m = h^2 s(s+1) f_s^m$$

$$S_x f_s^m = hm f_s^m$$

$$S_\pm f_s^m = h\sqrt{s(s+1) - m(m\pm 1)}\, f_s^{m\pm 1}$$

involving the operators $S^2 = S_x^2 + S_y^2 + S_z^2$, S_x and $S_\pm = S_y \pm iS_z$.

Proof. Here the first two formulae are something that we already know, from the previous section, with L, j being replaced by S, s. As for the last formula, this is something that we did not need, in the L, j context, but that we will need now. We want to compute the constants $C_{s,\pm}^m$ making work the raising and lowering formula, namely:

$$S_\pm f_s^m = C_{s,\pm}^m f_s^{m\pm 1}$$

But this can be done by using $S^2 = S_\pm S_\mp + S_x^2 \mp hS_x$ and $S_\pm^* = S_\mp$, and we get:

$$C_{s,+}^m = h\sqrt{s(s+1) - m(m+1)}$$

$$C_{s,-}^m = h\sqrt{s(s+1) - m(m-1)}$$

Thus, we are led to the last formula in the statement, and we are done. □

In practice now, let us look for the simplest mathematical realization of spin. We know from the Stern-Gerlach experiment that the spin is something binary, that can be either up, or down. Thus, we are led, for fixed particles, to a quantum mechanics over $H = \mathbb{C}^2$, with spin up and down being represented by the following two vectors:

$$e_1 = \begin{pmatrix}1\\0\end{pmatrix} \quad , \quad e_2 = \begin{pmatrix}0\\1\end{pmatrix}$$

It remains now to see how the equations in Theorem 6.14 reformulate, in this $H = \mathbb{C}^2$ setting. But here, not many choices, and we are led to:

Definition 6.15. *In the quantum mechanics of the spin, over $H = \mathbb{C}^2$, with*

$$e_1 = \begin{pmatrix}1\\0\end{pmatrix} \quad , \quad e_2 = \begin{pmatrix}0\\1\end{pmatrix}$$

being spin up and down, the spin is subject to the following equations, for $f = e_1, e_2$,

$$S^2 f = h^2 s(s+1) f$$

$$S_x f = h m_f f$$

$$S_\pm f = h\sqrt{s(s+1) - m_f(m_f \pm 1)}\, \check{f}$$

with parameters $s = 1/2$, $m_{e_1} = 1/2$, $m_{e_2} = -1/2$, and with $\{e_1, e_2\} = \{f, \check{f}\}$.

Here all the choices, and notably $s = 1/2$, are very natural in view of Theorem 6.14, because these are the choices providing a "minimal" realization of the equations in Theorem 6.14, in the smallest possible number of dimensions, namely $N = 2$. However, all this comes with a shade of mystery, or at least is not rock-solid enough as to be called theorem, and it is probably safer to use the term "definition", as we did above.

The point now is that the above questions can be solved, the result being:

Theorem 6.16. *In the above $H = \mathbb{C}^2$ context, of the mechanics of a single, fixed electron, the components of the normalized spin $\sigma = 2S/h$ are as follows,*

$$\sigma_x = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \quad , \quad \sigma_y = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix} \quad , \quad \sigma_z = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix}$$

called Pauli matrices. In the general, dynamic context, where we already have a Hilbert space H for the wave function, spin can be introduced by using the space

$$H' = H \otimes \mathbb{C}^2$$

and using the above Pauli matrices for it, acting on the $\mathbb{C}^2$ part.

PROOF. The equations in Definition 6.15, written in full detail, are as follows:

$$S^2\begin{pmatrix}1\\0\end{pmatrix}=\frac{3h^2}{4}\begin{pmatrix}1\\0\end{pmatrix}\quad,\quad S^2\begin{pmatrix}0\\1\end{pmatrix}=\frac{3h^2}{4}\begin{pmatrix}0\\1\end{pmatrix}$$

$$S_x\begin{pmatrix}1\\0\end{pmatrix}=\frac{h}{2}\begin{pmatrix}1\\0\end{pmatrix}\quad,\quad S_x\begin{pmatrix}0\\1\end{pmatrix}=-\frac{h}{2}\begin{pmatrix}0\\1\end{pmatrix}$$

$$S_+\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\quad,\quad S_+\begin{pmatrix}0\\1\end{pmatrix}=h\begin{pmatrix}1\\0\end{pmatrix}$$

$$S_-\begin{pmatrix}1\\0\end{pmatrix}=h\begin{pmatrix}0\\1\end{pmatrix}\quad,\quad S_-\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

Thus, we have the following formulae, for the various matrices involved:

$$S^2=\frac{3h^2}{4}\begin{pmatrix}1&0\\0&1\end{pmatrix}\quad,\quad S_x=\frac{h}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}$$

$$S_+=h\begin{pmatrix}0&1\\0&0\end{pmatrix}\quad,\quad S_-=h\begin{pmatrix}0&0\\1&0\end{pmatrix}$$

In relation with what we want to prove, we have obtained the formula of S_x. Regarding now the formulae of S_y,S_z, these follow by solving the following system:

$$S_+=S_y+iS_z\quad,\quad S_-=S_y-iS_z$$

To be more precise, the computation for S_y goes as follows:

$$S_y=\frac{S_++S_-}{2}=\frac{h}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$$

As for the computation for S_z, this goes as follows:

$$S_z=\frac{S_+-S_-}{2i}=\frac{h}{2i}\begin{pmatrix}0&1\\-1&0\end{pmatrix}=\frac{h}{2}\begin{pmatrix}0&-i\\i&0\end{pmatrix}$$

Thus, we are led to the conclusions in the statement. □

As a first consequence of the above, coming from the formula $s=1/2$ in Definition 6.15, we can say, a bit formally, that electrons have spin 1/2. We will heavily use this fact in what follows, for all sorts of purposes. And we will talk about spin $\neq 1/2$ too, later, with a general particle discussion, involving bosons and fermions.

6d. Bloch sphere

Bloch sphere.

6e. Exercises

Exercises.

CHAPTER 7

Quantum gates

CHAPTER 8

Entanglement

Part III

Quantum computers

Blows your mind drastically
Fantastically
Blows your mind drastically
Fantastically

CHAPTER 9

Early attempts

9a.

9b.

9c.

9d.

9e. Exercises

CHAPTER 10

Laser, cooling

10a.

10b.

10c.

10d.

10e. Exercises

CHAPTER 11

Bugs and fixes

11a.

11b.

11c.

11d.

11e. Exercises

CHAPTER 12

Modern devices

12a.

12b.

12c.

12d.

12e. Exercises

Part IV

Coding, bugs

Ooh Daytona
Shine your light on me
Ooh Daytona
Shine your light on me

CHAPTER 13

Information theory

13a.

13b.

13c.

13d.

13e. Exercises

CHAPTER 14

Quantum information

14a.

14b.

14c.

14d.

14e. Exercises

CHAPTER 15

Random information

15a.

15b.

15c.

15d.

15e. Exercises

CHAPTER 16

Quantum algebra

16a.

16b.

16c.

16d.

16e. Exercises

Bibliography

[1] M.F. Atiyah, The geometry and physics of knots, Cambridge Univ. Press (1990).

[2] G. Aubrun and S.J. Szarek, Alice and Bob meet Banach: the interface of asymptotic geometric analysis and quantum information theory, AMS (2017).

[3] T. Banica, Graphs and their symmetries (2023).

[4] T. Banica, Introduction to modern physics (2023).

[5] T. Banica, Methods of electrodynamics (2023).

[6] R.J. Baxter, Exactly solved models in statistical mechanics, Academic Press (1982).

[7] I. Bengtsson and K. Życzkowski, Geometry of quantum states, Cambridge Univ. Press (2006).

[8] H. Bercovici and V. Pata, Stable laws and domains of attraction in free probability theory, *Ann. of Math.* **149** (1999), 1023–1060.

[9] B. Bollobás, Modern graph theory, Springer (1998).

[10] B. Bollobás, Random graphs, Cambridge Univ. Press (1985).

[11] B. Bollobás, Extremal graph theory, Dover (1978).

[12] A.H. Chamseddine and A. Connes, The spectral action principle, *Comm. Math. Phys.* **186** (1997), 731–750.

[13] A.H. Chamseddine and A. Connes, Why the standard model, *J. Geom. Phys.* **58** (2008), 38–47.

[14] V. Chari and A. Pressley, A guide to quantum groups, Cambridge Univ. Press (1994).

[15] B. Collins and I. Nechita, Random quantum channels I: graphical calculus and the Bell state phenomenon, *Comm. Math. Phys.* **297** (2010), 345–370.

[16] B. Collins and I. Nechita, Random quantum channels II: entanglement of random subspaces, Rényi entropy estimates and additivity problems, *Adv. Math.* **226** (2011), 1181–1201.

[17] B. Collins and I. Nechita, Gaussianization and eigenvalue statistics for random quantum channels (III), *Ann. Appl. Probab.* **21** (2011), 1136–1179.

[18] B. Collins and P. Śniady, Integration with respect to the Haar measure on unitary, orthogonal and symplectic groups, *Comm. Math. Phys.* **264** (2006), 773–795.

[19] A. Connes, Noncommutative geometry, Academic Press (1994).

[20] A. Connes and M. Marcolli, Noncommutative geometry, quantum fields and motives, AMS (2008).

[21] W.N. Cottingham and D.A. Greenwood, An introduction to the standard model of particle physics, Cambridge Univ. Press (2012).

[22] W. de Launey and D. Flannery, Algebraic design theory, AMS (2011).

[23] P. Di Francesco, Meander determinants, *Comm. Math. Phys.* **191** (1998), 543–583.

[24] P. Di Francesco, P. Mathieu and D. Sénéchal, Conformal field theory, Springer (1996).

[25] P.A.M. Dirac, Principles of quantum mechanics, Oxford Univ. Press (1930).

[26] S.T. Dougherty, Combinatorics and finite geometry, Springer (2020).

[27] M. Dresher, The mathematics of games of strategy, Dover (1981).

[28] R. Durrett, Probability: theory and examples, Cambridge Univ. Press (1990).

[29] A. Einstein, Relativity: the special and the general theory, Dover (1916).

[30] L.C. Evans, Partial differential equations, AMS (1998).

[31] W. Feller, An introduction to probability theory and its applications, Wiley (1950).

[32] E. Fermi, Thermodynamics, Dover (1937).

[33] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics I: mainly mechanics, radiation and heat, Caltech (1963).

[34] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics II: mainly electromagnetism and matter, Caltech (1964).

[35] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics III: quantum mechanics, Caltech (1966).

[36] R.P. Feynman and A.R. Hibbs, Quantum mechanics and path integrals, Dover (1965).

[37] P. Flajolet and R. Sedgewick, Analytic combinatorics, Cambridge Univ. Press (2009).

[38] M.H. Freedman, A. Kitaev and Z. Wang, Simulation of topological field theories by quantum computers, *Comm. Math. Phys.* **227** (2002), 587–603.

[39] M.H. Freedman, M. Larsen and Z. Wang, A modular functor which is universal for quantum computation, *Comm. Math. Phys.* **227** (2022), 605–622.

[40] M.H. Freedman, A. Kitaev, M. Larsen and Z. Wang, Topological quantum computation, *Bull. Amer. Math. Soc.* **40** (2003), 31–38.

[41] C. Godsil and G. Royle, Algebraic graph theory, Springer (2001).

[42] D.J. Griffiths, Introduction to electrodynamics, Cambridge Univ. Press (2017).

[43] D.J. Griffiths and D.F. Schroeter, Introduction to quantum mechanics, Cambridge Univ. Press (2018).

[44] D.J. Griffiths, Introduction to elementary particles, Wiley (2020).

[45] D.J. Griffiths, Revolutions in twentieth-century physics, Cambridge Univ. Press (2012).

[46] L.K. Grover, A fast quantum mechanical algorithm for database search, in *STOC '96 Philadelphia* (1996), 212–219.

[47] K. Huang, Introduction to statistical physics, CRC Press (2001).

[48] K. Huang, Quantum field theory, Wiley (1998).

[49] K. Huang, Quarks, leptons and gauge fields, World Scientific (1982).

[50] K. Huang, Fundamental forces of nature, World Scientific (2007).

[51] C. Itzykson and J.B. Zuber, Quantum field theory, Dover (1980).

[52] V.F.R. Jones, Index for subfactors, *Invent. Math.* **72** (1983), 1–25.

[53] V.F.R. Jones, On knot invariants related to some statistical mechanical models, *Pacific J. Math.* **137** (1989), 311–334.

[54] V.F.R. Jones, Planar algebras I, preprint (1999).

[55] L.P. Kadanoff, Statistical physics: statics, dynamics and renormalization, World Scientific (2000).

[56] A. Kitaev, Fault-tolerant quantum computation by anyons, *Annals of Physics* **303** (1997), 2–30.

[57] D.E. Knuth, The art of computer programming, Addison-Wesley (1968).

[58] M. Kumar, Quantum: Einstein, Bohr, and the great debate about the nature of reality, Norton (2009).

[59] T. Lancaster and K.M. Blundell, Quantum field theory for the gifted amateur, Oxford Univ. Press (2014).

[60] L.D. Landau and E.M. Lifshitz, Mechanics, Pergamon Press (1960).

[61] L.D. Landau and E.M. Lifshitz, The classical theory of fields, Addison-Wesley (1951).

[62] L.D. Landau and E.M. Lifshitz, Quantum mechanics: non-relativistic theory, Pergamon Press (1959).

[63] V.B. Berestetskii, E.M. Lifshitz and L.P. Pitaevskii, Quantum electrodynamics, Butterworth-Heinemann (1982).

[64] S. Ling and C. Xing, Coding theory: a first course, Cambridge Univ. Press (2004).

[65] M. Lupini, L. Mančinska and D.E. Roberson, Nonlocal games and quantum permutation groups, *J. Funct. Anal.* **279** (2020), 1–39.

[66] V.A. Marchenko and L.A. Pastur, Distribution of eigenvalues in certain sets of random matrices, *Mat. Sb.* **72** (1967), 507–536.

[67] M. Marcolli, Noncommutative cosmology, World Scientific (2018).

[68] M.L. Mehta, Random matrices, Elsevier (2004).

[69] M.A. Nielsen and I.L. Chuang, Quantum computation and quantum information, Cambridge Univ. Press (2000).

[70] A. Peres, Quantum theory: concepts and methods, Kluwer (1993).

[71] M. Peskin and D.V. Schroeder, An introduction to quantum field theory, CRC press (1995).

[72] J. Preskill, Quantum information and computation, Caltech (1998).

[73] D.E. Roberson and S. Schmidt, Quantum symmetry vs nonlocal symmetry (2020).

[74] J.J. Sakurai and J. Napolitano, Modern quantum mechanics, Cambridge Univ. Press (1985).

[75] J. Schwinger, Einstein's legacy: the unity of space and time, Dover (1986).

[76] J. Schwinger, L.L. DeRaad Jr., K.A. Milton and W.Y. Tsai, Classical electrodynamics, CRC Press (1998).

[77] J. Schwinger and B.H. Englert, Quantum mechanics: symbolism of atomic measurements, Springer (2001).

[78] C.E. Shannon, A mathematical theory of communication, *Bell System Tech. J.* **27** (1948), 379–423.

[79] P.W. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *IEEE Comput. Soc. Press* (1994), 124–134.

[80] N.J.A. Sloane and S. Plouffe, Encyclopedia of integer sequences, Academic Press (1995).

[81] D.R. Stinson, Combinatorial designs: constructions and analysis, Springer-Verlag (2006).

[82] N.H. Temperley and E.H. Lieb, Relations between the "percolation" and "colouring" problem and other graph-theoretical problems associated with regular planar lattices: some exact results for the "percolation" problem, *Proc. Roy. Soc. London* **322** (1971), 251–280.

[83] J.S. Townsend, A modern approach to quantum mechanics, Univ. Science Books (1992).

[84] D.V. Voiculescu, K.J. Dykema and A. Nica, Free random variables, AMS (1992).

[85] J. von Neumann, Mathematical foundations of quantum mechanics, Princeton Univ. Press (1955).

[86] J. von Neumann and O. Morgenstern, Theory of games and economic behavior, Princeton Univ. Press (1944).

[87] S. Wang, Quantum symmetry groups of finite spaces, *Comm. Math. Phys.* **195** (1998), 195–211.

[88] J. Watrous, The theory of quantum information, Cambridge Univ. Press (2018).

[89] S. Weinberg, Foundations of modern physics, Cambridge Univ. Press (2011).

[90] S. Weinberg, Lectures on quantum mechanics, Cambridge Univ. Press (2012).

[91] S. Weinberg, Lectures on astrophysics, Cambridge Univ. Press (2019).

[92] S. Weinberg, Cosmology, Oxford Univ. Press (2008).

[93] D. Weingarten, Asymptotic behavior of group integrals in the limit of infinite rank, *J. Math. Phys.* **19** (1978), 999–1001.

[94] H. Weyl, The theory of groups and quantum mechanics, Princeton Univ. Press (1931).

[95] H. Weyl, The classical groups: their invariants and representations, Princeton Univ. Press (1939).

[96] H. Weyl, Space, time, matter, Princeton Univ. Press (1918).

[97] E. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. of Math.* **62** (1955), 548–564.

[98] E. Witten, Quantum field theory and the Jones polynomial, *Comm. Math. Phys.* **121** (1989), 351–399.

[99] S.L. Woronowicz, Compact matrix pseudogroups, *Comm. Math. Phys.* **111** (1987), 613–665.

[100] B. Zwiebach, A first course in string theory, Cambridge Univ. Press (2004).