Invitation to lattice models

Teo Banica

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

2010 *Mathematics Subject Classification.* 60B20

Key words and phrases. Lattice model, Quantum group

Abstract. A lattice model is a discretization of a usual, continuous problem taking place in $\mathbb{R}^3$, over the rescaled integer lattice $\varepsilon\mathbb{Z}^3 \subset \mathbb{R}^3$, with $\varepsilon \simeq 0$. Examples of such models abound, with the general idea of discretization being the main idea of statistical mechanics, and with a main example being the Ising model for ferromagnetism. We discuss here such questions, with an introduction to the subject, and with an introduction to more advanced aspects too, using methods from probability and quantum algebra.

Preface

A lattice model is a discretization of a usual, continuous problem taking place in $\mathbb{R}^3$, over the lattice $\varepsilon\mathbb{Z}^3 \subset \mathbb{R}^3$, with $\varepsilon \simeq 0$. That is, our modeling method consists in replacing the space $\mathbb{R}^3$ with the discrete lattice $\varepsilon\mathbb{Z}^3 \subset \mathbb{R}^3$, chosen to be reasonably fine, $\varepsilon \simeq 0$, then reformulating our question in terms of this lattice $\varepsilon\mathbb{Z}^3$, in the hope that the mathematics will get simpler, and that with $\varepsilon \to 0$ we will get to interesting results.

All this is extremely vague and general, and actually reminds computers, and you would probably say "yes, sounds like a lazy scientist who wants the computer to do all the work". Which is a good point, but it is not about numeric methods that we want to talk about in this book. Our claim is that discretization can be a powerful method even in the context of pure mathematics and physics, with no computers allowed.

Here are a few examples for all this, and judge for yourself:

(1) Waves. Not that the wave equation $\ddot{\varphi} = v^2\Delta\varphi$ really needs help from discretization, but one remarkable thing is that this wave equation can be deduced by approximating $\mathbb{R}^3$ with a lattice $\varepsilon\mathbb{Z}^3 \subset \mathbb{R}^3$, with $\varepsilon \to 0$, consisting of balls connected by springs. So, forgetting now about the wave equation itself, we can declare that our universe is simply something made of tiny little balls, connected by tiny little springs.

(2) Along the same lines, by making use of some imagination, we can discretize many other things, with every time interesting conclusions. Want to talk about the propagation of heat, $\dot{\varphi} = \alpha\Delta\varphi$? There surely should be a discrete model for that. What about fluid mechanics, and the various equations there? Just discretize the fluids, as being formed by tiny little spheres or cubes, and that should be your model. And so on.

(3) In fact, we can even discretize gravity, by saying that if we place a mass $M > 0$ at the origin, the "news" about it will spread quickly among the points of $\varepsilon\mathbb{Z}^3 \subset \mathbb{R}^3$, from neighbor to neighbor, with everyone saying around something of type "hey buddies, there's a mass $M > 0$ somewhere towards the origin, please comply and get attracted by it, and pass the word to your neighbors". And so, in the end, gravity works.

(4) And with the remark that, with this latter example, we might be well now into cutting edge physics. Indeed, one of the main findings of modern physics is that there is

a "Planck scale" $\varepsilon > 0$ where the discretization of physics might actually be something exact, and understanding gravity at the Planck scale is a big open problem. Which would put an end to the "physics of the void" that we are so used to, for centuries.

(5) Less speculatively now, examples of discretization and lattice models abound in statistical mechanics. As a basic example here, lattices $\varepsilon\mathbb{Z}^3 \subset \mathbb{R}^3$ are naturally met in the context of materials like metals, where electrons can be thought of as being arranged on such a lattice. And, with this picture in mind, a main example of lattice model, having countless concrete applications, is the Ising model for ferromagnetism.

(6) And for ending with more speculation, quantum mechanics in its advanced form used nowadays, which is the quantum electrodynamics (QED) developed by Feynman and others, is in fact something of a lattice model too, with a particle, in the QED picture, hesitating a billion times per second on where to go, left, right, up, down and so on, a bit like being totally lost inside a certain lattice $\varepsilon\mathbb{Z}^3 \subset \mathbb{R}^3$, with ε being very small.

So, this was for the general idea of lattice models, and as you can see, this idea is as wide as physics itself. And by throwing now into the picture computers and numeric methods, that we have not talked about in the above, we are led to the conclusion that, in practice, lattice models might well account for more than half of physics.

The present book is an introduction to this. Our goal with be that of discussing lattice models, with as many interesting examples as possible, and by being reasonably strong on mathematical aspects, with a preference for algebraic and probabilistic methods.

We will mostly insist on (5), statistical mechanics and the Ising model, which are the topics where there are many things that can be said, and by using for our study quantum groups, along the lines of the Yang-Baxter equation, and of the work of Vaughan Jones. We will also talk a bit about more speculative aspects, such as (1,2,3,4,6).

This book uses some quantum group mathematics that I developed since long, and I would like to thank my coworkers. Thanks as well to my cats, for some help with computations. The mouse discretization result at the end of chapter 16 is due to them.

Cergy, February 2025
Teo Banica

Contents

Part I

Lattice models

By the rivers of Babylon
There we sat down
Yeah, we wept
When we remembered Zion

CHAPTER 1

Wave equation

1a. Laplace operator

Welcome to the waves. The idea for this Part I of the present book will be to have a tour through basic physics, guided by the waves, and by our discretization and lattice model intentions. This might sound perhaps a bit original, but thinking well, each major branch of physics is guided by its own "wave equation", so following the waves looks like a reasonable way of planning our journey, and learning some physics.

In practice now, and coming a bit in advance, the truth about waves is as follows:

FACT 1.1. *Waves can be of many types, and basically fall into two classes:*

(1) *Mechanical waves, such as the usual water waves, but also the sound waves, or the seismic waves. In all these cases, the wave propagates mechanically, via a certain medium, which can be solid, liquid or gaseous.*
(2) *Electromagnetic waves, coming via a more complicated mechanism, namely an accelerating charge in the context of electromagnetism. These are the radio waves, microwaves, IR, visible light, UV, X-rays and γ-rays.*

Quite remarkably, the behavior of all the above waves is basically described by the same wave equation, which looks as follows, and details on this later:

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

Understanding all this, what are the above waves and why they all satisfy this equation, will need us to have a tour through classical mechanics, elasticity, and solids, liquids and gases, in short states of matter, and then electrodynamics, and the Maxwell equations. Exciting enough you would say, and with the remark that in each of the above cases, the equation $\ddot{\varphi} = v^2 \Delta \varphi$ must be fine-tuned a bit. We will discuss this too.

Let us mention that gravity will not be forgotten, with a wave oriented discussion about it, in chapter 3. And the same goes for advanced thermodynamics, we will talk about this in chapter 4, by adopting the main equation there, called heat diffusion equation, and which looks a bit different, namely $\dot{\varphi} = \alpha \Delta \varphi$, as being a "heat wave" equation. And with this being a good deal, because we will end up in this way of talking about the Schrödinger equation too, as a close cousin of this heat equation.

Getting started for good now, some mathematics first. What is the Laplace operator Δ, appearing in the above? The answer here is very simple, coming from:

PRINCIPLE 1.2. *The second derivative of a function $\varphi : \mathbb{R}^N \to \mathbb{R}$, making the formula*

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h + \frac{<\varphi''(x)h, h>}{2}$$

work, is its Hessian matrix $\varphi''(x) \in M_N(\mathbb{R})$, given by the following formula:

$$\varphi''(x) = \left(\frac{d^2\varphi}{dx_i dx_j}\right)_{ij}$$

However, when needing a number, as second derivative, the trace of $\varphi''(x)$, denoted

$$\Delta\varphi = \sum_{i=1}^{N} \frac{d^2\varphi}{dx_i^2}$$

and called Laplacian of φ, usually does the job.

So, let us try to understand this principle. In one variable, things are quite simple, and you certainly know them well. Given a function $\varphi : \mathbb{R} \to \mathbb{R}$, the first job is that of finding a quantity $\varphi'(x) \in \mathbb{R}$ making the following formula work:

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h$$

But here, there are not so many choices, and the solution is that of defining the number $\varphi'(x) \in \mathbb{R}$ by the following formula, provided that the limit converges indeed:

$$\varphi'(x) = \lim_{h \to 0} \frac{\varphi(x+h) - \varphi(x)}{h}$$

This number is called derivative of φ at the point $x \in \mathbb{R}$, and as you surely know, geometrically, we have $\varphi'(x) = \tan\alpha$, with α being the slope of φ at the point x.

Still in one variable, we can talk as well about second derivatives, as follows:

THEOREM 1.3. *The second derivative of a function $\varphi : \mathbb{R} \to \mathbb{R}$, making the formula*

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h + \frac{\varphi''(x)h^2}{2}$$

work, is the derivative φ'' of the derivative $\varphi' : \mathbb{R} \to \mathbb{R}$.

PROOF. Assume indeed that φ is twice differentiable at x, and let us try to construct an approximation of φ around x by a quadratic function, as follows:

$$\varphi(x+h) \simeq a + bh + ch^2$$

We must have $a = \varphi(x)$, and we also know that $b = \varphi'(x)$ is the correct choice for the coefficient of h. Thus, our approximation must be as follows:

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h + ch^2$$

In order to find the correct choice for $c \in \mathbb{R}$, observe that the function $\psi(h) = \varphi(x+h)$ matches with $P(h) = \varphi(x) + \varphi'(x)h + ch^2$ in what regards the value at $h = 0$, and also in what regards the value of the derivative at $h = 0$. Thus, the correct choice of $c \in \mathbb{R}$ should be the one making match the second derivatives at $h = 0$, and this gives:

$$c = \frac{\varphi''(x)}{2}$$

Thus, we are led to the formula in the statement. In order to prove now this formula, we can use L'Hôpital's rule, which states that the $0/0$ type limits are given by:

$$\frac{f(x)}{g(x)} \simeq \frac{f'(x)}{g'(x)}$$

Indeed, if we denote by $\psi(h) \simeq P(h)$ the formula to be proved, we have:

$$\begin{aligned}
\frac{\psi(h) - P(h)}{h^2} &\simeq \frac{\psi'(h) - P'(h)}{2h} \\
&\simeq \frac{\psi''(h) - P''(h)}{2} \\
&= \frac{\varphi''(h) - \varphi''(h)}{2} \\
&= 0
\end{aligned}$$

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

Many other things can be said, as a continuation of the above, and you surely know all this. Before leaving the subject, however, let us record the following statement, which is something a bit heuristic, which will play an important role, in what follows:

PROPOSITION 1.4. *Intuitively speaking, the second derivative $\varphi''(x) \in \mathbb{R}$ computes how much different is $\varphi(x)$, compared to the average of $\varphi(y)$, with $y \simeq x$.*

PROOF. As already mentioned, this is something a bit heuristic, but which is good to know. Let us write the formula in Theorem 1.3, as such, and with $h \to -h$ too:

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h + \frac{\varphi''(x)}{2}\,h^2$$

$$\varphi(x-h) \simeq \varphi(x) - \varphi'(x)h + \frac{\varphi''(x)}{2}\,h^2$$

By making the average, we obtain the following formula:

$$\frac{\varphi(x+h) + \varphi(x-h)}{2} \simeq \varphi(x) + \frac{\varphi''(x)}{2}\,h^2$$

Thus, thinking a bit, we are led to the conclusion in the statement, modulo some discussion regarding what the average of $h^2/2$ exactly is. It is of course possible to say more here, but we will not really need all the details, in what follows. □

Moving now to several variables, $N \geq 2$, as a first job, given a function $\varphi : \mathbb{R}^N \to \mathbb{R}$, we would like to find a quantity $\varphi'(x)$ making the following formula work:

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h$$

But here, again there are not so many choices, and the solution is that of defining $\varphi'(x)$ as being the row vector formed by the partial derivatives at x:

$$\varphi'(x) = \begin{pmatrix} \frac{d\varphi}{dx_1} & \cdots & \frac{d\varphi}{dx_N} \end{pmatrix}$$

To be more precise, with this value for $\varphi'(x)$, our approximation formula $\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h$ makes sense indeed, as an equality of real numbers, with $\varphi'(x)h \in \mathbb{R}$ being obtained as the matrix multiplication of the row vector $\varphi'(x)$, and the column vector h. As for the fact that our formula holds indeed, this follows by putting together the approximation properties of each of the partial derivatives $d\varphi/dx_i$, which give:

$$\varphi(x+h) \simeq \varphi(x) + \sum_{i=1}^{N} \frac{d\varphi}{dx_i} \cdot h_i = \varphi(x) + \varphi'(x)h$$

Before moving forward, you might say, why bothering with horizontal vectors, when it is so simple and convenient to have all vectors vertical, by definition. Good point, and in answer, we can indeed talk about the gradient of φ, constructed as follows:

$$\nabla\varphi = \begin{pmatrix} \frac{d\varphi}{dx_1} \\ \vdots \\ \frac{d\varphi}{dx_N} \end{pmatrix}$$

With this convention, $\nabla\varphi$ geometrically describes the slope of φ at the point x, in the obvious way. However, the approximation formula must be rewritten as follows:

$$\varphi(x+h) \simeq \varphi(x) + < \nabla\varphi(x), h >$$

In what follows we will use both φ' and $\nabla\varphi$, depending on the context. Moving now to second derivatives, the result here, generalizing Theorem 1.3, is as follows:

THEOREM 1.5. *The second derivative of a function $\varphi : \mathbb{R}^N \to \mathbb{R}$, making the formula*

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h + \frac{< \varphi''(x)h, h >}{2}$$

work, is its Hessian matrix $\varphi''(x) \in M_N(\mathbb{R})$, given by the following formula:

$$\varphi''(x) = \left(\frac{d^2\varphi}{dx_i dx_j} \right)_{ij}$$

Moreover, this Hessian matrix is symmetric, $\varphi''(x)_{ij} = \varphi'(x)_{ji}$.

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

(1) As a first observation, at $N = 1$ the Hessian matrix constructed above is simply the 1×1 matrix having as entry the second derivative $\varphi''(x)$, and the formula in the statement is something that we know well from Theorem 1.3, namely:

$$\varphi(x + h) \simeq \varphi(x) + \varphi'(x)h + \frac{\varphi''(x)h^2}{2}$$

(2) At $N = 2$ now, we obviously need to differentiate φ twice, and the point is that we come in this way upon the following formula, called Clairaut formula:

$$\frac{d^2\varphi}{dxdy} = \frac{d^2\varphi}{dydx}$$

But, is this formula correct or not? As an intuitive justification for it, let us consider a product of power functions, $\varphi(z) = x^p y^q$. We have then our formula, due to:

$$\frac{d^2\varphi}{dxdy} = \frac{d}{dx}\left(\frac{dx^p y^q}{dy}\right) = \frac{d}{dx}\left(qx^p y^{q-1}\right) = pqx^{p-1}y^{q-1}$$

$$\frac{d^2\varphi}{dydx} = \frac{d}{dy}\left(\frac{dx^p y^q}{dx}\right) = \frac{d}{dy}\left(px^{p-1} y^q\right) = pqx^{p-1}y^{q-1}$$

Next, let us consider a linear combination of power functions, $\varphi(z) = \sum_{pq} c_{pq} x^p y^q$, which can be finite or not. We have then, by using the above computation:

$$\frac{d^2\varphi}{dxdy} = \frac{d^2\varphi}{dydx} = \sum_{pq} c_{pq} pq x^{p-1} y^{q-1}$$

Thus, we can see that our commutation formula for derivatives holds indeed, due to the fact that the functions in x, y commute. Of course, all this does not fully prove our formula, in general. But exercise for you, to have this idea fully working, or to look up the standard proof of the Clairaut formula, using the mean value theorem.

(3) Moving now to $N = 3$ and higher, we can use here the Clairaut formula with respect to any pair of coordinates, which gives the Schwarz formula, namely:

$$\frac{d^2\varphi}{dx_i dx_j} = \frac{d^2\varphi}{dx_j dx_i}$$

Thus, the second derivative, or Hessian matrix, is symmetric, as claimed.

(4) Getting now to the main topic, namely approximation formula in the statement, in arbitrary N dimensions, this is in fact something which does not need a new proof, because it follows from the one-variable formula in (1), applied to the restriction of φ to the following segment in $\mathbb{R}^N$, which can be regarded as being a one-variable interval:

$$I = [x, x + h]$$

To be more precise, let $y \in \mathbb{R}^N$, and consider the following function, with $r \in \mathbb{R}$:

$$f(r) = \varphi(x + ry)$$

We know from (1) that the Taylor formula for f, at the point $r = 0$, reads:

$$f(r) \simeq f(0) + f'(0)r + \frac{f''(0)r^2}{2}$$

And our claim is that, with $h = ry$, this is precisely the formula in the statement.

(5) So, let us see if our claim is correct. By using the chain rule, we have the following formula, with on the right, as usual, a row vector multiplied by a column vector:

$$f'(r) = \varphi'(x + ry) \cdot y$$

By using again the chain rule, we can compute the second derivative as well:

$$\begin{aligned} f''(r) &= (\varphi'(x + ry) \cdot y)' \\ &= \left(\sum_i \frac{d\varphi}{dx_i}(x + ry) \cdot y_i \right)' \\ &= \sum_i \sum_j \frac{d^2\varphi}{dx_i dx_j}(x + ry) \cdot \frac{d(x + ry)_j}{dr} \cdot y_i \\ &= \sum_i \sum_j \frac{d^2\varphi}{dx_i dx_j}(x + ry) \cdot y_i y_j \\ &= < \varphi''(x + ry)y, y > \end{aligned}$$

(6) Time now to conclude. We know that we have $f(r) = \varphi(x + ry)$, and according to our various computations above, we have the following formulae:

$$f(0) = \varphi(x) \quad , \quad f'(0) = \varphi'(x) \quad , \quad f''(0) =< \varphi''(x)y, y >$$

Buit with this data in hand, the usual Taylor formula for our one variable function f, at order 2, at the point $r = 0$, takes the following form, with $h = ry$:

$$\begin{aligned} \varphi(x + ry) &\simeq \varphi(x) + \varphi'(x)ry + \frac{< \varphi''(x)y, y > r^2}{2} \\ &= \varphi(x) + \varphi'(x)t + \frac{< \varphi''(x)h, h >}{2} \end{aligned}$$

Thus, we have obtained the formula in the statement. □

As before in the one variable case, many more things can be said, as a continuation of the above. For instance the local minima and maxima of $\varphi : \mathbb{R}^N \to \mathbb{R}$ appear at the points $x \in \mathbb{R}^N$ where the derivative vanishes, $\varphi'(x) = 0$, and where the second derivative $\varphi''(x) \in M_N(\mathbb{R})$ is positive, respectively negative. But, you surely know all this.

Getting back now to what we wanted to do, namely understand Principle 1.2, it remains to talk about the Laplace operator Δ. Things are quite tricky here, basically requiring some physics that we still need to develop, but as something mathematical to start with, we have the following higher dimensional analogue of Proposition 1.4:

PROPOSITION 1.6. *Intuitively, the following quantity, called Laplacian of φ,*

$$\Delta\varphi = \sum_{i=1}^{N} \frac{d^2\varphi}{dx_i^2}$$

measures how much different is $\varphi(x)$, compared to the average of $\varphi(y)$, with $y \simeq x$.

PROOF. As before with Proposition 1.4, this is something a bit heuristic, but good to know. Let us write the formula in Theorem 1.5, as such, and with $h \to -h$ too:

$$\varphi(x+h) \simeq \varphi(x) + \varphi'(x)h + \frac{<\varphi''(x)h, h>}{2}$$

$$\varphi(x-h) \simeq \varphi(x) - \varphi'(x)h + \frac{<\varphi''(x)h, h>}{2}$$

By making the average, we obtain the following formula:

$$\frac{\varphi(x+h) + \varphi(x-h)}{2} \simeq \varphi(x) + \frac{<\varphi''(x)h, h>}{2}$$

Thus, thinking a bit, we are led to the conclusion in the statement, modulo some discussion about integrating this, that we will not really need, in what follows. □

With this understood, the problem is now, what can we say about the mathematics of Δ? As a first observation, which is a bit speculative, the Laplace operator appears by applying twice the gradient operator, in a somewhat formal sense, as follows:

$$\begin{aligned}
\Delta\varphi &= \sum_{i=1}^{N} \frac{d^2\varphi}{dx_i^2} \\
&= \sum_{i=1}^{N} \frac{d}{dx_i} \cdot \frac{d\varphi}{dx_i} \\
&= \left\langle \begin{pmatrix} \frac{d}{dx_1} \\ \vdots \\ \frac{d}{dx_N} \end{pmatrix}, \begin{pmatrix} \frac{d\varphi}{dx_1} \\ \vdots \\ \frac{d\varphi}{dx_N} \end{pmatrix} \right\rangle \\
&= <\nabla, \nabla\varphi>
\end{aligned}$$

Thus, it is possible to write a formula of type $\Delta = \nabla^2$, with the convention that the square of the gradient ∇ is taken in a scalar product sense, as above. However, this can be a bit confusing, and in what follows, we will not use this notation.

Instead of further thinking at this, and at double derivatives in general, let us formulate a more straightforward question, inspired by linear algebra, as follows:

QUESTION 1.7. *The Laplace operator being linear,*

$$\Delta(a\varphi + b\psi) = a\Delta\varphi + b\Delta\psi$$

what can we say about it, inspired by usual linear algebra?

In answer now, the space of functions $\varphi : \mathbb{R}^N \to \mathbb{R}$, on which Δ acts, being infinite dimensional, the usual tools from linear algebra do not apply as such, and we must be extremely careful. For instance, we cannot really expect to diagonalize Δ, via some sort of explicit procedure, as we usually do in linear algebra, for the usual matrices.

Thinking some more, there is actually a real bug too with our problem, because at $N = 1$ this problem becomes "what can we say about the second derivatives $\varphi'' : \mathbb{R} \to \mathbb{R}$ of the functions $\varphi : \mathbb{R} \to \mathbb{R}$, inspired by linear algebra", with answer "not much".

And by thinking even more, still at $N = 1$, there is a second bug too, because if $\varphi : \mathbb{R} \to \mathbb{R}$ is twice differentiable, nothing will guarantee that its second derivative $\varphi'' : \mathbb{R} \to \mathbb{R}$ is twice differentiable too. Thus, we have some issues with the domain and range of Δ, regarded as linear operator, and these problems will persist at higher N.

So, shall we trash Question 1.7? Not so quick, because, very remarkably, some magic comes at $N = 2$ and higher in relation with complex analysis, according to:

PRINCIPLE 1.8. *The functions $\varphi : \mathbb{R}^N \to \mathbb{R}$ which are 0-eigenvectors of Δ,*

$$\Delta\varphi = 0$$

called harmonic functions, have the following properties:

(1) *At $N = 1$, nothing spectacular, these are just the linear functions.*
(2) *At $N = 2$, these are, locally, the real parts of holomorphic functions.*
(3) *At $N \geq 3$, these still share many properties with the holomorphic functions.*

In order to understand this, or at least get introduced to it, let us first look at the case $N = 2$. Here, any function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ can be regarded as function $\varphi : \mathbb{C} \to \mathbb{R}$, depending on $z = x + iy$. But, in view of this, it is natural to enlarge the attention to the functions $\varphi : \mathbb{C} \to \mathbb{C}$, and ask which of these functions are harmonic, $\Delta\varphi = 0$. And here, we have the following remarkable result, making the link with complex analysis:

THEOREM 1.9. *Any holomorphic function $\varphi : \mathbb{C} \to \mathbb{C}$, when regarded as function*

$$\varphi : \mathbb{R}^2 \to \mathbb{C}$$

is harmonic. Moreover, the conjugates $\bar{\varphi}$ of holomorphic functions are harmonic too.

PROOF. The first assertion comes from the following computation, with $z = x + iy$:

$$\begin{aligned}\Delta z^n &= \frac{d^2z^n}{dx^2} + \frac{d^2z^n}{dy^2}\\ &= \frac{d(nz^{n-1})}{dx} + \frac{d(inz^{n-1})}{dy}\\ &= n(n-1)z^{n-2} - n(n-1)z^{n-2}\\ &= 0\end{aligned}$$

As for the second assertion, this follows from $\Delta\bar{\varphi} = \overline{\Delta\varphi}$, which is clear from definitions, and which shows that if φ is harmonic, then so is its conjugate $\bar{\varphi}$. □

Many more things can be said, along these lines, notably a proof of the assertion (2) in Principle 1.8, which is however a quite tough piece of mathematics, and then with a clarification of the assertion (3) too, from that same principle, which again requires some substantial mathematics. We will be back to both these topics, in due time.

As a conclusion to all this, we have now a reasonable understanding of the second derivatives $\varphi''(x)$, and of the related quantities $\Delta\varphi(x)$, for the functions $\varphi : \mathbb{R}^N \to \mathbb{R}$, coming from Principle 1.2, which is something that we understand now well, and from Principle 1.8 too, which is something which still remains to be explored. And, good news, this is all we need, for the moment, in order to get into physics and waves.

1b. Wave equation

Getting now to physics, we first need to talk about the simplest mechanical waves. But, these simplest mechanical waves, you can produce them yourself, by playing with a chain, or cracking a whip. And with the mathematics here being pure classical mechanics, of standard type, we can eat raw our modeling problem, and we find:

THEOREM 1.10. *The wave equation in $\mathbb{R}^N$ is*

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

where $v > 0$ is the propagation speed.

PROOF. This is indeed something straightforward, coming from some routine classical mechanics computations for a real-life chain, and then letting the number of chains, by keeping the total length fixed, to go to infinity, $C \to \infty$. Alternatively, we can speculate as well directly on a continuous model, of whip type, again via classical mechanics. Finally, all this study can be performed in arbitrary $N \in \mathbb{N}$ dimensions, as indicated. □

We would like to present now another model for the mechanical waves, this time based on a "balls and springs" vision of the problem, and perhaps even of the whole universe, which is something very exciting, and will be our main model, afterwards. Here is the result, identical to our previous Theorem 1.10, but coming with a different proof:

THEOREM 1.11. *The wave equation in $\mathbb{R}^N$ is*

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

where $v > 0$ is the propagation speed.

PROOF. As already mentioned before, the equation in the statement is what comes out of experiments. However, allowing us a bit of imagination, and trust in this imagination, we can mathematically "prove" this equation, by discretizing, 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 the 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 $B >> 0$ 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:

$$m = \frac{M}{B} \quad , \quad k = KB \quad , \quad l = \frac{L}{B}$$

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

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

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

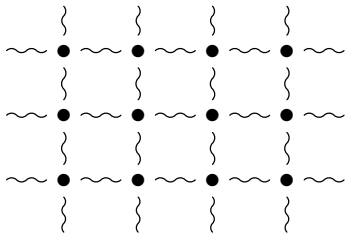
$$\ddot{\varphi}(x) = \frac{KL^2}{M} \cdot \frac{\varphi(x+l) - 2\varphi(x) + \varphi(x-l)}{l^2}$$

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

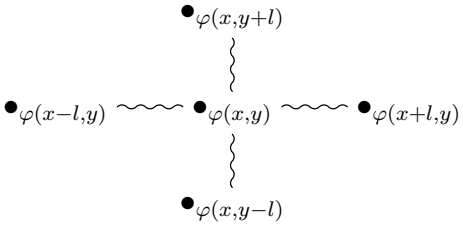
$$\ddot{\varphi}(x) = \frac{KL^2}{M} \cdot \frac{d^2\varphi}{dx^2}(x)$$

We are therefore led to the wave equation in the statement, which is $\ddot{\varphi} = v^2\varphi''$ in our present $N = 1$ dimensional case, the propagation speed being $v = \sqrt{K/M} \cdot L$.

(3) In 2 dimensions now, the same argument carries on. Indeed, we can use here a lattice model as follows, with all the edges standing for small springs:



As before in one dimension, we send an impulse, and we zoom on one ball. The situation here is as follows, with l being the spring length:



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

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

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

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

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

$$\begin{aligned} m\cdot\ddot{\varphi}(x,y) &= k(\varphi(x+l,y)-2\varphi(x,y)+\varphi(x-l,y)) \\ &+ k(\varphi(x,y+l)-2\varphi(x,y)+\varphi(x,y-l)) \end{aligned}$$

(4) 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 $B^2>>0$ balls, having a total mass M, and spanning a total area L^2. Thus, our previous infinitesimal parameters are as follows, with K being the spring constant of the total system, taken to be equal to k:

$$m=\frac{M}{B^2}\quad,\quad k=K\quad,\quad l=\frac{L}{B}$$

With these changes, our equation of motion found in (3) reads:

$$\begin{aligned} \ddot{\varphi}(x,y) &= \frac{KB^2}{M}(\varphi(x+l,y)-2\varphi(x,y)+\varphi(x-l,y)) \\ &+ \frac{KB^2}{M}(\varphi(x,y+l)-2\varphi(x,y)+\varphi(x,y-l)) \end{aligned}$$

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

$$\begin{aligned} \ddot{\varphi}(x,y) &= \frac{KL^2}{M}\times\frac{\varphi(x+l,y)-2\varphi(x,y)+\varphi(x-l,y)}{l^2} \\ &+ \frac{KL^2}{M}\times\frac{\varphi(x,y+l)-2\varphi(x,y)+\varphi(x,y-l)}{l^2} \end{aligned}$$

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

$$\ddot{\varphi}(x,y)=\frac{KL^2}{M}\left(\frac{d^2\varphi}{dx^2}+\frac{d^2\varphi}{dy^2}\right)(x,y)$$

As a conclusion to this, we are led to the following wave equation in two dimensions, with $v=\sqrt{K/M}\cdot L$ being the propagation speed of our wave:

$$\ddot{\varphi}(x,y)=v^2\left(\frac{d^2\varphi}{dx^2}+\frac{d^2\varphi}{dy^2}\right)(x,y)$$

But we recognize at right the Laplace operator, and we are done. As before in 1D, there is of course some discussion to be made here, arguing that our spring model in (3) is indeed the correct one. But do not worry, experiments confirm our findings.

(5) In 3 dimensions now, which is the case of the main interest, corresponding to our real-life world, the same argument carries over, and the wave equation is as follows:

$$\ddot{\varphi}(x,y,z)=v^2\left(\frac{d^2\varphi}{dx^2}+\frac{d^2\varphi}{dy^2}+\frac{d^2\varphi}{dz^2}\right)(x,y,z)$$

(6) Finally, the same argument, namely a lattice model, carries on in arbitrary N dimensions, and the wave equation here is as follows:

$$\ddot{\varphi}(x_1,\ldots,x_N)=v^2\sum_{i=1}^{N}\frac{d^2\varphi}{dx_i^2}(x_1,\ldots,x_N)$$

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

Summarizing, done with our second mathematical proof of the wave equation, and with this being certainly a good thing. Observe however that there are some subtleties in the above, for instance in relation with our conventions for the total spring constant K, which varies with the dimension N. We will be back to these issues, later.

As another comment, once we are in $N\geq 2$ dimensions, modifying the springs in our lattice model, as to allow a dissymmetry between horizontal and vertical, either at the level of spring lengths, or spring constants, will most likely lead to different results.

In order to reach now to some further insight into our spring models above, we must get deeper into elasticity. Indeed, the Hooke law that we used has behind it some trivial elasticity, of "linear" type, and understanding all this, and further modifying our models, according to what elasticity theory exactly says, is certainly an interesting question.

Observe that all this can only lead us too into a better understanding of the fact that the propagation speed is finite, $v<c$. Indeed, the Hooke law is something static, and for better understanding what happens dynamically, we must certainly go into elasticity.

As a starting point for all this, we have the following result:

THEOREM 1.12. *The wave equation can be understood as well directly, as a wave propagating through a linear elastic medium, via stress.*

PROOF. This is indeed something very standard, with the $N=1$ picture involving a pulse propagating through a bar, and with at $N\geq 2$ something of a similar type:

(1) In the 1D case, assume that we have a bar of length L, made of linear elastic material. The stiffness of the bar is then the following quantity, with A being the cross-sectional area, and with E being the Young modulus of the material:

$$K=\frac{EA}{L}$$

Now when sending a pulse, this propagates as follows, M being the total mass:

$$\ddot{\varphi}=\frac{EAL}{M}\cdot\varphi''(x)$$

Bur since $V = AL$ is the volume, with $\rho = M/V$ being the density, we have:

$$\ddot{\varphi} = \frac{E}{\rho} \cdot \varphi''(x)$$

Thus, as a conclusion, the wave propagates with speed $v = \sqrt{E/\rho}$.

(2) In two or more dimensions, the study, and final result, are similar. □

As mentioned in the above, the next question which appears is that of understanding how exactly the various mechanical waves propagate through solids, liquids and gases, and what corrections to the wave equation are needed, in each case. We will defer the discussion here, which needs some thermodynamics preliminaries, to chapter 2 below.

1c. Basic solutions

We have seen so far that the mechanical waves propagate according to the wave equation $\ddot{\varphi} = v^2\Delta\varphi$. Let us do now some math. In 1D, we have the following result:

THEOREM 1.13. *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 linear combination of such basic solutions is a solution of the wave equation.

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 clear. It is in fact possible to prove, using Fourier analysis, that any solution of the 1D wave equation appears in this way. □

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:

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

In fact, with a bit more work, we can fully solve the 1D wave equation. In order to explain this, we will need a standard calculus result, as follows:

PROPOSITION 1.14. *The derivative of a function of type*

$$\varphi(x) = \int_{g(x)}^{h(x)} f(s)ds$$

is given by the formula $\varphi'(x) = f(h(x))h'(x) - f(g(x))g'(x)$.

PROOF. Consider a primitive of the function that we integrate, $F' = f$. We have:

$$\begin{aligned}\varphi(x) &= \int_{g(x)}^{h(x)} f(s)ds\\ &= \int_{g(x)}^{h(x)} F'(s)ds\\ &= F(h(x)) - F(g(x))\end{aligned}$$

By using now the chain rule for derivatives, we obtain from this:

$$\begin{aligned}\varphi'(x) &= F'(h(x))h'(x) - F'(g(x))g'(x)\\ &= f(h(x))h'(x) - f(g(x))g'(x)\end{aligned}$$

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

Now back to the 1D waves, the general result here, due to d'Alembert, along with a little more, in relation with our lattice models above, is as follows:

THEOREM 1.15. *The solution of the 1D wave equation with initial value conditions* $\varphi(x,0) = f(x)$ *and* $\dot{\varphi}(x,0) = g(x)$ *is given by the d'Alembert formula, namely:*

$$\varphi(x,t) = \frac{f(x-vt) + f(x+vt)}{2} + \frac{1}{2v}\int_{x-vt}^{x+vt} g(s)ds$$

In the context of our previous lattice model discretizations, what happens is more or less that the above d'Alembert integral gets computed via Riemann sums.

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

(1) Let us first check that the d'Alembert solution is indeed a solution of the wave equation $\ddot{\varphi} = v^2\varphi''$. The first time derivative is computed as follows:

$$\dot{\varphi}(x,t) = \frac{-vf'(x-vt) + vf'(x+vt)}{2} + \frac{1}{2v}(vg(x+vt) + vg(x-vt))$$

The second time derivative is computed as follows:

$$\ddot{\varphi}(x,t) = \frac{v^2f''(x-vt) + v^2f(x+vt)}{2} + \frac{vg'(x+vt) - vg'(x-vt)}{2}$$

Regarding now space derivatives, the first one is computed as follows:

$$\varphi'(x,t) = \frac{f'(x-vt) + f'(x+vt)}{2} + \frac{1}{2v}(g'(x+vt) - g'(x-vt))$$

As for the second space derivative, this is computed as follows:

$$\varphi''(x,t) = \frac{f''(x-vt) + f''(x+vt)}{2} + \frac{g''(x+vt) - g''(x-vt)}{2v}$$

Thus we have indeed $\ddot{\varphi} = v^2\varphi''$. As for the initial conditions, $\varphi(x,0) = f(x)$ is clear from our definition of φ, and $\dot{\varphi}(x,0) = g(x)$ is clear from our above formula of $\dot{\varphi}$.

(2) Conversely now, we must show that our solution is unique, but instead of going here into abstract arguments, we will simply solve our equation, which among others will doublecheck the computations in (1). Let us make the following change of variables:

$$\xi = x - vt \quad , \quad \eta = x + vt$$

With this change of variables, which is quite tricky, mixing space and time variables, our wave equation $\ddot{\varphi} = v^2\varphi''$ reformulates in a very simple way, as follows:

$$\frac{d^2\varphi}{d\xi d\eta} = 0$$

But this latter equation tells us that our new ξ, η variables get separated, and we conclude from this that the solution must be of the following special form:

$$\varphi(x,t) = F(\xi) + G(\eta) = F(x-vt) + G(x+vt)$$

Now by taking into account the intial conditions $\varphi(x,0) = f(x)$ and $\dot{\varphi}(x,0) = g(x)$, and then integrating, we are led to the d'Alembert formula in the statement.

(3) In regards now with our discretization questions, by using a 1D lattice model with balls and springs as before, what happens to all the above is more or less that the above d'Alembert integral gets computed via Riemann sums, in our model, as stated. □

1d. Spherical coordinates

In higher dimensions things become more complicated, because the study of the wave equation here is usually done via spherical coordinates, leading to spherical harmonics and some related mathematics, and with this being of course something very natural, motivated by the various spherical waves that we can meet in the real life.

In relation with our discretization questions, we run here into a bit of trouble, because we would like our lattice model to be rotationally invariant, and this is certainly not the case with our models so far, unless of course we are in $N = 1$ dimensions.

In order to comment on this, in usual $N = 3$ dimensions, let us first leave aside the wave equation $\ddot{\varphi} = v^2\Delta\varphi$, and focus on the Laplace operator Δ itself. We will need:

PROPOSITION 1.16. *We have spherical coordinates in* 3 *dimensions,*

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

the corresponding Jacobian being $J(r, s, t) = r^2\sin s$.

PROOF. The fact that we have indeed spherical coordinates is clear. Regarding now the Jacobian, this is given by the following formula:

$$\begin{aligned} & J(r,s,t) \\ = & \begin{vmatrix} \cos s & -r\sin s & 0 \\ \sin s\cos t & r\cos s\cos t & -r\sin s\sin t \\ \sin s\sin t & r\cos s\sin t & r\sin s\cos t \end{vmatrix} \\ = & r^2\sin s\sin t\begin{vmatrix} \cos s & -r\sin s \\ \sin s\sin t & r\cos s\sin t \end{vmatrix} + r\sin s\cos t\begin{vmatrix} \cos s & -r\sin s \\ \sin s\cos t & r\cos s\cos t \end{vmatrix} \\ = & r\sin s\sin^2 t\begin{vmatrix} \cos s & -r\sin s \\ \sin s & r\cos s \end{vmatrix} + r\sin s\cos^2 t\begin{vmatrix} \cos s & -r\sin s \\ \sin s & r\cos s \end{vmatrix} \\ = & r\sin s(\sin^2 t + \cos^2 t)\begin{vmatrix} \cos s & -r\sin s \\ \sin s & r\cos s \end{vmatrix} \\ = & r\sin s\times 1\times r \\ = & r^2\sin s \end{aligned}$$

Thus, we have indeed the formula in the statement. □

In relation now with the Laplace operator Δ, we have:

THEOREM 1.17. *The Laplace operator in spherical coordinates is*

$$\Delta=\frac{1}{r^2}\cdot\frac{d}{dr}\left(r^2\cdot\frac{d}{dr}\right)+\frac{1}{r^2\sin s}\cdot\frac{d}{ds}\left(\sin s\cdot\frac{d}{ds}\right)+\frac{1}{r^2\sin^2 s}\cdot\frac{d^2}{dt^2}$$

with our standard conventions for these coordinates, in 3D.

PROOF. There are several proofs here, a short, elementary one being as follows:

(1) Let us first see how Δ behaves under a change of coordinates $\{x_i\}\to\{y_i\}$, in arbitrary N dimensions. Our starting point is the chain rule for derivatives, namely:

$$\frac{d}{dx_i}=\sum_j\frac{d}{dy_j}\cdot\frac{dy_j}{dx_i}$$

By using this rule, then Leibnitz for products, then again this rule, we obtain:

$$\begin{aligned}\frac{d^2f}{dx_i^2}&=\sum_j\frac{d}{dx_i}\left(\frac{df}{dy_j}\cdot\frac{dy_j}{dx_i}\right)\\&=\sum_j\frac{d}{dx_i}\left(\frac{df}{dy_j}\right)\cdot\frac{dy_j}{dx_i}+\frac{df}{dy_j}\cdot\frac{d}{dx_i}\left(\frac{dy_j}{dx_i}\right)\\&=\sum_j\left(\sum_k\frac{d}{dy_k}\cdot\frac{dy_k}{dx_i}\right)\left(\frac{df}{dy_j}\right)\cdot\frac{dy_j}{dx_i}+\frac{df}{dy_j}\cdot\frac{d^2y_j}{dx_i^2}\\&=\sum_{jk}\frac{d^2f}{dy_kdy_j}\cdot\frac{dy_k}{dx_i}\cdot\frac{dy_j}{dx_i}+\sum_j\frac{df}{dy_j}\cdot\frac{d^2y_j}{dx_i^2}\end{aligned}$$

(2) Now by summing over i, we obtain the following formula, with A being the derivative of $x\to y$, that is to say, the matrix of partial derivatives dy_i/dx_j:

$$\begin{aligned}\Delta f&=\sum_{ijk}\frac{d^2f}{dy_kdy_j}\cdot\frac{dy_k}{dx_i}\cdot\frac{dy_j}{dx_i}+\sum_{ij}\frac{df}{dy_j}\cdot\frac{d^2y_j}{dx_i^2}\\&=\sum_{ijk}A_{ki}A_{ji}\frac{d^2f}{dy_kdy_j}+\sum_{ij}\frac{d^2y_j}{dx_i^2}\cdot\frac{df}{dy_j}\\&=\sum_{jk}(AA^t)_{jk}\frac{d^2f}{dy_kdy_j}+\sum_j\Delta(y_j)\frac{df}{dy_j}\end{aligned}$$

So, this will be the formula that we will need. Observe that this formula can be further compacted as follows, with all the notations being self-explanatory:

$$\Delta f=Tr(AA^tH_y(f))+<\Delta(y),\nabla_y(f)>$$

(3) Getting now to spherical coordinates, $(x,y,z)\to(r,s,t)$, the derivative of the inverse, obtained by differentiating x,y,z with respect to r,s,t, is given by:

$$A^{-1}=\begin{pmatrix}\cos s & -r\sin s & 0\\ \sin s\cos t & r\cos s\cos t & -r\sin s\sin t\\ \sin s\sin t & r\cos s\sin t & r\sin s\cos t\end{pmatrix}$$

The product $(A^{-1})^tA^{-1}$ of the transpose of this matrix with itself is then:

$$\begin{pmatrix}\cos s & \sin s\cos t & \sin s\sin t\\ -r\sin s & r\cos s\cos t & r\cos s\sin t\\ 0 & -r\sin s\sin t & r\sin s\cos t\end{pmatrix}\begin{pmatrix}\cos s & -r\sin s & 0\\ \sin s\cos t & r\cos s\cos t & -r\sin s\sin t\\ \sin s\sin t & r\cos s\sin t & r\sin s\cos t\end{pmatrix}$$

But everything simplifies here, and we have the following remarkable formula, which by the way is something very useful, worth to be memorized:

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

Now by inverting, we obtain the following formula, in relation with the above:

$$AA^t=\begin{pmatrix}1 & 0 & 0\\ 0 & 1/r^2 & 0\\ 0 & 0 & 1/(r^2\sin^2 s)\end{pmatrix}$$

(4) As a last ingredient, let us compute now the Laplacian of r,s,t. The formulae here are as follows, all following from some straightforward computations:

$$\Delta(r)=\frac{2}{r}\quad,\quad\Delta(s)=\frac{\cos s}{r^2\sin s}\quad,\quad\Delta(t)=0$$

We can now plug this data into the general formula found in (2), and we obtain:

$$\begin{aligned}\Delta f&=\frac{d^2f}{dr^2}+\frac{1}{r^2}\cdot\frac{d^2f}{ds^2}+\frac{1}{r^2\sin^2 s}\cdot\frac{d^2f}{dt^2}+\frac{2}{r}\cdot\frac{df}{dr}+\frac{\cos s}{r^2\sin s}\cdot\frac{df}{ds}\\ &=\frac{2}{r}\cdot\frac{df}{dr}+\frac{d^2f}{dr^2}+\frac{\cos s}{r^2\sin s}\cdot\frac{df}{ds}+\frac{1}{r^2}\cdot\frac{d^2f}{ds^2}+\frac{1}{r^2\sin^2 s}\cdot\frac{d^2f}{dt^2}\\ &=\frac{1}{r^2}\cdot\frac{d}{dr}\left(r^2\cdot\frac{df}{dr}\right)+\frac{1}{r^2\sin s}\cdot\frac{d}{ds}\left(\sin s\cdot\frac{df}{ds}\right)+\frac{1}{r^2\sin^2 s}\cdot\frac{d^2f}{dt^2}\end{aligned}$$

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

In connection now with our lattice model questions, still staying with a usual 3D lattice, $\varepsilon\mathbb{Z}^3\subset\mathbb{R}^3$, what we have to do is to modify the discrete geometry of our model, as to make it more in tune with the formula of Δ from Theorem 1.17.

However, such things are quite technical, and in what regards the final conclusions, regarding the discretization of the wave equation $\ddot{\varphi} = v^2\Delta\varphi$ itself, and its solutions, things remain quite unclear, at least at this stage of things. We will be back to this.

1e. Exercises

Exercises:

Exercise 1.18.

Exercise 1.19.

Exercise 1.20.

Exercise 1.21.

Exercise 1.22.

Exercise 1.23.

Bonus exercise.

CHAPTER 2

Fluid mechanics

2a. Gases, pressure

So far, we have met rather difficult problems, in relation with our discretization and lattice model program. Perhaps the waves, while being so familiar and intuitive, from our real life experience, and also mathematically quite well understood, are not the good targets for our program? Most likely yes, and in the present chapter we will orient ourselves towards certain seemingly more difficult questions, mainly belonging to fluid mechanics, with the hope that our lattice models can really help here.

And good news, coming in advance, this will be the case. In fact, later in this book, we will get so enthusiastic about our lattice model techniques, in the context of fluids, thermodynamics and related topics, that we will dedicate the whole Part II to statistical mechanics, and Part III too. And in what regards the waves, do not worry. We will keep them on our watchlist, and we will be back to them first in chapter 4, and then in Part IV, armed with some heavy statistical mechanics knowledge, accumulated until then.

This was for the plan. In practice now, in order to get started, we must first talk about gases. Things however are quite tricky here, and leaving aside excessive rigor, which does not fit well with thermodynamics, let us start with the following basic fact, which was the beginning of everything, going back to work of Boyle, Charles, Avogardo, Gay-Lussac, Clapeyron and others from the 17th, 18th and 19th centuries, and with final touches from Maxwell, Boltzmann, Gibbs and others, in the late 19th and early 20th centuries:

FACT 2.1. *The ideal gases satisfy the equation $PV = kT$, where:*

(1) *V is the volume of the gas, independently of the shape of the container used.*
(2) *P is the pressure of the gas, measured with a manometer.*
(3) *T is the temperature of the gas, measured with a thermometer.*
(4) *k is a constant, depending on the gas.*

That is, $PV = kT$ basically tells us that "pressure and temperature are the same thing".

At the first glance, for instance if you are a mathematician not used to this, this looks more like a joke. Why not defining then $P = T$ or vice-versa, you would say, and what is the point with that long list of distinguished gentlemen having worked hard on this.

Error. The point indeed comes from the following:

EXPLANATION 2.2. *In the equation of state $PV = kT$, as formulated above, the pressure P and the temperature T appear more precisely as follows,*

(1) *The manometer read comes from the gas molecules pushing a piston, so P is a statistical quantity, coming from the statistics of the molecular speeds,*
(2) *The thermometer read is something even more complicated, and T is as well a statistical quantity, coming from the statistics of the molecular speeds,*

so $PV = kT$ is something non-trivial, telling us that the mathematical machinery producing P, T, via manometer and thermometer, out of the molecular speeds, is the same.

Hope you got my point, and getting back now to historical details, Boyle, Charles, Avogardo, Gay-Lussac, Clapeyron, joined by Clausius, Carnot, Joule, Lord Kelvin and others, first observed $PV = kT$, and then reached to a good understanding of what this means, via an axiomatization of P and T. Later Maxwell started to look into the molecular speeds and their statistics, then Boltzmann came with a tough mathematical computation, proving $PV = kT$, and then, even later, Gibbs and others further built on all this, by formalizing modern thermodynamics, in the form that is still used today.

But probably too much talking, let us get to work. As a first result now, dealing with pressure only, and for the gases without collisions between molecules, we have:

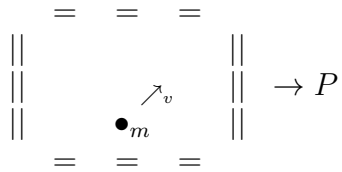
THEOREM 2.3. *The pressure P, volume V and total kinetic energy K of a gas, having point molecules, with no collisions between them, satisfy*

$$PV = \frac{2K}{d}$$

where $d = 1, 2, 3$ is the dimensionality of the gas, $d = 3$ for usual $3D$ gases.

PROOF. We can do this in several steps, as follows:

(1) Let us first assume that the gas is enclosed in a cubic volume, $V = L^3$. We want to compute the pressure P on the right wall. Since there are no collisions, we can assume by linearity that our gas has 1 molecule, having mass m and travelling at speed v. We must compute the pressure P exterted by this molecule on the right wall:



(2) We first look at a 1D gas. Our molecule hits the right wall at every $\Delta t = 2L/v$ interval, with its change of momentum being $\Delta p = 2mv$. We obtain, as desired:

$$P = \frac{F}{L^2} = \frac{\Delta p}{L^2 \Delta t} = \frac{2mv}{L^2 \cdot 2L/v} = \frac{mv^2}{L^3} = \frac{2K}{V}$$

(3) In the case of a d-dimensional gas, exactly the same computation takes place, but this time with v being replaced by its horizontal component v_1. Thus, we have:

$$P = \frac{mv_1^2}{V}$$

But, we have the following formula, with the equality on the right being understood in a statistical sense, our molecule being assumed to follow a random direction:

$$||v||^2 = v_1^2 + \ldots + v_d^2 = dv_1^2$$

Thus, the pressure in this case is given by the following formula, as desired:

$$P = \frac{m||v||^2}{dV} = \frac{2K}{dV}$$

(4) It remains to extend our result to arbitrary volume shapes. For this purpose, let us first redo the above computations for a parallelepiped, $V = L_1L_2L_3$. Here the above 1D gas computation carries on, and gives the same result, as follows:

$$P = \frac{F}{L_2L_3} = \frac{\Delta p}{L_2L_3\Delta t} = \frac{2mv}{L_2L_3 \cdot 2L_1/v} = \frac{mv^2}{L_1L_2L_3} = \frac{2K}{V}$$

Thus the d-dimensional computation carries on too, and gives the result.

(5) In order now to reach to arbitrary shapes, the idea will be that of stacking thin parallelepipeds, best approximating the shape that we have in mind, as follows:

(6) But for this purpose it is better to drop our assumption that the gas has 1 molecule, and use N molecules instead. With $\rho = N/V$ being the molecular density, and K_0 being the kinetic energy of a single molecule, our computation in (4) for the parallelepiped, with now N molecules instead of 1, reformulates as follows:

$$P = \frac{2K}{dV} = \frac{2NK_0}{dV} = \frac{2\rho VK_0}{dV} = \frac{2\rho K_0}{d}$$

(7) But this latter formula shows that the pressure has nothing to do with the precise volume V, but just with the molecular density $\rho = N/V$. Thus, we can stack indeed parallelepipeds, with of course the assumption that ρ is constant over these parallelepipeds, and we obtain that the above formula holds for an arbitrary volume shape V:

$$P = \frac{2\rho K_0}{d}$$

Now by getting back to the volume V, we obtain the following formula:

$$P = \frac{2\rho K_0}{d} = \frac{2NK_0}{dV} = \frac{2K}{dV}$$

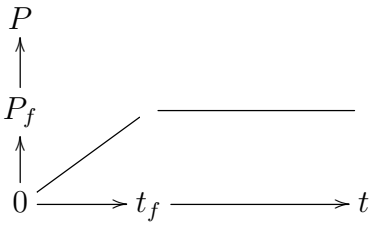
Thus, we are led to the conclusion in the statement. □

Observe that, when carefully looking at the above proof, and more specifically at step (3), we can see that pressure P is not something instantaneous, but rather something statistical. Thus, in practice, pressure reading is not instantaneous, and we have:

THEOREM 2.4. *In the context of a gas consisting of point molecules, with no collisions between them, the correct time for reading the correct pressure is*

$$t_f = \frac{2\sqrt{d}\,V^{1/3}}{||v||} \quad : \quad P_f = \frac{2K}{dV}$$

with $||v||$ being the average molecular speed, with the precise pressure reading being



and with this being taken in an approximate, statistical sense.

PROOF. We can do this in two steps, as follows:

(1) Let us first look at a 1D gas. We can assume that we are in a cubic container, $V = L^3$, and we know that each molecule i hits the right wall, where P is measured, at $\Delta t_i = 2L/|v_i|$ intervals. But with this picture in hand, it is quite clear that, on average, the pressure reading process will be linear, starting from $P = 0$, up to time $t_f = 2L/|v|$, with $|v|$ being the average molecular speed, where the correct pressure $P_f = 2K/V$ will be read, and constant at P_f afterwards. Now since $L = V^{1/3}$, this gives, as desired:

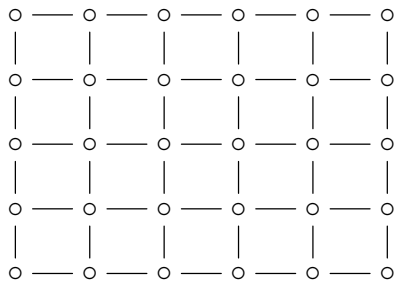
$$t_f = \frac{2V^{1/3}}{|v|} \quad : \quad P_f = \frac{2K}{V}$$

(2) In the general case now, that of a d-dimensional gas, with $d = 1, 2, 3$, the same argument carries on, with the only change being that each molecular speed $v_i \in \mathbb{R}^d$ is now replaced by its horizontal component $v_{i1} \in \mathbb{R}$, which by statistical reasons has squared magnitude as follows, as explained in the proof of Theorem 2.3:

$$v_{i1}^2 = ||v_i||^2/d$$

Thus, with respect to (1), the correct final pressure must be adjusted by a d factor, and becomes $P_f = 2K/dV$, as in Theorem 2.3. As for the correct reading time, this must be adjusted by a $\sqrt{d}$ factor, and becomes $t_f = 2\sqrt{d}\,V^{1/3}/||v||$, as claimed. □

Getting back now to our lattice model questions, the above results, dealing with pressure P only, are quite interesting, and suggest doing a number of things. For instance, we can focus on the piston surface, and model it with a square grid, as follows:



The point indeed is that what happens on the piston surface is some sort of "game" between the gas and the piston, reminding board games or arcade games, with the gas pushing here and there, with this or that speed, from time to time, and with the piston itself recording all these impulses, and converting them, via some mathematics, into pressure P. And this type of "game" can be obviously modelled on a grid, as above.

Along the same lines, a more complicated question is that of getting to 3D, with a lattice model for the whole gas, in the form of small 3D cubes constantly changing their status, depending on their neighbors. Note that this reminds a bit Conway's Game of Life, and also, importantly, that we can model as well internal collisions, in this way.

Speaking collisions between the gas molecules, at a more advanced level now, dealing with the internal mechanism of the $PV = kT$ formula, we have:

THEOREM 2.5 (Maxwell, Boltzmann). *The molecular speeds $v \in \mathbb{R}^3$ of a gas in thermal equilibrium are subject to the Maxwell-Boltzmann distribution formula*

$$P(v) = \left(\frac{m}{2\pi bT}\right)^{3/2} \exp\left(-\frac{m||v||^2}{2bT}\right)$$

with m being the mass of the molecules, and b being the Boltzmann constant.

PROOF. As before with other things, this is something in between fact and theorem. Maxwell came upon it as a fact, or perhaps as a sort of pseudo-theorem, and a bit later Boltzmann came with a proof. In what follows we will present the original argument of Maxwell, then briefly discuss Boltzmann's proof. Here is Maxwell's argument:

(1) We are looking for the precise probability distribution P of the molecular speeds $v = (v_1, v_2, v_3)$ which makes the mechanics of gases work. Intuition tells us that P has no corellations between the x, y, z directions of space, and so we must have:

$$P(v) = f(v_1)g(v_2)h(v_3)$$

Moreover, by rotational symmetry the functions f, g, h must coincide, and so:

$$P(v) = f(v_1)f(v_2)f(v_3)$$

(2) Further thinking, again invoking rotational symmetry, leads to the conclusion that $P(v)$ must depend only on the magnitude $||v||$ of the velocity $v \in \mathbb{R}^3$, and not on the direction. Thus, we must have as well a formula of the following type:

$$P(v) = \varphi(||v||^2)$$

(3) Now by comparing the requirements in (1) and (2), we are led via some math to the conclusion that φ must be an exponential, which amounts in saying that:

$$P(v) = \lambda \exp\left(-C||v||^2\right)$$

(4) Obviously we must have $C > 0$, for things to be bounded, and then by integrating we can obtain λ as function of C, and our formula becomes:

$$P(v) = \left(\frac{C}{\pi}\right)^{3/2} \exp\left(-C||v||^2\right)$$

(5) It remains to find the value of $C > 0$. But for this purpose, observe that, now that we have our distribution, be that still depending on $C > 0$, we can compute everything that we want to, just by integrating. In particular, we find that on average:

$$v_1^2 = v_2^2 = v_3^2 = \frac{1}{2C}$$

Thus the average magnitude of the molecular speed is given by:

$$||v|| = \frac{3}{2C}$$

It follows that the average kinetic energy of the molecules is:

$$K_0 = \frac{m||v||^2}{2} = \frac{3m}{4C}$$

(6) On the other hand, recall from our discussions above that one of the many equivalent formulations of $PV = kT$, using $PV = 2K/3$, was as follows:

$$\frac{2K_0}{3} = bT$$

Thus we obtain $m/(2C) = bT$, and so $C = m/(2bT)$, as desired.

(7) Observe that the above proof has in fact little physical content, with the whole thing being basically obtained by using $PV = 2K/3$, which is a mathematical theorem,

then $PV = kT$, which is a crucial physics fact, and finally by invoking several times a number of clever symmetry arguments, and doing some calculus.

(8) However, Boltzmann came later with a more rigorous argument, fully establishing the above formula, via detailed computations in relation with the internal collisions. □

So long for basic thermodynamics. As before, an interesting question is that of coming up with a 3D lattice model for the gas, in the form of small 3D cubes constantly changing their status, depending on their neighbors, a bit like in Conway's Game of Life.

2b. States of matter

Let us discuss now some basic theory, regarding the states of matter. Our scope will be quite broad, because we would like to talk about all sorts of matter, solid, liquid or gaseous, with a look into the extremes $T \to 0$ and $T >> 0$ too, where other forms of matter appear, and finally with our matter being as 3D as possible, with this meaning occupying a precise body $B \subset \mathbb{R}^3$, instead of just a volume $V = vol(B) \in \mathbb{R}$.

In order to discuss this, we first need to talk about work, motivated by:

QUESTIONS 2.6. *What is the work done by a gas pushing a piston? What about the work done by a gas evolving on a path γ, in the state space $f(P, V, T) = 0$?*

So, work. You have certainly experienced that, for instance when solving our exercises, and you might also know that what we call in physics "work", or more specifically "mechanical work", is something quite subtle, requiring both non-trivial math and physics preliminaries. So, let us start with this, math and physics preliminaries.

Regarding the mathematics, you surely know what a derivative or an integral is, and what the main properties of the derivatives and the integrals are. As a continuation of that, we will need to know as well about contour integrals. We have here:

THEOREM 2.7. *Given a path $\gamma \subset \mathbb{R}^3$, we can talk about integrals of type*

$$I = \int_\gamma f(x)dx_1 + g(x)dx_2 + h(x)dx_3$$

with $f, g, h : \mathbb{R}^3 \to \mathbb{R}$, which are independent on the chosen parametrization of the path.

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

(1) Regarding the statement itself, assume indeed that we have a path in $\mathbb{R}^3$, which can be best thought of as corresponding to a function as follows:

$$\gamma : [a, b] \to \mathbb{R}^3$$

Observe that this function γ is not exactly the path itself, for instance because the following functions produce the same path, parametrized differently:

$$\begin{array}{c}\delta : [0, b-a] \to \mathbb{R}^3 \quad , \quad \delta(t) = \gamma(t+a)\\ \varepsilon : [0,1] \to \mathbb{R}^3 \quad , \quad \varepsilon(t) = \delta((b-a)t)\\ \varphi : [0,1] \to \mathbb{R}^3 \quad , \quad \varphi(t) = \varepsilon(t^2)\\ \psi : [0,1] \to \mathbb{R}^3 \quad , \quad \psi(t) = \varepsilon(1-t)\\ \vdots\end{array}$$

Our claim, however, is that we can talk about integrals as follows, with $f, g, h : \mathbb{R}^3 \to \mathbb{R}$, which are independent on the chosen parametrization of our path:

$$I = \int_\gamma f(x)dx_1 + g(x)dx_2 + h(x)dx_3$$

(2) In order to prove this, let us choose a parametrization $\gamma : [a, b] \to \mathbb{R}^3$ as above. This parametrization has as components three functions $\gamma_1, \gamma_2, \gamma_3$, given by:

$$\gamma = (\gamma_1, \gamma_2, \gamma_3) : [a, b] \to \mathbb{R}^3$$

In order to construct the integral I, it is quite clear, by suitably cutting our path into pieces, that we can restrict the attention to the case where all three components $\gamma_1, \gamma_2, \gamma_3 : [a, b] \to \mathbb{R}$ are increasing, or decreasing. Thus, we can assume that these three components are as follows, increasing or decreasing, and bijective on their images:

$$\gamma_i : [a, b] \to [a_i, b_i]$$

(3) Moreover, by using the obvious symmetry between the coordinates x_1, x_2, x_3, in order to construct I, we just need to construct integrals of the following type:

$$I_1 = \int_\gamma f(x)dx_1$$

(4) So, let us construct this latter integral I_1, under the assumptions in (2). The simplest case is when the first path, $\gamma_1 : [a, b] \to [a_1, b_1]$, is the identity:

$$\gamma_1 : [a, b] \to [a, b] \quad , \quad \gamma_1(x) = x$$

In other words, the simplest case is when our path is of the following form, with $\gamma_2, \gamma_3 : [a, b] \to \mathbb{R}$ being certain functions, that should be increasing or decreasing, as per our conventions (2) above, but in what follows we will not need this assumption:

$$\gamma(x_1) = (x_1, \gamma_2(x_1), \gamma_3(x_1))$$

But now, with this convention made, we can define our contour integral, or rather its first component, as explained above, as a usual one-variable integral, as follows:

$$I_1 = \int_a^b f(x_1, \gamma_2(x_1), \gamma_3(x_1))dx_1$$

(5) With this understood, let us examine now the general case, where the first path, $\gamma_1 : [a, b] \to [a_1, b_1]$, is arbitrary, increasing or decreasing, and bijective on its image. In this case we can reparametrize our curve, as to have it as in (4) above, as follows:

$$\tilde{\gamma} = (id, \gamma_2\gamma_1^{-1}, \gamma_3\gamma_1^{-1}) : [a_1, b_1] \to \mathbb{R}^3$$

Now since we want our integral $I_1 = \int_\gamma f(x)dx_1$ to be independent of the parametrization, we are led to the following formula for it, coming from the formula in (4):

$$\begin{aligned} I_1 &= \int_{\tilde{\gamma}} f(x)dx_1 \\ &= \int_{a_1}^{b_1} f(x_1, \gamma_2\gamma_1^{-1}(x_1), \gamma_3\gamma_1^{-1}(x_1))dx_1 \\ &= \int_a^b f(\gamma_1(y_1), \gamma_2(y_1), \gamma_3(y_1))\gamma_1'(y_1)dy_1 \end{aligned}$$

Here we have used at the end the change of variable formula, with $x_1 = \gamma_1(y_1)$.

(6) Thus, job done, we have our definition for the contour integrals, with the formula being as follows, obtained by using (5) for all three coordinates x_1, x_2, x_3:

$$\begin{aligned} I &= \int_a^b f(\gamma_1(y_1), \gamma_2(y_1), \gamma_3(y_1))\gamma_1'(y_1)dy_1 \\ &\quad + \int_a^b g(\gamma_1(y_2), \gamma_2(y_2), \gamma_3(y_2))\gamma_2'(y_2)dy_2 \\ &\quad + \int_a^b h(\gamma_1(y_3), \gamma_2(y_3), \gamma_3(y_3))\gamma_3'(y_3)dy_3 \end{aligned}$$

And with this, we are led to the conclusion in the statement. □

Let us record as well the following more compact form of Theorem 2.7 and its proof, containing what is needed to know, in practice, for computing contour integrals:

THEOREM 2.8. *The contour integrals over a curve $\gamma : [a, b] \to \mathbb{R}^3$ are given by*

$$\int_\gamma < F(x), dx >= \int_a^b < F(\gamma(y)), \gamma'(y)dy >$$

valid for any $F : \mathbb{R}^3 \to \mathbb{R}^3$, where on the right $\gamma'(y)dy = (\gamma_i'(y_i)dy_i)_i$.

PROOF. This is a fancy reformulation of what we did in Theorem 2.7 and its proof. Indeed, with the notation $F = (F_1, F_2, F_3) = (f, g, h)$, the integral computed there is:

$$\int_\gamma < F(x), dx >= \int_\gamma F_1(x)dx_1 + F_2(x)dx_2 + F_3(x)dx_3$$

As for the value of this integral, according to the proof of Theorem 2.7, this is:

$$\begin{aligned}\int_\gamma < F(x), dx > &= \int_a^b F_1(\gamma_1(y_1), \gamma_2(y_1), \gamma_3(y_1))\gamma_1'(y_1)dy_1\\ &+ \int_a^b F_2(\gamma_1(y_2), \gamma_2(y_2), \gamma_3(y_2))\gamma_2'(y_2)dy_2\\ &+ \int_a^b F_3(\gamma_1(y_3), \gamma_2(y_3), \gamma_3(y_3))\gamma_3'(y_3)dy_3\end{aligned}$$

Now observe that we can write this in a more compact way, as follows:

$$\begin{aligned}\int_\gamma < F(x), dx > &= \int_a^b F_1(\gamma(y_1))\gamma'(y_1)dy_1\\ &+ \int_a^b F_2(\gamma(y_2))\gamma'(y_2)dy_2\\ &+ \int_a^b F_3(\gamma(y_3))\gamma'(y_3)dy_3\end{aligned}$$

And we can do even better. Indeed, we have only one integral here, $\int_a^b$, and in order to best express the integrand, consider the formal vector in the statement, namely:

$$\gamma'(y)dy = \begin{pmatrix}\gamma_1'(y_1)dy_1\\ \gamma_2'(y_2)dy_2\\ \gamma_3'(y_3)dy_3\end{pmatrix}$$

Our integrand appears then as the scalar product of $F(\gamma(y))$ with this latter vector $\gamma'(y)dy$, so our formula above for the contour integral takes the following form:

$$\int_\gamma < F(x), dx >= \int_a^b < F(\gamma(y)), \gamma'(y)dy >$$

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

Many other things can be said about the contour integrals, and we will be back to this with some further mathematics, whenever needed. For the moment, the above material, which is more or less the definition of these contour integrals, will do.

More concretely now, let us temporarily forget about the paths γ, and have a look at the quantities which are to be integrated, namely:

$$\alpha = F_1(x)dx_1 + F_2(x)dx_2 + F_3(x)dx_3$$

Obviously, these are something rather mathematical, and many things can be said here. However, we can have some physical intuition on them. Assume indeed that we

are given a function as follows, that you can think for instance as corresponding to an external force, with $F(x) \in \mathbb{R}^3$ being the force vector applied at a given point $x \in \mathbb{R}^3$:

$$F : \mathbb{R}^3 \to \mathbb{R}^3$$

By writing $F = (F_1, F_2, F_3)$, we can then consider the following quantity, and when $F : \mathbb{R}^3 \to \mathbb{R}^3$ varies, we obtain exactly the abstract quantities α considered above:

$$< F(x), dx >= F_1(x)dx_1 + F_2(x)dx_2 + F_3(x)dx_3$$

Thus, all in all, what we have done in the above with our construction of contour integrals, was to define quantities as follows, with γ being a path in $\mathbb{R}^3$, and with $F : \mathbb{R}^3 \to \mathbb{R}^3$ being a certain function, that we can think of, if we want, as being a force:

$$I = \int_\gamma < F(x), dx >$$

Which brings us into physics. Indeed, by assuming now that $F : \mathbb{R}^3 \to \mathbb{R}^3$ does correspond to a force, we can formulate the following definition:

DEFINITION 2.9. *The work done by a force $F = F(x)$ for moving a particle from point $p \in \mathbb{R}^3$ to point $q \in \mathbb{R}^3$ via a given path $\gamma : p \to q$ is the following quantity:*

$$W(\gamma) = \int_\gamma < F(x), dx >$$

We say that F is conservative if this work quantity $W(\gamma)$ does not depend on the chosen path $\gamma : p \to q$, and in this case we denote this quantity by $W(p, q)$.

Now back to gases, we want to compute the work W done by a gas, pushing a piston. This work is quite easy to compute, and we are led to the following statement:

PROPOSITION 2.10. *The mechanical work done by a gas pushing a piston is*

$$W = \int_a^b PdV$$

where $[a, b]$ is the interval where the piston has traveled.

PROOF. In order to compute mechanical work W done by the gas, observe that, infinitesimally, if we denote by dl the distance traveled by the piston in time dt, we have the following formula, with S being the area of the piston:

$$dW = Fdl = PSdl = Pd(Sl) = PdV$$

Here we have used the standard interpretation of the pressure P, as explained before. Now assuming that the piston has traveled from a to b, we obtain:

$$W = \int_a^b dW = \int_a^b PdV$$

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

Obviously what we have above is a general formula, which will hold in more general situations, not necessarily involving a container with a piston. So, let us formulate:

THEOREM 2.11. *The mechanical work done by a gas evolving on a path γ, in the state space $f(P,V,T)=0$, is given by:*

$$W_\gamma=\int_\gamma PdV$$

When representing the transformation γ as a one-variable function $P=\varphi(V)$, with T being determined at each moment by $f(P,V,T)=0$, this formula reads

$$W_\gamma=\int_{V_0}^{V_1}\varphi(V)dV$$

with V_0 being the initial volume, and V_1 being the final volume.

PROOF. This is quite clear from the previous discussion, involving the cylinder with a piston, the argument there carrying on to the general case without problems, and giving the first formula. As for the second formula, this follows from it, with again the case of the cylinder with a piston being a good illustration, φ being there linear. □

In order to discuss now examples, let us introduce as well the following notions:

DEFINITION 2.12. *A transformation of a gas is called:*

(1) *Isobaric, if P is constant.*
(2) *Isochoric, if V is constant.*
(3) *Isothermal, if T is constant.*

Here all terminology comes from Greek, with isos meaning equal, baros meaning weight, chora meaning space, and therme meaning heat. There is as well a 4th class of transformations, the adiabatic ones, coming from the Greek adiabatos, meaning impassable. These latter transformations are the scary ones, and more on them later.

As an illustration for Theorem 2.11, for an ideal gas, we have the following result:

THEOREM 2.13. *For an ideal gas, $PV=kT$, isothermally expanding, we have*

$$W=T\log\frac{V_1}{V_0}$$

where V_0 is the initial volume, and V_1 is the final volume.

PROOF. This follows indeed from Theorem 2.11, in either formulation:

(1) With the first formula the work is readily computed, as follows:

$$W=\int_\gamma PdV=kT\int_\gamma\frac{dV}{V}=T\log\frac{V_1}{V_0}$$

(2) With the second formulation, since we have $PV = kT$ by the equation of state, and T constant by assumption, the transformation is given by $\varphi(V) = kT/V$, and so:

$$W = \int_{V_0}^{V_1} \frac{kT}{V}\, dV = kT \int_{V_0}^{V_1} \frac{dV}{V} = T \log \frac{V_1}{V_0}$$

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

Finally, the first law of thermodynamics invites us to introduce, as a complement to the various elementary notions from Definition 2.12:

DEFINITION 2.14. *A transformation is called adiabatic if:*

$$Q = 0$$

Equivalently, we must have the energy conservation law $\Delta E = -W$.

Obviously, this is something a bit complicated. As basic examples here we have the transformations of the thermally insulated systems obeying to $Q = C\Delta T$, with $C > 0$. As an important result about adiabatics, regarding the ideal gases, we have:

THEOREM 2.15. *The adiabatic transformations of an ideal gas, $PV = kT$, satisfy*

$$TV^{K-1} = \text{constant}$$

where K is a modified version of k, appearing as follows:

$$K = \frac{C_P}{C_V} = 1 + \frac{k}{C_V}$$

Together with $PV = kT$, the above equation produces certain curves, called adiabatics.

PROOF. We use here the Joule formula $C_V dT + PdV = dE$. Since for an adiabatic transformation we have $dE = 0$, the Joule formula becomes:

$$C_V dT + PdV = 0$$

Now by using $PV = kT$, we succesively obtain:

$$\begin{aligned} PV = kT &\implies C_V dT + \frac{kT}{V}\, dV = 0 \\ &\implies \frac{dT}{T} + \frac{k}{C_V} \cdot \frac{dV}{V} = 0 \\ &\implies \log T + \frac{k}{C_V} \log V = \text{constant} \\ &\implies TV^{k/C_V} = \text{constant} \end{aligned}$$

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

All the above theory concerns the ideal gases, but in practice, things are more complicated than this. The key result here, due to Van der Waals, is as follows:

THEOREM 2.16. *Beyond the ideal gas setting, stating that we should have $PV = kT$, the gases are subject to the Van der Waals equation*

$$\left(P + \frac{\alpha}{V^2}\right)(V - \beta) = kT$$

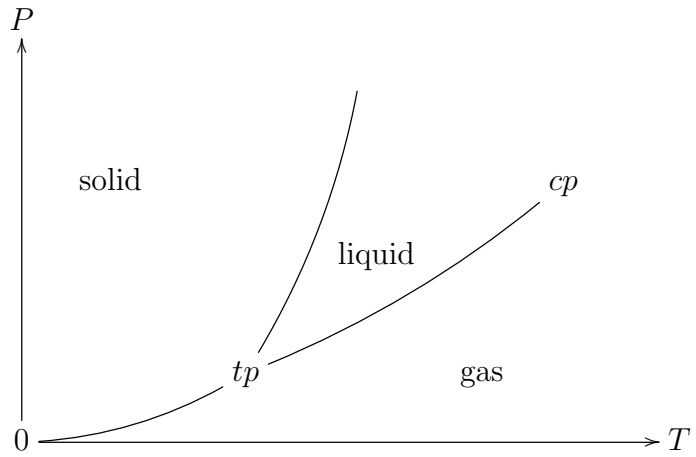
depending on two parameters $\alpha, \beta > 0$.

PROOF. This is something quite tricky, with the correction parameters $\alpha, \beta > 0$ appearing from a detailed study of the gas, from a kinetic viewpoint. It it possible of course to specify the parameters $\alpha, \beta > 0$, via various lists and tables, and even make some speculations on the precise meaning of these two parameters, using basic chemistry. □

The above result is of key importance, and takes us into rethinking everything that what we know about the ideal gases, which must be replaced with Van der Waals gases, at the advanced level. Among the main consequences of this replacement, the isobars, isochores, isothermals and adiabatics of the ideal gases, given by simple formulae, must be replaced by isobars, isochores, isothermals and adiabatics for the Van der Waals gases, which are no longer something trivial, with some interesting math being now involved.

Among others, the Van der Waals gas study makes appear some interesting points on the isothermals, called triple and critical points of the gas. Which makes the connection with the general theory of matter, whose basics can be summarized as follows:

FACT 2.17. *Ordinary matter appears in 3 forms, namely solid, liquid and gaseous, roughly appearing according to the following generic diagram*



with tp, cp standing for the triple and critical points. Also, at low or high temperatures we have interesting phenonema like Bose-Einstein condensation, and plasma.

Passed the gases, that we are quite familiar with, the problem is with the liquids and solids, plus of course with all sorts of exotic matter that can appear at the extremes.

2c. Fluid dynamics

Generally speaking, the fluids are by definition the non-solids, and as such, they fall into liquids, gases and plasma. Thus, a fluid is something quite complicated, whose understanding requires advanced thermodynamics, and states of matter theory. Also, we can also see that the modeling of a fluid can only be something quite complicated too, again requiring advanced thermodynamics, and states of matter theory.

This being said, our claim is that, at least for certain very simple types of fluids, and very basic questions regarding them, we can do some interesting modeling work, using models which are not that complicated, of similar type to those from chapter 1.

Getting started now, fluid mechanics is a complex science, which is organized by taking into account the two main properties that a fluid can have or not, which are:

(1) Compressibility.

(2) Viscosity.

Regarding compressibility, this is certainly a property of the gases, but not of most of the liquids, at least when idealized. We will assume here that our fluid is incompressible, which in practice means more or less that we are dealing with liquids, of rather "regular" type. However, this will be not the general rule, for instance because materials like sand, or snow, that we are particularly interested in, in view of their obvious strong link with classical mechanics, and with our modelling abilities so far, are incompressible too.

Regarding viscosity, this is something far more tricky. Intuitively, this comes from the mutual "friction" of the constituent molecules, when the fluid is moving, and with this being something quite difficult to model and understand, via precise mathematical equations. Without getting into details, for the moment, let us mention that, from this point of view, the fluids, or rather liquids, fall into 3 main classes, namely:

– Inviscid. This is intuitively the case of regular water, and other familiar liquids. However, this remains an idealization, with the true inviscid fluids, in the real life, being basically only the superfluids, met at very low temperatures.

– Newtonian. This is intuitively the case for most of the familiar visquous fluids, from the real life, whose viscosity is proportional to the applied stress, and with this proportionality being known as Newton's law on viscosity.

– Non-Newtonian. These are visquous fluids which do not obey to Newton's law on viscosity, and there are plenty of them, all very interesting, such as paint, toothpaste, ketchup and many more, not to forget basic things like snow or sand.

As a conclusion to this, we can see that, even when looking at the simplest 2 possible things that can be said about a fluid, namely compressibility and viscosity, we end up with a whole menagerie of fluids, with each of them corresponding to its own branch of fluid mechanics. In what concerns us, let us fix the following goal:

GOAL 2.18. *We are interested in the basic modelling of incompressible fluids, which can be either inviscid, or non-Newtonian.*

Here our incompressibility assumption comes from our discussion above regarding compressibility, and is certainly somehing quite natural, and with this being the simplest situation anyway, mathematically speaking, that we would like to study first. As for our second assumption, "either inviscid or non-Newtonian", this might sound a bit exotic, but the point here is that we want to avoid, at least in the beginning, the Newtonian, visquous case, whose basic mathematics is notoriously quite complicated.

Of course you might argue that non-Newtonian is more complicated than Newtonian, but my answer is that this is not exactly the case, with a pile of sand, and its mathematics, being probably something simpler than a visquous, Newtonian fluid flowing.

Speaking goals, let us discuss now as well the main PDE for the fluids, that we would like to recover, whenever possible, for the fluids that we are mainly interested in. As a first observation, which is of key importance in advanced fluid dynamics, we have:

OBSERVATION 2.19. *An incompressible fluid is right away something mathematical, whose dynamics is described by a diffeomorphism, evolving in time $t > 0$.*

Obviously, this is something very basic, coming from the very nature of the incompressibility property. To be more precise, assuming for instance that we have colored our fluid initially, say into tiny little cubes of red, yellow, green and so on, after some time $t > 0$ we will obviously still have a mixture of red, yellow, green and so on, appearing in equal parts, and the function $f : \mathbb{R}^3 \to \mathbb{R}^3$ mapping red to red, yellow to yellow, green to green, and so on, is what we call in mathematics a diffeomorphism.

In practice now, in order to reach from this to a PDE, let us assume that we are in the simplest case, where our fluid is inviscid, and also adiabatic, or with zero thermal conductivity. The dynamics of the diffeomorphisms $f : \mathbb{R}^3 \to \mathbb{R}^3$ will be then, intuitively speaking, basically governed by the mechanics of the red, yellow, green and so on components, and we are led in this way to the Euler equations, which are as follows:

$$\dot{u} + < u, \nabla > u = -\nabla w \quad , \quad < \nabla, u >= 0$$

Here u is the vector velocity field, and w is the thermodynamic work, with the quantity on the right from the first equation being as follows, p being the pressure:

$$\nabla w = \frac{\nabla p}{\rho}$$

With this replacement made, the Euler equations become:

$$\dot{u}+ < u, \nabla > u = -\frac{\nabla p}{\rho} \quad , \quad < \nabla, u >= 0$$

Here the first equation, which is the important one, is called the Euler momentum equation. This equation can be further complicated by adding an acceleration term $+g$ on the right, accounting for exterior forces, gravitational, or magnetic or of some other kind. As for the second equation, this is the incompressibility constraint.

Although we will not really need this here, at least at this stage of our discussion, let us record as well what happens in the visquous case, by staying as before in the incompressible setting. Here the Euler momentum equation gets replaced by the Navier-Stokes equation, which is as follows, with v being the kinematic viscosity:

$$\dot{u}+ < u, \nabla > u = -\frac{\nabla p}{\rho} + v\Delta u$$

Summarizing, in relation with our modelling questions for the incompressible fluids, we will be mostly interested in the Euler equation, and its versions.

Getting now to modeling questions, let us begin by recalling, and we insist on this, the fact that fluid mechanics is something quite complicated, and so that what we will be doing here, of rather elementary nature, will be quite speculatory. But, let's give this a go. We have several intuitive ways of modelling the incompressible fluids, as follows:

(1) With fluids modeled by little cubes, it is a bit unclear what we can get, because these cubes will obviously not slide exactly like fluids do. However, the little cubes will model well things like a pile of sand, because no friction is needed between our cubes for our pile to stand, as a pile of sand does stand, and this looks quite interesting.

(2) With little spheres, we are certainly close to reality, for somehow obvious reasons. However, at the mathematical level, things are quite tricky. In the static setting already, we are led to non-trivial questions, of sphere-packing flavor. And in the dynamic setting, which is the one that we are interested in, things can only be more complicated.

(3) In order to avoid the mathematics of the little cubes or little spheres, which obviously can be quite complicated, we can assume instead that we are dealing with "random-shaped" little pieces of material. Which is in fact, when thinking well, not that far from what happens in the real life, with materials like sand.

As before with some technical aspects of the wave equation, in relation with the PDE work on the subject, such questions are a bit obscure in the context of mainstream mathematics and physics, because one you have your PDE, you will go of course with that PDE, without bothering on where that PDE comes from. However, for our present

purposes, and more specifically in view of what we want to do in Parts II, III, IV in this book, such abstract aspects are important, and must be understood.

There are many interesting questions here, coming as a continuation of the modeling work that we did before, and with phenomena like soil liquefaction confirming this. In fact, one interesting feature of what happens here concerns the liquid/solid mixtures.

Solids, granularity, and with all this being mainly motivated by soil liquefaction. There are many interesting modeling questions here, and at the purely mathematical level, we are led as well into some quite subtle questions, regarding sphere packings.

2d. Heat diffusion

We discuss here the basics of heat diffusion, with more advanced aspects left for later, in chapter 4. Let us start with some generalities. We have, as general principle:

PRINCIPLE 2.20. *The amount of heat Q received during a transformation is a thermodynamical quantity, in relation with temperature, and which normally depends on the chosen path in the state space. However, we usually have*

$$Q = C\Delta T$$

with $C > 0$ being a constant, called heat capacity of the material, independently on the chosen path in the state space, between the initial and the final point.

Observe that C depends on the given quantity of the material. A better parameter would be the specific heat capacity, given by the following formula, m being the mass:

$$c = \frac{C}{m}$$

But more on heat capacity in a moment. Perhaps the most illustrating for all this, Principle 2.20 and its consequences, are some simple examples, as follows:

PROPOSITION 2.21. *The following happen:*

(1) *For a cyclic transformation we have $Q = W$, regardless of the chosen path in the state space, between the initial and the final point.*
(2) *Assuming that the heat received is subject to $Q = C\Delta T$, with $C > 0$, for a thermally insulated system we have $\Delta E = -W$.*

PROOF. Both these assertions are clear from definitions, as follows:

(1) For a cyclic transformation we have $\Delta E = 0$, and so $Q = W$, as claimed.

(2) Our combined assumptions mean $Q = C\Delta T = 0$, so $\Delta E = -W$, as claimed. □

Let us do now some math, based on all this. We first have:

PROPOSITION 2.22. *The heat capacity $C = Q/\Delta T$ is given by*

$$C = \frac{\Delta E + W}{\Delta T}$$

its version, the heat capacity at constant volume, is given by

$$C_V = \left(\frac{dE}{dT}\right)_V$$

and its other version, the heat capacity at constant pressure, is given by:

$$C_P = \left(\frac{dE}{dT}\right)_P + P\left(\frac{dV}{dT}\right)_P$$

Also, we have in general $C > C_V$, since objects tend to expand when heated.

PROOF. The first law of theormodynamics gives right away the first formula:

$$C = \frac{Q}{\Delta T} = \frac{\Delta E + W}{\Delta T}$$

Regarding now the second formula, at constant volume V, we have indeed:

$$C_V = \left(\frac{\Delta E}{\Delta T}\right)_V = \left(\frac{dE}{dT}\right)_V$$

As for the third formula, the proof here is similar, as follows:

$$C_P = \left(\frac{\Delta E + P\Delta V}{\Delta T}\right)_P = \left(\frac{dE}{dT}\right)_P + P\left(\frac{dV}{dT}\right)_P$$

Finally, in what regards the last assertion, this is something well-known. □

Let us record as well a related result for the ideal gases, as follows:

PROPOSITION 2.23. *For an ideal gas, $PV = kT$, we have the formula*

$$C_P = C_V + k$$

relating the heat capacity at constant pressure, and at constant volume.

PROOF. We have indeed the following computation, based on the above:

$$\begin{aligned} C_P &= \left(\frac{dE}{dT}\right)_P + P\left(\frac{dV}{dT}\right)_P \\ &= \left(\frac{dE}{dT}\right)_P + P \cdot \frac{d}{dT}\left(\frac{kT}{P}\right) \\ &= \left(\frac{dE}{dT}\right)_P + P \cdot \frac{k}{P} \\ &= C_V + k \end{aligned}$$

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

Moving ahead now, in the general case, we have the following key result:

THEOREM 2.24. *We have the following heat equations,*

$$dQ = \left(\frac{dE}{dP}\right)_V dP + \left[\left(\frac{dE}{dV}\right)_P + P\right] dV$$

$$dQ = \left[\left(\frac{dE}{dP}\right)_T + P\left(\frac{dV}{dP}\right)_T\right] dP + \left(\frac{dH}{dT}\right)_P dT$$

$$dQ = \left[\left(\frac{dE}{dV}\right)_T + P\right] dV + \left(\frac{dE}{dT}\right)_V dT$$

where $H = E + PV$, called enthalpy of the system.

PROOF. This follows indeed by doing some computations, as follows:

(1) We first have the following formulae, coming from definitions:

$$dE(P,V) = \left(\frac{dE}{dP}\right)_V dP + \left(\frac{dE}{dV}\right)_P dV$$

$$dE(P,T) = \left(\frac{dE}{dP}\right)_T dP + \left(\frac{dE}{dT}\right)_P dT$$

$$dE(V,T) = \left(\frac{dE}{dV}\right)_T dV + \left(\frac{dE}{dT}\right)_V dT$$

(2) In order to advance, we use the usual work formula, rewritten as $dW = PdV$. By plugging this formula into the first law, the first law becomes:

$$dQ = dE + dW = dE + PdV$$

(3) Now by using this latter formula, the equations found in (1) become:

$$dQ = \left(\frac{dE}{dP}\right)_V dP + \left[\left(\frac{dE}{dV}\right)_P + P\right] dV$$

$$dQ = \left(\frac{dE}{dP}\right)_T dP + \left(\frac{dE}{dT}\right)_P dT + PdV$$

$$dQ = \left[\left(\frac{dE}{dV}\right)_T + P\right] dV + \left(\frac{dE}{dT}\right)_V dT$$

(4) In what regards the first and third equations, these are those in the statement. In what regards the second equation, we can work a bit more on it, and make it look better. Indeed, we can use here the following formula, coming from definitions:

$$dV = \left(\frac{dV}{dP}\right)_T dP + \left(\frac{dV}{dT}\right)_P dT$$

With this formula in hand, the second equation can be rewritten as:

$$dQ = \left[\left(\frac{dE}{dP}\right)_T + P\left(\frac{dV}{dP}\right)_T\right]dP + \left(\frac{d(E+PV)}{dT}\right)_P dT$$

(5) Still in regards with this second equation, with the aim of making it even better, let us introduce a new state function of the system, called enthalpy, as follows:

$$H = E + PV$$

In terms of this function, the second equation as modified in (4) becomes:

$$dQ = \left[\left(\frac{dE}{dP}\right)_T + P\left(\frac{dV}{dP}\right)_T\right]dP + \left(\frac{dH}{dT}\right)_P dT$$

Thus, we have as well the second equation in the statement, and we are done. □

As an immediate consequence of the above result, we have:

PROPOSITION 2.25. *The heat capacities at constant V and P are*

$$C_V = \left(\frac{dE}{dT}\right)_V \quad , \quad C_P = \left(\frac{dE}{dT}\right)_P + P\left(\frac{dV}{dT}\right)_P$$

with the derivatives being taken as usual, as above.

PROOF. This follows indeed from the formulae in Theorem 2.24. □

Before going ahead, let us comment on the enthalpy function $H = E + PV$ appearing in Theorem 2.24. We came upon this function via some math, but we have:

COMMENT 2.26. *The enthalpy function of a system, given by*

$$H = E + PV$$

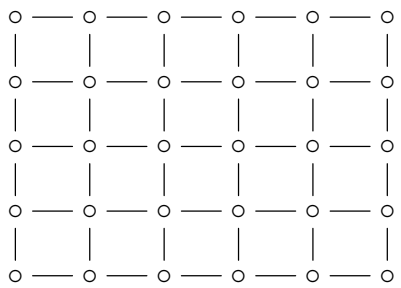
is what's needed for creating the system out of nothing, and putting it into environment.

To be more precise, assume that you are a magician, and want to create a rabbit. You first need energy E for creating the rabbit. But then you also need energy for putting it into environment, with this meaning pushing out the atmosphere, for creating space for the rabbit. And assuming that we are at 1 atm, this extra energy needed is PV.

Due to this, enthalpy is an important quantity in chemistry computations, and this even for simple reactions like $2H_2 + O_2 = 2H_2O$, because the basic quantum chemistry formulae for them, involving electrons and so on, do not take into account the energy needed for putting the output into environment, and must be therefore fine-tuned.

Regarding now heat diffusion, the simplest question here, studied and understood since long, concerns a container containing two gases, having initial different temperatures

$T_1 < T_2$, separated by a membrane. Heat transfer goes on, and obviously, we can model this by focusing on the membrane, with a basic grid model for it, as follows:



As before with questions regarding pistons and pressure, there is some sort of "game" played by the two gases, over this grid, and the question is whether we can model all this, and then recover the known results about heat diffusion, in this setting.

At a more advanced level, we can remove the membrane. Again, there is some sort of "game" here, played by the two gases, which can be 2D or 3D, depending on modeling. Also, in this setting, we can actually keep the membrane, but allow it to inflate.

2e. Exercises

Exercises:

EXERCISE 2.27.

EXERCISE 2.28.

EXERCISE 2.29.

EXERCISE 2.30.

EXERCISE 2.31.

EXERCISE 2.32.

Bonus exercise.

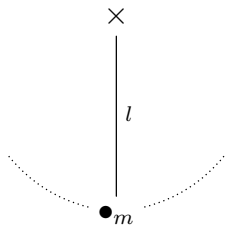
CHAPTER 3

Gravity models

3a. The pendulum

In this chapter we go back to a more normal pace, and to more mathematics, and systematically investigate discretization questions and lattice models. We will start, and no surprise here, with the beginning, namely classical mechanics. We first have:

DEFINITION 3.1. *A simple pendulum is a device of type*



consisting of a bob of mass m, attached to a rigid rod of length l.

In order to study the physics of the pendulum, which can easily lead to a lot of complicated computations, when approached with bare hands, the most convenient is to use the notion of energy. For a particle moving under the influence of a force F, the position x, speed v and acceleration a are related by the following formulae:

$$v = \dot{x} \quad , \quad a = \dot{v} = \ddot{x} \quad , \quad F = ma$$

The kinetic energy of our particle is then given by the following formula:

$$T = \frac{mv^2}{2}$$

By differentiating with respect to time t, we obtain the following formula:

$$\dot{T} = mv\dot{v} = mva = Fv$$

Now by integrating, also with respect to t, this gives the following formula:

$$T = \int Fvdt = \int F\dot{x}dt = \int Fdx$$

But this suggests to define the potential energy V by the following formula, up to a constant, with the derivative being with respect to the space variable x:

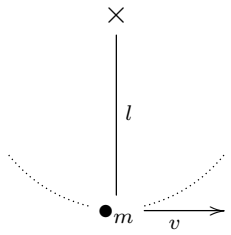
$$V' = -F$$

Indeed, we know from the above that we have $T' = F$, so if we define the total energy to be $E = T + V$, then this total energy is constant, as shown by:

$$E' = T' + V' = 0$$

Very nice all this, and by getting back now to the pendulum from Definition 3.1, we can have this understood with not many computations involved, as follows:

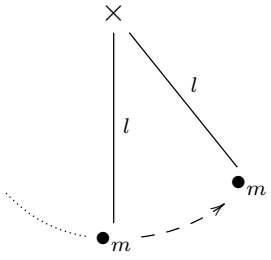
THEOREM 3.2. *For a pendulum starting with speed v from the equilibrium position,*



the motion will be confined if $v^2 < 4gl$, and circular if $v^2 > 4gl$.

PROOF. There are many ways of proving this result, along with working out several other useful related formulae, for which we will refer to the proof below, and with a quite elegant approach to this, using no computations or almost, being as follows:

(1) Let us first examine what happens when the bob has traveled an angular distance $\theta > 0$, with respect to the vertical. The picture here is as follows:



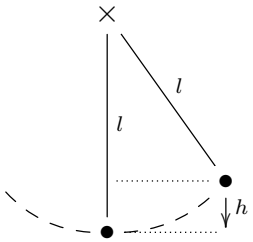
The distance traveled is then $x = l\theta$. As for the force acting, this is $F_{total} = mg$ oriented downwards, with the component alongside x being given by:

$$\begin{aligned} F &= -||F_{total}|| \sin\theta \\ &= -mg\sin\theta \\ &= -mg\sin\left(\frac{x}{l}\right) \end{aligned}$$

(2) But with this, we can compute the potential energy. With the convention that this vanishes at the equilibrium position, $V(0) = 0$, we obtain the following formula:

$$\begin{aligned} V' = -F &\implies V' = mg \sin\left(\frac{x}{l}\right) \\ &\implies V = mgl\left(1 - \cos\left(\frac{x}{l}\right)\right) \\ &\implies V = mgl(1 - \cos\theta) \end{aligned}$$

(3) Alternatively, in case this sounds too wizarding, we can compute the potential energy in the old fashion, by letting the bob fall, the picture being as follows:



The height of the fall is then $h = l - l\cos\theta$, and since for this fall the force is constant, $\mathcal{F} = -mg$, we obtain the following formula for the potential energy:

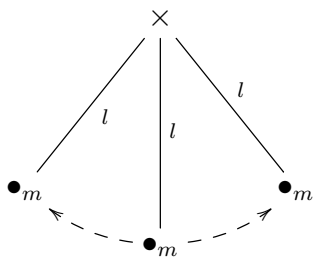
$$\begin{aligned} V' = -\mathcal{F} &\implies V' = mg \\ &\implies V = mgh \\ &\implies V = mgl(1 - \cos\theta) \end{aligned}$$

Summarizing, one way or another we have our formula for the potential energy V.

(4) Now comes the discussion. The motion will be confined when the initial kinetic energy, namely $E = mv^2/2$, satisfies the following condition:

$$\begin{aligned} E < \sup_\theta V = 2mgl &\iff \frac{mv^2}{2} < 2mgl \\ &\iff v^2 < 4gl \end{aligned}$$

In this case, the motion will be confined between two angles $-\theta, \theta$, as follows:



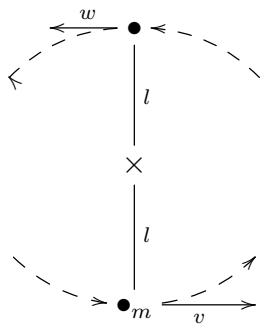
To be more precise here, the two extreme angles $-\theta, \theta \in (-\pi, \pi)$ can be explicitly computed, as being solutions of the following equation:

$$\begin{aligned} V = E &\iff mgl(1-\cos\theta) = \frac{mv^2}{2} \\ &\iff 1-\cos\theta = \frac{v^2}{2gl} \end{aligned}$$

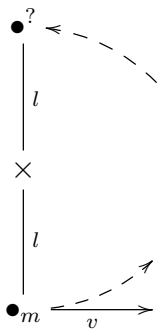
(5) Regarding now the case $v^2 > 4gl$, here the bob will certainly reach the upwards position, with the speed $w > 0$ there being given by the following formula:

$$\begin{aligned} \frac{mw^2}{2} = E - 2mgl &\implies \frac{mw^2}{2} = \frac{mv^2}{2} - 2mgl \\ &\implies w^2 = v^2 - 4gl \\ &\implies w = \sqrt{v^2 - 4gl} \end{aligned}$$

Thus, with the convention in the statement for v, that is, going to the right, the motion of the pendulum will be counterclockwise circular, and perpetual:



(6) Finally, in the case $v^2 = 4gl$, the bob will also reach the upwards position, but with speed $w = 0$ there, and then, at least theoretically, will remain there:



(7) Actually, it is quite interesting in this latter situation, $v^2 = 4gl$, to further speculate on what can happen, when making our problem more realistic. For instance, we can add to our setting the assumption that when the bob is stuck on top, with speed 0, there is a

1/3 chance for it to keep going, to the left, a 1/3 chance for it to come back, to the right, and a 1/3 chance for it to remain stuck. In this case there are infinitely many possible trajectories, which are best investigated by using probability. Welcome to chaos. □

Let us discuss now the motion of a particle near an equilibrium point. We have two basic examples of such points provided by the pendulum, namely the downwards one, which is stable, and the upwards one, which is unstable. But our discussion here will be valid for any other types of particles moving, under the influence of forces.

As a first observation, our generalities about motion and energy provide us with:

THEOREM 3.3. *For a particle moving near an equilibrium point $x = 0$, the following equivalent conditions must be satisfied, infinitesimally:*

(1) *The potential energy is $V = kx^2/2$, when assuming $V(0) = 0$.*
(2) *The force acting on our particle is $F = -kx$.*
(3) *The equation of motion is $m\ddot{x} + kx = 0$, with m being the mass.*

PROOF. This is something very standard, the idea being as follows:

(1) Let us start with some generalities regarding the potential energy V. Around any given point, that we can choose by translation to be $x = 0$, we can write:

$$V(x) = V(0) + V'(0)x + \frac{V''(0)x^2}{2} + \frac{V'''(0)x^3}{6} + \ldots$$

By definition of V, we can assume $V(0) = 0$. Regarding now the second term, this vanishes too, because our condition of equilibrium reads:

$$V'(0) = -F(0) = 0$$

Thus, with the above normalizations $x = 0$ and $V(0) = 0$ made, our general formula above for V takes at equilibrium the following form, with $k = V''(0)$:

$$V(x) = \frac{kx^2}{2} + \ldots$$

Thus, we are led to the conclusion in the statement, provided that we are indeed in the non-degenerate case, where $k \neq 0$, which amounts in saying that $F'(0) \neq 0$.

(2) This follows indeed from (1), and from $V' = -F$.

(3) This follows indeed from (2), and from $F = ma = m\ddot{x}$. □

The above result suggests formulating the following definition:

DEFINITION 3.4. *A harmonic oscillator is a particle moving as above, following*

$$m\ddot{x} + kx = 0$$

with $k \neq 0$. In the case $k > 0$, we say that we have a simple harmonic oscillator.

Here the last convention comes from the fact that our oscillator is unstable when $k < 0$, and stable $k > 0$, and it is in this latter case that we are mostly interested in. And with this, stability depending on the sign of k, coming either from some abstract reasoning along the lines of Theorem 3.3, or from the explicit formulae below.

Very nice, so let us solve now the equation of motion. We have here:

THEOREM 3.5. *The solutions of the equation of motion $m\ddot{x} + kx = 0$ for the harmonic oscillators are as follows:*

(1) $x = ae^{pt} + be^{-pt}$ *with* $p = \sqrt{-k/m}$, *when* $k < 0$.
(2) $x = c\cos wt + d\sin wt$ *with* $w = \sqrt{k/m}$, *when* $k > 0$.

PROOF. This is standard mathematics, as follows:

(1) Assume first that we are in the case $k < 0$. Here, with $p = \sqrt{-k/m}$ as in the statement, the equation of motion takes the following form:

$$\ddot{x} = p^2 x$$

But the functions e^{pt}, e^{-pt} being solutions of this equation, by linearity we obtain that the solutions are exactly the linear combinations of these two functions, as claimed.

(2) Assume now that we are in the case $k > 0$. Here, with $w = \sqrt{k/m}$ as in the statement, the equation of motion takes the following form:

$$\ddot{x} = -w^2 x$$

But the functions $\cos wt, \sin wt$ being solutions, by linearity we obtain that the solutions are exactly the linear combinations of these two functions, as claimed. □

Many other things can be said, along these lines. We will be back to this.

3b. Kepler and Newton

Getting now to Kepler and Newton, their famous theorem is as follows:

THEOREM 3.6. *The following happen:*

(1) *Planets and other celestial bodies move around the Sun on conics, that is, curves given by $P(x, y) = 0$, with $P \in \mathbb{R}[x, y]$ being of degree 2.*
(2) *The conics are the curves which appear by cutting a 2-sided cone with a plane, and can be classified into ellipses, parabolas and hyperbolas.*

PROOF. The idea here is that (1) is something tough, due to Kepler and Newton, and (2), due to the ancient Greeks, is elementary. Getting started now:

(1) According to observations and calculations performed over the centuries, since the ancient times, and first formalized by Newton, following some groundbreaking work of Kepler, the force of attraction between two bodies of masses M, m is given by:

$$||F|| = G \cdot \frac{Mm}{d^2}$$

Here d is the distance between the two bodies, and $G \simeq 6.674 \times 10^{-11}$ is a constant. Now assuming that M is fixed at $0 \in \mathbb{R}^3$, the force exterted on m positioned at $x \in \mathbb{R}^3$, regarded as a vector $F \in \mathbb{R}^3$, is given by the following formula:

$$F = -||F|| \cdot \frac{x}{||x||} = -\frac{GMm}{||x||^2} \cdot \frac{x}{||x||} = -\frac{GMmx}{||x||^3}$$

But $F = ma = m\ddot{x}$, with $a = \ddot{x}$ being the acceleration, second derivative of the position, so the equation of motion of m, assuming that M is fixed at 0, is:

$$\ddot{x} = -\frac{GMx}{||x||^3}$$

Obviously, the problem happens in 2 dimensions, and you can even find, as an exercise, a formal proof of that, based on the above equation, if your really want to. Now here the most convenient is to use standard x, y coordinates, and denote our point as $z = (x, y)$. With this change made, and by setting $K = GM$, the equation of motion becomes:

$$\ddot{z} = -\frac{Kz}{||z||^3}$$

(2) The idea now is that the problem can be solved via some calculus. Let us write indeed our vector $z = (x, y)$ in polar coordinates, as follows:

$$x = r\cos\theta \quad , \quad y = r\sin\theta$$

We have then $||z|| = r$, and our equation of motion becomes:

$$\ddot{z} = -\frac{Kz}{r^3}$$

Let us differentiate now x, y. By using the standard calculus rules, we have:

$$\dot{x} = \dot{r}\cos\theta - r\sin\theta \cdot \dot{\theta} \quad , \quad \dot{y} = \dot{r}\sin\theta + r\cos\theta \cdot \dot{\theta}$$

Differentiating one more time gives the following formulae:

$$\ddot{x} = \ddot{r}\cos\theta - 2\dot{r}\sin\theta \cdot \dot{\theta} - r\cos\theta \cdot \dot{\theta}^2 - r\sin\theta \cdot \ddot{\theta}$$

$$\ddot{y} = \ddot{r}\sin\theta + 2\dot{r}\cos\theta \cdot \dot{\theta} - r\sin\theta \cdot \dot{\theta}^2 + r\cos\theta \cdot \ddot{\theta}$$

Consider now the following two quantities, appearing as coefficients in the above:

$$a = \ddot{r} - r\dot{\theta}^2 \quad , \quad b = 2\dot{r}\dot{\theta} + r\ddot{\theta}$$

In terms of these quantities, our second derivative formulae read:

$$\ddot{x} = a\cos\theta - b\sin\theta \quad , \quad \ddot{y} = a\sin\theta + b\cos\theta$$

(3) We can now solve the equation of motion from (1). Indeed, with the formulae that we found for $\ddot{x}, \ddot{y}$, our equation of motion takes the following form:

$$a\cos\theta - b\sin\theta = -\frac{K}{r^2}\cos\theta \quad , \quad a\sin\theta + b\cos\theta = -\frac{K}{r^2}\sin\theta$$

But these two formulae can be written in the following way:

$$\left(a + \frac{K}{r^2}\right)\cos\theta = b\sin\theta \quad , \quad \left(a + \frac{K}{r^2}\right)\sin\theta = -b\cos\theta$$

By making now the product, and assuming that we are in a non-degenerate case, where the angle θ varies indeed, we obtain by positivity that we must have:

$$a + \frac{K}{r^2} = b = 0$$

(4) Let us first examine the second equation, $b = 0$. This can be solved as follows:

$$\begin{aligned} b = 0 &\iff 2\dot{r}\dot{\theta} + r\ddot{\theta} = 0 \\ &\iff \frac{\ddot{\theta}}{\dot{\theta}} = -2\frac{\dot{r}}{r} \\ &\iff (\log\dot{\theta})' = (-2\log r)' \\ &\iff \log\dot{\theta} = -2\log r + c \\ &\iff \dot{\theta} = \frac{\lambda}{r^2} \end{aligned}$$

As for the first equation the we found, namely $a + K/r^2 = 0$, this becomes:

$$\ddot{r} - \frac{\lambda^2}{r^3} + \frac{K}{r^2} = 0$$

As a conclusion to all this, in polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$, our equations of motion are as follows, with λ being a constant, not depending on t:

$$\ddot{r} = \frac{\lambda^2}{r^3} - \frac{K}{r^2} \quad , \quad \dot{\theta} = \frac{\lambda}{r^2}$$

Even better now, by writing $K = \lambda^2/c$, these equations read:

$$\ddot{r} = \frac{\lambda^2}{r^2}\left(\frac{1}{r} - \frac{1}{c}\right) \quad , \quad \dot{\theta} = \frac{\lambda}{r^2}$$

(5) In order to study the first equation, we use a trick. Let us write:

$$r(t) = \frac{1}{f(\theta(t))}$$

With the convention that dots mean as usual derivatives with respect to t, and that the primes will denote derivatives with respect to $\theta = \theta(t)$, we have:

$$\dot{r} = -\frac{f'\dot{\theta}}{f^2} = -\frac{f'}{f^2} \cdot \frac{\lambda}{r^2} = -\lambda f'$$

By differentiating one more time with respect to t, we obtain:

$$\ddot{r} = -\lambda f''\dot{\theta} = -\lambda f'' \cdot \frac{\lambda}{r^2} = -\frac{\lambda^2}{r^2} f''$$

On the other hand, our equation for $\ddot{r}$ found in (4) above reads:

$$\ddot{r} = \frac{\lambda^2}{r^2}\left(\frac{1}{r} - \frac{1}{c}\right) = \frac{\lambda^2}{r^2}\left(f - \frac{1}{c}\right)$$

Thus, in terms of $f = 1/r$ as above, our equation for $\ddot{r}$ simply reads:

$$f'' + f = \frac{1}{c}$$

But this latter equation is elementary to solve. Indeed, both functions $\cos t, \sin t$ satisfy $g" + g = 0$, so any linear combination of them satisfies as well this equation. But the solutions of $f'' + f = 1/c$ being those of $g'' + g = 0$ shifted by $1/c$, we obtain:

$$f = \frac{1 + \varepsilon\cos\theta + \delta\sin\theta}{c}$$

Now by inverting, we obtain the following formula:

$$r = \frac{c}{1 + \varepsilon\cos\theta + \delta\sin\theta}$$

(6) But this leads to the conclusion that the trajectory is a conic. Indeed, in terms of the parameter θ, the formulae of the coordinates are:

$$x = \frac{c\cos\theta}{1 + \varepsilon\cos\theta + \delta\sin\theta} \quad , \quad y = \frac{c\sin\theta}{1 + \varepsilon\cos\theta + \delta\sin\theta}$$

Now observe that these two functions x, y satisfy the following formula:

$$\begin{aligned} x^2 + y^2 &= \frac{c^2(\cos^2\theta + \sin^2\theta)}{(1 + \varepsilon\cos\theta + \delta\sin\theta)^2} \\ &= \frac{c^2}{(1 + \varepsilon\cos\theta + \delta\sin\theta)^2} \end{aligned}$$

On the other hand, these two functions satisfy as well the following formula:

$$\begin{aligned} (\varepsilon x + \delta y - c)^2 &= \frac{c^2\big(\varepsilon\cos\theta + \delta\sin\theta - (1 + \varepsilon\cos\theta + \delta\sin\theta)\big)^2}{(1 + \varepsilon\cos\theta + \delta\sin\theta)^2} \\ &= \frac{c^2}{(1 + \varepsilon\cos\theta + \delta\sin\theta)^2} \end{aligned}$$

We conclude that our coordinates x, y satisfy the following equation:

$$x^2 + y^2 = (\varepsilon x + \delta y - c)^2$$

But what we have here is an equation of a conic, and this ends the proof of the first assertion of the theorem. Which is not bad at all, because you probably know now more classical mechanics than the average nerd. Just ask around, you will be surprised.

(7) Before getting into the mathematics of conics, a bit more physics. Astronomy and Kepler tell us that for planets the trajectory should be an ellipsis, and this can be deduced from what we have, the missing piece of math, which is elementary, being that of proving that a bounded non-degenerate conic must be an ellipsis. However, this will follow as well from the classification results below, so we will stop physics here.

(8) The classification of the conics, going back to the ancient Greeks, is standard. Consider indeed one of these conics:

$$C = \left\{(x, y) \in \mathbb{R}^2 \middle| P(x, y) = 0, \deg P \leq 2\right\}$$

By doing some suitable manipulations on the degree 2 polynomial $P \in \mathbb{R}[x, y]$, up to affine transformations of the curve, we can have this curve written in some simple, "standard" form, with standard depending a bit on you, matter of taste. But this standard form can only lead to the 3 cases in the statement, namely ellipses, parabolas and hyperbolas, up to degeneration, with the degenerate cases being the lines, double lines, points, empty set, and $\mathbb{R}^2$ itself, basically appearing when $\deg P \leq 1$.

(9) The fact that the conics appear by cutting a 2-sided cone with a plane is also elementary, and also known since the ancient Greeks. A first proof is by doing some abstract algebra, and verifying that the cut must be indeed a curve of degree 2. A second proof is by computing the cut in the various cases that might appear, depending on the angle of the plane with respect to the cone, with this leading to the curves found in (8), namely ellipses, parabolas and hyperbolas, up to degeneration.

(10) Summarizing, we are done with what was announced in the theorem, but there is still some discussion to be made, in relation with degeneration. With respect to what was found in (8), when cutting cones with a plane, if you want to get exactly the same list of conics, you have to allow the degenerate cone, of angle $180°$, which in practice means a plane, in order to have as examples the empty set, and $\mathbb{R}^2$ itself.

(11) Also in relation to what was found in (8), what comes out of gravity basically agrees, namely ellipses, parabolas and hyperbolas. However, at the level of degenerate examples, these are different, consisting of the point, and then of the segment, which corresponds to the object m falling into the object M on a perfectly straight line.

(12) Finally, there is as well a discussion concerning normalization, because in the Kepler problem we assumed M to be fixed at 0. However, when changing coordinates via a translation, we can obtain in this way all the ellipses, parabolas and hyperbolas. □

Still with me, I hope, after all these computations. For further applications, here is a sort of "best of" the formulae found in the proof of Theorem 3.6:

THEOREM 3.7. *In the context of a 2-body problem, with M fixed at 0, and m starting its movement from Ox, the equation of motion of m, namely*

$$\ddot{z} = -\frac{Kz}{||z||^3}$$

with $K = GM$, and $z = (x, y)$, becomes in polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$,

$$\ddot{r} = \frac{\lambda^2}{r^2}\left(\frac{1}{r} - \frac{1}{c}\right) \quad , \quad \dot{\theta} = \frac{\lambda}{r^2}$$

for some $\lambda, c \in \mathbb{R}$, related by $\lambda^2 = Kc$. The value of r in terms of θ is given by

$$r = \frac{c}{1 + \varepsilon\cos\theta + \delta\sin\theta}$$

for some $\varepsilon, \delta \in \mathbb{R}$. At the level of the affine coordinates x, y, this means

$$x = \frac{c\cos\theta}{1 + \varepsilon\cos\theta + \delta\sin\theta} \quad , \quad y = \frac{c\sin\theta}{1 + \varepsilon\cos\theta + \delta\sin\theta}$$

with $\theta = \theta(t)$ being subject to $\dot{\theta} = \lambda^2/r$, as above. Finally, we have

$$x^2 + y^2 = (\varepsilon x + \delta y - c)^2$$

which is a degree 2 equation, and so the resulting trajectory is a conic.

PROOF. As already mentioned, this is a sort of "best of" the formulae found in the proof of Theorem 3.6. And in the hope of course that we have not forgotten anything. Finally, let us mention that the simplest illustration for this is the circular motion, and for details on this, not included in the above, we refer to the proof of Theorem 3.6. □

As a last question, we would like to understand how the various parameters appearing above, namely $\lambda, c, \varepsilon, \delta$, which via some basic math can only tell us more about the shape of the orbit, appear from the initial data. The formulae here are as follows:

PROPOSITION 3.8. *In the context of Theorem 3.7, and in polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$, the initial data is as follows, with $R = r_0$:*

$$r_0 = \frac{c}{1 + \varepsilon} \quad , \quad \theta_0 = 0$$

$$\dot{r}_0 = -\frac{\delta\sqrt{K}}{\sqrt{c}} \quad , \quad \dot{\theta}_0 = \frac{\sqrt{Kc}}{R^2}$$

$$\ddot{r}_0 = \frac{\varepsilon K}{R^2} \quad , \quad \ddot{\theta}_0 = \frac{4\delta K}{R^2}$$

The corresponding formulae for the affine coordinates x, y can be deduced from this. Also, the various motion parameters c, ε, δ and $\lambda = \sqrt{Kc}$ can be recovered from this data.

PROOF. This is something quite routine, and we will leave this as an instructive exercise. In fact, things are still not over with this, because in practice, you still have to do some reverse engineering at the level of parameters, and work out how various inital speeds and accelerations lead to various types of conics. The computations here are very interesting, and we will leave them as well as an instructive exercise. □

3c. N body systems

The discretization problem for gravity is quite tricky, because we have in fact to fight with two questions, none of them being trivial. The first question is as follows:

PROBLEM 3.9. *Avoiding the void.*

To be more precise, this is certainly something that we need, for doing any sort of statistical mechanics, coming even before discretizing our problems. That is, void is certainly something unwanted in statistical mechanics.

As a first idea now, in order to deal with this question, we have the well-known "gravitational well" interpretation of the gravitational field, for instance with the 2D gravity field of a system of N bodies being interpreted as corresponding to some sort of sheet, say metal sheet in 3D, with quadratic infinite wells at the location of each body, on which a supplementary particle m will have to travel.

Of course, this interpretation is still something a bit gravitational, because the supplementary particle m will move on this metal sheet by obeying to some sort of gravity pulling it downwards. But that latter, unique gravity which acts, once the metal sheet is suitably built, is in fact of the simplest form, namely uniform.

At the level of concrete questions to be solved, the first problem is that of recovering the Kepler circular trajectories via this model, and then recovering the general Kepler elliptic trajectories, that we found before via equations, via this model.

(2) Discretizing. This is something quite tricky, with all sorts of approaches being possible, and with many such approaches pointing towards the efforts in quantum gravity, where everything is discretized anyway, at the Planck scale.

Note however that these latter theories are relativistic, and to start with, we would like to have something non-relativistic. Ideally, we would like to use a suitable modification of the balls and springs model for the mechanical waves, discussed in chapter 1, but the problem is how exactly gravity acts, inside such balls and springs models.

What we discussed in the above, in relation with gravity, was of course nice, but remains quite theoretical. At a more advanced level, the problem is that of doing some "random classical mechanics", with applications, starting from such ideas.

There are many things that can be done here, starting with topics like billiards, and chaos, which have been heavily studied, in the context of classical mechanics and dynamical systems. Many interesting problems appear as well in the connection with the N-body problem, for solid, non-pointwise bodies, with the aim of computing the probability of collisions, based on our real-life knowledge, that this probability is very close to 0.

However, if there is one very interesting thing in connection with all this, random classical mechanics, this is the organization of matter in the universe, in the form of solar systems, then galaxies, then clusters of galaxies, and so on. Obviously, some form of "clustering" is going on here, at various scales, which needs to be understood.

In order to discuss this, we will need a more advanced approach to gravity, due to Lagrange and Hamilton. Let us start with the following basic fact:

THEOREM 3.10. *Given a conservative force F, appearing as follows, with V being uniquely determined up to an additive constant,*

$$F = -\nabla V$$

the movements of a particle under F preserve the total energy, given by

$$E = T + V$$

with $T = m||v||^2/2$ being the kinetic energy, and with V being called potential energy.

PROOF. This is something very standard, which for gravity can be established via a computation using the chain rule for derivatives, the idea being as follows:

$$\dot{T} =< F, v >= - < \nabla V, v >= -\dot{V}$$

More generally, all this is known to hold for any conservative force, as stated. □

With the above in hand, we can now reformulate the classical mechanics formalism, in a far more efficient way, following Lagrange and Hamilton. Let us begin with some mathematics. We have the following theorem, due to Euler and Lagrange:

THEOREM 3.11. *Given a function $f : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$, the integral*

$$I = \int_{x_0}^{x_1} f(u, \dot{u}) dx$$

is stationary, in the sense that it is left unchanged by small variations of $u = u(x)$, which vanish at the endpoints x_0, x_1, precisely when $u = u(x)$ satisfies the equations

$$\frac{df}{du_i} = \frac{d}{dx}\left(\frac{df}{d\dot{u}_i}\right)$$

called Euler-Lagrange equations.

PROOF. Let us just work out the case $N = 1$, the general case being similar. Consider a small variation $\Delta u(x)$, which vanishes at the endpoints x_0, x_1, as required above:

$$\Delta u(x_0) = \Delta u(x_1) = 0$$

The corresponding variation of $f(u, \dot{u})$, at first order, is then given by:

$$\Delta f = \frac{df}{du}\,\Delta u + \frac{df}{d\dot{u}}\,\Delta\dot{u}$$

Thus the corresponding variation of the integral in the statement is given by:

$$\begin{aligned}
\Delta I &= \int_{x_0}^{x_1} \frac{df}{du}\,\Delta u\,dx + \int_{x_0}^{x_1} \frac{df}{d\dot{u}}\,\Delta\dot{u}\,dx \\
&= \int_{x_0}^{x_1} \frac{df}{du}\,\Delta u\,dx + \int_{x_0}^{x_1} \frac{df}{d\dot{u}} \cdot \frac{d(\Delta u)}{dx}\,dx \\
&= \int_{x_0}^{x_1} \frac{df}{du}\,\Delta u\,dx + \left[\frac{df}{d\dot{u}}\,\Delta u\right]_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx}\left(\frac{df}{d\dot{u}}\right)\Delta u\,dx \\
&= \int_{x_0}^{x_1} \frac{df}{du}\,\Delta u\,dx - \int_{x_0}^{x_1} \frac{d}{dx}\left(\frac{df}{d\dot{u}}\right)\Delta u\,dx \\
&= \int_{x_0}^{x_1} \left(\frac{df}{du} - \frac{d}{dx}\left(\frac{df}{d\dot{u}}\right)\right)\Delta u\,dx
\end{aligned}$$

We conclude that I is stationary precisely when the following equation is satisfied:

$$\frac{df}{du} = \frac{d}{dx}\left(\frac{df}{d\dot{u}}\right)$$

But this is the Euler-Lagrange equation in the statement, as desired. □

The point now with the above is that, when looking at the usual motion equations of mechanics, but written in a somewhat bizarre way, we will get precisely into the Euler-Lagrange equations. To be more precise, let us first formulate:

DEFINITION 3.12. *In the context of a conservative force F acting, the quantity*

$$L = T - V$$

is called Lagrangian of the system.

This is something quite tricky. In relation now with Theorem 3.11, the connection is very simple, called Hamilton principle, as follows:

THEOREM 3.13. *The following integral, called action integral, is stationary,*

$$I = \int_{t_0}^{t_1} L\,dt$$

and the corresponding Euler-Lagrange equations are precisely the equations of motion.

PROOF. According to Definition 3.12, the Lagrangian is given by the following formula, with V being the potential associated to our conservative force, via $F = -\nabla V$:

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z)$$

Thus, we are in the general framework of Theorem 3.11, with the function u there being played by the coordinates x, y, z. Now let us pick one of these coordinates, $s = x, y, z$, and compute the derivatives of L with respect to $s, \dot{s}$. By using $F = -\nabla V$ we have:

$$\frac{dL}{ds} = -\frac{dV}{ds} = F_s$$

Also since the potential V is time-independent, we have:

$$\frac{dL}{d\dot{s}} = m\dot{s} = ma_s$$

Now consider the equation of motion, under the influence of the force F:

$$F = ma$$

This is a vector equation, with 3 components, and according to the above formulae its 3 components can be written as follows, in terms of the Lagrangian L:

$$\frac{dL}{ds} = \frac{d}{dt}\left(\frac{dL}{d\dot{s}}\right)$$

But these are precisely the Euler-Lagrange equations for the stationarity of the action integral $I = \int L$, and we are therefore led to the conclusions in the statement. □

The point now with the above result is that it leads right away into another result, which this time is something fundamental and powerful, as we will soon discover:

THEOREM 3.14. *The Euler-Lagrange equations for the action integral*

$$I = \int_{t_0}^{t_1} L\, dt$$

hold in any system of coordinates (q_1, q_2, q_3), and are as follows:

$$\frac{dL}{dq} = \frac{d}{dt}\left(\frac{dL}{d\dot{q}}\right)$$

These latter equations are called the Lagrange equations of motion.

PROOF. We know from Theorem 3.13 that the action integral is stationary, with respect to the standard coordinates (x, y, z). But this shows that the action integral is stationary with respect to any system of coordinates (q_1, q_2, q_3), and so the corresponding Euler-Lagrange equations, which are the equations in the statement, hold indeed. □

All this might seem a bit complicated, but we will see examples in what follows. The idea every time will be the same, namely thinking a bit, than picking up a suitable system of coordinates (q_1, q_2, q_3) for our problem, and then instead of doing all sorts of computations for reformulating the equations of motion in terms of these coordinates (q_1, q_2, q_3), simply writing the Lagrange equations, which are there for that.

Moving ahead now, let us discuss some further interesting manipulations on the Lagrangian and the Lagrange equations, due to Hamilton. Let us start with:

DEFINITION 3.15. *Given a system of coordinates $q_1, \ldots, q_n$, the quantities*

$$p_i = \frac{dL}{d\dot{q}_i}$$

are called generalized momenta. In terms of them, the Lagrange equations read

$$\frac{dL}{dq_i} = \dot{p}_i$$

analogously to the usual motion equations $F = \dot{p}$.

What are these new variables good for? Let us recall from Definition 3.12 that the Lagrangian was by definition a function as follows:

$$L = L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n)$$

The point now, which is something quite subtle, and useful for all sorts of purposes, in classical mechanics, and especially in its versions and generalizations, is that we can get rid of the derivatives $\dot{q}_1, \ldots, \dot{q}_n$, which are variables in the Lagrange formulation of mechanics, by replacing them by the generalized momenta $p_1, \ldots, p_n$ constructed above. In order to do so we need a clever new quantity H, replacing L, and we have here:

DEFINITION 3.16. *With $q_1, \ldots, q_n$ and $p_1, \ldots, p_n$ being as above, the quantity*

$$H(q, p) = \sum_i p_i \dot{q}_i - L$$

is called Hamiltonian of the system.

As we will soon see, for many simple systems H is in fact the total energy. Before that, however, let us explain how H replaces L, as a quantity which encapsulates as well what is going on, namely the equations of motion. The result here is as follows:

THEOREM 3.17. *The Hamiltonian $H = H(q, p)$ is subject to the equations*

$$\frac{dH}{dp_i} = \dot{q}_i \quad , \quad \frac{dH}{dq_i} = -\dot{p}_i$$

called Hamilton equations, which are equivalent to the usual equations of motion.

PROOF. As a first observation, this reminds right away Theorem 3.14, namely the Lagrange formulation of mechanics, who was claiming the same type of thing. However, there are some differences. On one hand, the new variables $q_1, \ldots, q_n$ and $p_1, \ldots, p_n$ are certainly a bit more abstract than the old ones $q_1, \ldots, q_n$ and $\dot{q}_1, \ldots, \dot{q}_n$. On the other hand, the new equations look great. Regarding now the proof, everything follows from the definition of the variables p_i and from the Lagrange equations, namely:

$$p_i = \frac{dL}{d\dot{q}_i} \quad , \quad \frac{dL}{dq_i} = \dot{p}_i$$

(1) By using the definition of the variables p_i, we obtain the following formula:

$$\begin{aligned}
\frac{dH}{dp_i} &= \frac{d\left(\sum_j p_j\dot{q}_j - L\right)}{dp_i} \\
&= \dot{q}_i + \sum_j p_j \frac{d\dot{q}_j}{dp_i} - \sum_j \frac{dL}{d\dot{q}_j} \cdot \frac{d\dot{q}_j}{dp_i} \\
&= \dot{q}_i + \sum_j p_j \frac{d\dot{q}_j}{dp_i} - \sum_j p_j \frac{d\dot{q}_j}{dp_i} \\
&= \dot{q}_i
\end{aligned}$$

(2) By using the Lagrange equations, we obtain the following formula:

$$\begin{aligned}
\frac{dH}{dq_i} &= \frac{d\left(\sum_j p_j\dot{q}_j - L\right)}{dq_i} \\
&= -\frac{dL}{dq_i} + \sum_j p_j \frac{d\dot{q}_j}{dq_i} - \sum_j \frac{dL}{d\dot{q}_j} \cdot \frac{d\dot{q}_j}{dq_i} \\
&= -\dot{p}_i + \sum_j p_j \frac{d\dot{q}_j}{dq_i} - \sum_j p_j \frac{d\dot{q}_j}{dq_i} \\
&= -\dot{p}_i
\end{aligned}$$

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

With this technology in hand, we can now talk about the N-body problem, and more specifically about this problem in the $N >> 0$ regime. There is some obvious form of "clustering" going on here, at various scales, which needs to be understood. The ideas in order to deal with this question are described by Arnold in his book [6].

Finally, along the same lines, many interesting questions appear in relation with interstellar gases and dust. But more on this later, when talking thermodynamics.

3d. A word on relativity

Let us end with a word on relativity, which will certainly complicate certain things, but will clarify some other. Following Einstein, we first have the following result:

THEOREM 3.18. *In special relativity, where all speeds are smaller than the speed of light, $v < c$, the speeds add according to the Einstein formula*

$$u +_e v = \frac{u + v}{1 + uv/c^2}$$

in 1 dimension, and according to a similar formula, a bit more complicated, in $N \geq 2$ dimensions. Time and length are subject to Lorentz dilation and contraction,

$$t \to \gamma t \quad , \quad L \to L/\gamma$$

and the spacetime frame change along Ox is given by the Lorentz transform

$$x' = \gamma(x - vt)$$

$$y' = y$$

$$z' = z$$

$$t' = \gamma(t - vx/c^2)$$

with $\gamma = 1/\sqrt{1 - v^2/c^2}$, called the Lorentz factor.

PROOF. As explained by Einstein in [29] all this basically follows from nothing, just by thinking. So, philosophy, and here is a possible account of that thinking:

(1) Let us first discuss speeds only. For this purpose, since no time and distance will be present, we will make the normalization $c = 1$. According to $v < 1$, which by the way is not an abstraction, but something observed in practice, we must have:

$$1 +_e v = 1$$

Thus, we are led into a math puzzle, namely how to define a new sum for speeds $+_e$, for the above to happen. In 1 dimension the answer to this is obvious, namely:

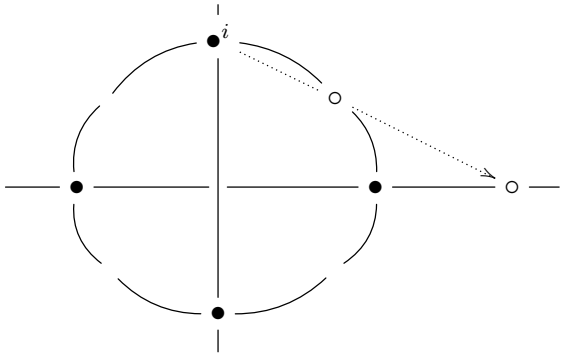
$$u +_e v = \frac{u + v}{1 + uv}$$

(2) And good news, this is the good formula, compatible with Galileo's $u +_g v = u + v$ at low speeds, $u, v \simeq 0$, and compatible as well, up to an epsilon, with the following key formula of Fizeau, for the speed of light through liquid flowing into a tube:

$$u +_f v = u + v\left(1 - \frac{1}{n^2}\right)$$

To be more precise, here light moves through liquid at speed $u < c$, with $n = c/u$ being the index of refraction, and the liquid moves through the tube at speed $v > 0$.

(3) Some mathematical comments as well. Getting back to the beginning, namely our math puzzle in (1), you might perhaps say that we are looking for a simple addition operation on $[-1,1]$, and that can only be the usual addition on $\mathbb{R}$, via wrapping $[-1,1]$ on the unit circle around i, and then stereographically projecting from i:



However, this is wrong, because the resulting function $\varphi : [-1,1] \to [-\infty,\infty]$, given by $\varphi(u) = \tan(\pi u/2)$, has a number of bugs, including a bad slope at the origin.

(4) This being said, never give up. After some thinking, the idea in (3) is after all not that bad, it's just that tan must be replaced by tanh. And with tanh, we have:

$$\tanh(x+y) = \frac{\tanh x + \tanh y}{1 + \tanh x \tanh y}$$

Thus, our original idea in (3) proved to be fruitful, and we ended up with the Einstein summation from (1). Of course, all might seem overly twisted, but note that with this second approach, we can also say now that we are into hyperbolic geometry. Nice.

(5) Getting now to $N \geq 2$ dimensions, assuming that you have a couple of days for playing with the math puzzle there, you will reach to the following solution:

$$u +_e v = \frac{1}{1 + < u, v >} \left(u + v + \frac{< u, v > u - < u, u > v}{1 + \sqrt{1 - ||u||^2}} \right)$$

In the case $N = 3$, and actually in the case $N = 2$ too, due to a lack of further simplifications there, the final formula is as follows, using the vector product $\times$:

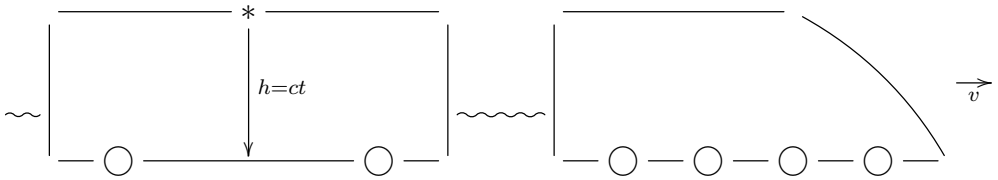
$$u +_e v = \frac{1}{1 + < u, v >} \left(u + v + \frac{u \times (u \times v)}{1 + \sqrt{1 - ||u||^2}} \right)$$

Observe that our formulae at $N \geq 2$ are not commutative, nor associative. This is something quite puzzling, but it is simply impossible to do otherwise. Just try.

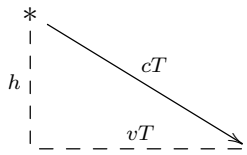
(6) Again good news, our formulae above are the good ones at $N = 2$, and consequently at $N = 3$ and higher too, this time due to their compatibility with what came out

of another experiment, further building on Fizeau's, due to Michelson and Morley. In addition, there is as well a key compatibility with electromagnetism and the Maxwell equations, via an invariance result of Lorentz, to be explained in a moment. For more on all this, physics supporting the above formula of $+_e$, we refer to Einstein [29].

(7) Getting now to what happens to distance and time, in all this, according to the formula $v = d/t$, it is convenient to waive the $c = 1$ convention, and get back to SI units. So, assume that we have a train, moving through vacuum with speed v. The passenger onboard switches on the ceiling light bulb, measures the time t that the light needs to hit the floor, by traveling at speed c, and concludes that the train height is $h = ct$:



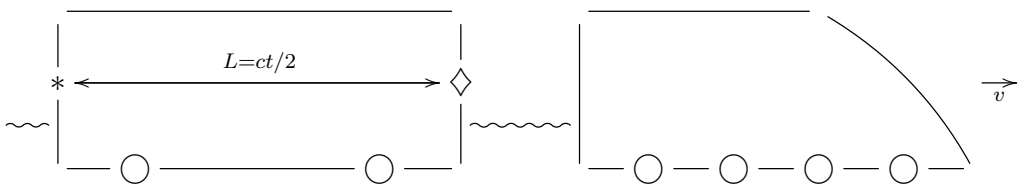
(8) On the other hand, an observer on the ground will see here something different, namely a right triangle, with on the vertical the height of the train h, on the horizontal the distance vT that the train has traveled, and on the hypotenuse the distance cT that light has travelled, with T being the duration of the event, according to his watch:



But this leads, via Pythagoras, to the Lorentz time dilation formula, namely:

$$T = \frac{h}{\sqrt{c^2 - v^2}} = \frac{ct}{\sqrt{c^2 - v^2}} = \frac{t}{\sqrt{1 - v^2/c^2}} = \gamma t$$

(9) In order to deal now with length, in the same train running at speed v, the passenger switches on the light bulb, now at the rear of the car, and measures the time t needed for the light to reach the front of the car, and get reflected back by a mirror installed there, as to conclude that the length of the car is $L = ct/2$:



(10) Viewed from the ground, the duration of the event is $T = T_1 + T_2$, where $T_1 > T_2$ are respectively the time needed for the light to travel forward, among others for beating

v, and the time for the light to travel back, helped this time by v. More precisely, if l denotes the length of the train car viewed from the ground, the formula of T is:

$$\begin{aligned} T &= T_1 + T_2 \\ &= \frac{l}{c-v} + \frac{l}{c+v} \\ &= \frac{2lc}{c^2-v^2} \end{aligned}$$

With this data, the formula $T = \gamma t$ of time dilation established before reads:

$$\frac{2lc}{c^2-v^2} = \gamma t = \frac{2\gamma L}{c}$$

Thus, the two lengths L and l are indeed not equal, but instead related by:

$$\begin{aligned} l &= \frac{\gamma L(c^2-v^2)}{c^2} \\ &= \gamma L\left(1 - \frac{v^2}{c^2}\right) \\ &= \frac{\gamma L}{\gamma^2} \\ &= \frac{L}{\gamma} \end{aligned}$$

(11) As a comment here, while the proof of time dilation from (7,8) was clean, the $c \pm v$ speeds in the proof of length contraction from (9,10) are not fully honest. But, this can be fixed either mathematically, by fine-tuning a bit the train experiment, via an added angle, and making this time use of the speed addition formula in 2D, from (5), or of course by arguing that length contraction is compatible with all the physics in (6).

(12) Finally, the Lorentz transform formula at the end is clear from the Lorentz dilation of time, and contraction of length. And with the comment that, shall you want that Lorentz transform formula to be true by axiom, all the rest, including the speed addition formulae in (5), follow easily from this, just by computing some derivatives. □

In relation now with our discretization questions, all this is obviously quite complicated, unless we are in 1D, where a number of things can be done, following the geometric observations from the proof of Theorem 3.18. In fact, we are led to similar difficulties to those in chapter 1, and as an interesting philosophical conclusion, we can now say that the difficulties in chapter 1 with waves were not exactly due to the spherical nature of spherical harmonics, but rather, via Lorentz, to the curved nature of spacetime.

I don't know about you, but at this point I am personally on the verge of exploding, due to too much abstract thinking. This being said, let us cool down, and move ahead with more relativity theory. Still following Einstein, we have the following key result:

THEOREM 3.19. *The theory of gravity can be suitably modified, and merged with relativity, into a theory called general relativity.*

PROOF. All this is a bit complicated, involving some geometry, as follows:

(1) As a starting point, let us go back to the formula $F = -\Delta V$, that we know well. Geometrically, as said before, this suggests looking at the gravitational field of k bodies $M_1, \ldots, M_k$ as being represented by $\mathbb{R}^3$ having k holes in it, and with the heavier the M_i, the bigger the hole, and with poor $m \simeq 0$ having to roll on all this.

(2) Of course we are here in 4D, for the full picture, that of the potential V, or rather of its graph, and in order to better understand this, it is of help to first consider the question where our bodies $M_1, \ldots, M_k$ lie in a plane $\mathbb{R}^2$.

(3) The point now is that, with this done, the passage to relativity can be understood as well, by modifying a bit the geometry there, as to fit with relativistic spacetime, and by having a $F = \dot{P}$ idea in mind too. That is the main idea behind general relativity, and in practice, all this needs a bit of technical geometry and formulae. □

Needless to say, in regards with our discretization and lattice model questions, all this becomes quite complicated. But we will be back to this, later in this book.

3e. Exercises

Exercises:

EXERCISE 3.20.

EXERCISE 3.21.

EXERCISE 3.22.

EXERCISE 3.23.

EXERCISE 3.24.

EXERCISE 3.25.

Bonus exercise.

CHAPTER 4

Light and heat

4a. Heat equation

Time to systematically talk about light and heat, which are related. We will first go ahead with more thermodynamics, by talking about the heat equation, and then we will discuss electromagnetism, light, quantum mechanics, and the work of Max Planck.

Getting started, we already know about heat from chapter 2, and we would like now to go into heavier, more powerful models and equations for heat diffusion. The general equation here is quite similar to the one for the waves, as follows:

THEOREM 4.1. *Heat diffusion in $\mathbb{R}^N$ is described by the heat equation*

$$\dot{\varphi} = \alpha\Delta\varphi$$

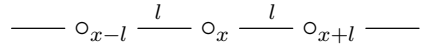
where $\alpha > 0$ is the thermal diffusivity of the medium, and Δ is the Laplace operator.

PROOF. The study here is quite similar to the study of waves, as follows:

(1) To start with, as an intuitive explanation for the equation, since the second derivative φ'' in one dimension, or the quantity $\Delta\varphi$ in general, computes the average value of a function φ around a point, minus the value of φ at that point, the heat equation as formulated above tells us that the rate of change $\dot{\varphi}$ of the temperature of the material at any given point must be proportional, with proportionality factor $\alpha > 0$, to the average difference of temperature between that given point and the surrounding material.

(2) The heat equation as formulated above is of course something approximative, and several improvements can be made to it, first by incorporating a term accounting for heat radiation, and then doing several fine-tunings, depending on the material involved. But more on this later, for the moment let us focus on the heat equation above.

(3) In relation with our modelling questions, we can recover this equation a bit as we did for the wave equation in chapter 1, by using a basic lattice model. Indeed, let us first assume, for simplifying, that we are in the one-dimensional case, $N = 1$. Here our model looks as follows, with distance $l > 0$ between neighbors:



In order to model heat diffusion, we have to implement the intuitive mechanism explained above, namely "the rate of change of the temperature of the material at any given point must be proportional, with proportionality factor $\alpha > 0$, to the average difference of temperature between that given point and the surrounding material".

(4) In practice, this leads to a condition as follows, expressing the change of the temperature φ, over a small period of time $\delta > 0$:

$$\varphi(x, t+\delta) = \varphi(x,t) + \frac{\alpha\delta}{l^2} \sum_{x \sim y} [\varphi(y,t) - \varphi(x,t)]$$

To be more precise, we have made several assumptions here, as follows:

– General heat diffusion assumption: the change of temperature at any given point x is proportional to the average over neighbors, $y \sim x$, of the differences $\varphi(y,t) - \varphi(x,t)$ between the temperatures at x, and at these neighbors y.

– Infinitesimal time and length conditions: in our model, the change of temperature at a given point x is proportional to small period of time involved, $\delta > 0$, and is inverse proportional to the square of the distance between neighbors, l^2.

(5) Regarding these latter assumptions, the one regarding the proportionality with the time elapsed $\delta > 0$ is something quite natural, physically speaking, and mathematically speaking too, because we can rewrite our equation as follows, making it clear that we have here an equation regarding the rate of change of temperature at x:

$$\frac{\varphi(x, t+\delta) - \varphi(x,t)}{\delta} = \frac{\alpha}{l^2} \sum_{x \sim y} [\varphi(y,t) - \varphi(x,t)]$$

As for the second assumption that we made above, namely inverse proportionality with l^2, this can be justified on physical grounds too, but again, perhaps the best is to do the math, which will show right away where this proportionality comes from.

(6) So, let us do the math. In the context of our 1D model the neighbors of x are the points $x \pm l$, and so the equation that we wrote above takes the following form:

$$\frac{\varphi(x, t+\delta) - \varphi(x,t)}{\delta} = \frac{\alpha}{l^2} \Big[(\varphi(x+l,t) - \varphi(x,t)) + (\varphi(x-l,t) - \varphi(x,t)) \Big]$$

Now observe that we can write this equation as follows:

$$\frac{\varphi(x, t+\delta) - \varphi(x,t)}{\delta} = \alpha \cdot \frac{\varphi(x+l,t) - 2\varphi(x,t) + \varphi(x-l,t)}{l^2}$$

(7) As it was the case with the wave equation in chapter 1, we recognize on the right the usual approximation of the second derivative, coming from calculus. Thus, when taking the continuous limit of our model, $l \to 0$, we obtain the following equation:

$$\frac{\varphi(x, t+\delta) - \varphi(x,t)}{\delta} = \alpha \cdot \varphi''(x,t)$$

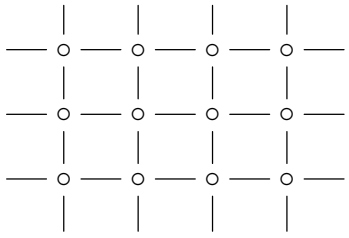
Now with $t \to 0$, we are led in this way to the heat equation, namely:

$$\dot{\varphi}(x,t) = \alpha \cdot \varphi''(x,t)$$

Summarizing, we are done with the 1D case, with our proof being quite similar to the one for the wave equation, from chapter 1.

(8) In practice now, there are of course still a few details to be discussed, in relation with all this, for instance at the end, in relation with the precise order of the limiting operations $l \to 0$ and $\delta \to 0$ to be performed, but these remain minor aspects, because our equation makes it clear, right from the beginning, that time and space are separated, and so that there is no serious issue with all this. And so, fully done with 1D.

(9) With this done, let us discuss now 2 dimensions. Here, as before for the waves, we can use a lattice model as follows, with all lengths being $l > 0$, for simplifying:



(10) We have to implement now the physical heat diffusion mechanism, namely "the rate of change of the temperature of the material at any given point must be proportional, with proportionality factor $\alpha > 0$, to the average difference of temperature between that given point and the surrounding material". In practice, this leads to a condition as follows, expressing the change of the temperature φ, over a small period of time $\delta > 0$:

$$\varphi(x,y,t+\delta) = \varphi(x,y,t) + \frac{\alpha\delta}{l^2} \sum_{(x,y)\sim(u,v)} [\varphi(u,v,t) - \varphi(x,y,t)]$$

In fact, we can rewrite our equation as follows, making it clear that we have here an equation regarding the rate of change of temperature at x:

$$\frac{\varphi(x,y,t+\delta) - \varphi(x,y,t)}{\delta} = \frac{\alpha}{l^2} \sum_{(x,y)\sim(u,v)} [\varphi(u,v,t) - \varphi(x,y,t)]$$

(11) So, let us do the math. In the context of our 2D model the neighbors of x are the points $(x \pm l, y \pm l)$, so the equation above takes the following form:

$$\begin{aligned} & \frac{\varphi(x,y,t+\delta)-\varphi(x,y,t)}{\delta} \\ = \ & \frac{\alpha}{l^2}\Big[(\varphi(x+l,y,t)-\varphi(x,y,t))+(\varphi(x-l,y,t)-\varphi(x,y,t))\Big] \\ + \ & \frac{\alpha}{l^2}\Big[(\varphi(x,y+l,t)-\varphi(x,y,t))+(\varphi(x,y-l,t)-\varphi(x,y,t))\Big] \end{aligned}$$

Now observe that we can write this equation as follows:

$$\begin{aligned} \frac{\varphi(x,y,t+\delta)-\varphi(x,y,t)}{\delta} &= \alpha\cdot\frac{\varphi(x+l,y,t)-2\varphi(x,y,t)+\varphi(x-l,y,t)}{l^2} \\ &+ \alpha\cdot\frac{\varphi(x,y+l,t)-2\varphi(x,y,t)+\varphi(x,y-l,t)}{l^2} \end{aligned}$$

(12) As it was the case when modelling the wave equation in chapter 1, we recognize on the right the usual approximation of the second derivative, coming from calculus. Thus, when taking the continuous limit of our model, $l \to 0$, we obtain the following equation:

$$\frac{\varphi(x,y,t+\delta)-\varphi(x,y,t)}{\delta} = \alpha\left(\frac{d^2\varphi}{dx^2}+\frac{d^2\varphi}{dy^2}\right)(x,y,t)$$

Now with $t \to 0$, we are led in this way to the heat equation, namely:

$$\dot{\varphi}(x,y,t) = \alpha\cdot\Delta\varphi(x,y,t)$$

Finally, in arbitrary N dimensions the same argument carries over, namely a straightforward lattice model, and gives the heat equation, as formulated in the statement. □

Observe that we can use if we want different lenghts $l > 0$ on the vertical and on the horizontal, because these will simplify anyway due to proportionality. Also, for some further mathematical fun, we can build our model on a cylinder, or a torus.

Also, as mentioned before, our heat equation above is something approximative, and several improvements can be made to it, first by incorporating a term accounting for heat radiation, and also by doing several fine-tunings, depending on the material involved. Some of these improvements can be implemented in the lattice model setting.

Regarding now the mathematics of the heat equation, many things can be said. As a first result here, often used by mathematicians, as to assume $\alpha = 1$, we have:

PROPOSITION 4.2. *Up to a time rescaling, we can assume $\alpha = 1$, as to deal with*

$$\dot{\varphi} = \Delta\varphi$$

called normalized heat equation.

PROOF. This is clear physically speaking, because according to our model, changing the parameter $\alpha > 0$ will result in accelerating or slowing the heat diffusion, in time $t > 0$. Mathematically, this follows via a change of variables, for the time variable t. □

Regarding the resolution of the heat equation, we have here:

THEOREM 4.3. *The heat equation, normalized as $\dot{\varphi} = \Delta\varphi$, and with initial condition $\varphi(x, 0) = f(x)$, has as solution the function*

$$\varphi(x, t) = (K_t * f)(x)$$

where the function $K_t : \mathbb{R}^N \to \mathbb{R}$, called heat kernel, is given by

$$K_t(x) = (4\pi t)^{-N/2} e^{-||x||^2/4t}$$

with $||x||$ being the usual norm of vectors $x \in \mathbb{R}^N$.

PROOF. According to the definition of the convolution operation $*$, we have to check that the following function satisfies $\dot{\varphi} = \Delta\varphi$, with initial condition $\varphi(x, 0) = f(x)$:

$$\varphi(x, t) = (4\pi t)^{-N/2} \int_{\mathbb{R}^N} e^{-||x-y||^2/4t} f(y) dy$$

But both checks are elementary, coming from definitions. □

In regards with our discretization questions, things are quite tricky. In relation with Theorem 4.3, and with the heat kernel, the first thought towards discretization goes to the Central Limit Theorem (CLT) from probability theory, which produces the normal laws, in dimension $N = 1$, but also in general, in arbitrary $N \geq 1$ dimensions.

In order to discuss this, let us briefly recall the CLT, and its proof. We first have:

THEOREM 4.4. *We have the following formula,*

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$$

called Gauss integral formula.

PROOF. Let I be the above integral. By using polar coordinates, we obtain:

$$\begin{aligned} I^2 &= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-x^2-y^2}dxdy \\ &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}rdrdt \\ &= 2\pi\int_0^{\infty}\left(-\frac{e^{-r^2}}{2}\right)'dr \\ &= 2\pi\left[0-\left(-\frac{1}{2}\right)\right] \\ &= \pi \end{aligned}$$

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

By using the Gauss formula we can introduce the normal laws, as follows:

DEFINITION 4.5. *The normal law of parameter* 1 *is the following measure:*

$$g_1 = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$$

More generally, the normal law of parameter $t > 0$ *is the following measure:*

$$g_t = \frac{1}{\sqrt{2\pi t}}e^{-x^2/2t}dx$$

These are also called Gaussian distributions, with "g" standing for Gauss.

Observe that the above laws have indeed mass 1, as they should. This follows indeed from the Gauss formula, which gives, with $x = \sqrt{2t}\,y$:

$$\begin{aligned} \int_{\mathbb{R}} e^{-x^2/2t}dx &= \int_{\mathbb{R}} e^{-y^2}\sqrt{2t}\,dy \\ &= \sqrt{2t}\int_{\mathbb{R}} e^{-y^2}dy \\ &= \sqrt{2\pi t} \end{aligned}$$

We will need the following key result, regarding the normal laws:

THEOREM 4.6. *With* $F_f(x) = E(e^{ixf})$ *we have the following formula:*

$$F_{g_t}(x) = e^{-tx^2/2}$$

In particular, the normal laws satisfy $g_s * g_t = g_{s+t}$*, for any* $s, t > 0$.

PROOF. The Fourier transform formula can be established as follows:

$$\begin{aligned}
F_{g_t}(x) &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-y^2/2t+ixy}dy\\
&= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-(y/\sqrt{2t}-\sqrt{t/2}ix)^2-tx^2/2}dy\\
&= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-z^2-tx^2/2}\sqrt{2t}dz\\
&= \frac{1}{\sqrt{\pi}}e^{-tx^2/2}\int_{\mathbb{R}} e^{-z^2}dz\\
&= \frac{1}{\sqrt{\pi}}e^{-tx^2/2}\cdot\sqrt{\pi}\\
&= e^{-tx^2/2}
\end{aligned}$$

As for the last assertion, this follows from the fact that $\log F_{g_t}$ is linear in t. □

We are now ready to state and prove the CLT, as follows:

THEOREM 4.7 (CLT). *Given random variables $f_1, f_2, f_3, \ldots \in L^\infty(X)$ which are i.i.d., centered, and with variance $t > 0$, we have, with $n \to \infty$, in moments,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} f_i \sim g_t$$

where g_t is the Gaussian law of parameter t, having as density $\frac{1}{\sqrt{2\pi t}}e^{-y^2/2t}dy$.

PROOF. In terms of moments, $F_f(x) = E(e^{ixf})$ is given by the following formula:

$$\begin{aligned}
F_f(x) &= E\left(\sum_{k=0}^{\infty}\frac{(ixf)^k}{k!}\right)\\
&= \sum_{k=0}^{\infty}\frac{(ix)^k E(f^k)}{k!}\\
&= \sum_{k=0}^{\infty}\frac{i^k M_k(f)}{k!}\,x^k
\end{aligned}$$

Thus, the Fourier transform of the variable in the statement is:

$$\begin{aligned} F(x) &= \left[F_f\left(\frac{x}{\sqrt{n}}\right)\right]^n \\ &= \left[1-\frac{tx^2}{2n}+O(n^{-2})\right]^n \\ &\simeq \left[1-\frac{tx^2}{2n}\right]^n \\ &\simeq e^{-tx^2/2} \end{aligned}$$

But this latter function being the Fourier transform of g_t, we obtain the result. □

With the above result in hand, complemented by its higher dimensional analogues, which follow from it, we can in principle talk about discretizing the heat kernel.

4b. Light waves

The wave equation as presented in chapter 1 belongs to classical mechanics, but an important phenomenon, with major consequences for physics in general, is that this wave equation applies as well to the waves coming from electromagnetism.

Let us start with the following standard fact, that you surely know well:

THEOREM 4.8 (Maxwell). *Electrodynamics is governed by the formulae*

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

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

called Maxwell equations.

PROOF. This is something in between a fact and a theorem, with parts of it coming from experiments, and other parts from mathematics, 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, namely the Maxwell equations in the statement. To be more precise, regarding first the math of these equations, the dots denote as usual derivatives with respect to time, and ∇ is the gradient operator, or space derivative, given by:

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

(4) 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.

(5) Finally, and of crucial importance for the considerations to follow, the constants μ_0 and ε_0 appearing in the Maxwell equations are related by the following magical formula, due to Biot-Savart, with $c = 299,792,458$ being the speed of light:

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

So, this was for the basic story of the Maxwell equations, needed in what follows. For more, there are countless good books on electrodynamics, waiting for you. □

We can now formulate a key result, in relation with the waves, as follows:

THEOREM 4.9. *In regions of space where there is no charge or current, both the electric field E and the magnetic field B are subject to the wave equation*

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

with c being as usual the speed of light.

PROOF. Under the circumstances in the statement, the Maxwell equations read:

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

$$\nabla \times E = -\dot{B}$$

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

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. □

Getting now to light, this is simply the wave predicted by Theorem 4.9, traveling at speed c, and with the important extra property that it depends on a real positive parameter, that can be called, upon taste, frequency, wavelength, or color:

FACT 4.10. *An accelerating or decelerating charge produces electromagnetic waves, traveling in vacuum at speed* $c = 299,792,458$*,*

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

and in non-vacuum at a lower speed $v < c$*. These waves are 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.

Let us try now to understand the simplest solutions of the wave equation $\ddot{\varphi} = c^2\Delta\varphi$, or more generally the wave equation $\ddot{\varphi} = v^2\Delta\varphi$, from Fact 4.10. Let us formulate:

DEFINITION 4.11. *A monochromatic plane wave is a solution of the 3D wave equation which moves in only 1 direction, that is, 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 the 1D analysis from chapter 1. And second is the assumption, in connection with the Fourier decomposition result from chapter 1, 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.

The above result opens a whole new perspective on electromagnetism, with the possibility of 3D electromagnetic waves traveling at the speed of light c appearing. We will discuss this in a moment, but in order now to finish our discussion here, let us just mention

that such electromagnetic waves exist indeed, and are very familiar objects, as follows:

Frequency	Wave type	Wavelength
	–	
10^{21}	γ rays	10^{-12}
10^{18}	$\mathrm{X}-$ rays	10^{-9}
10^{16}	UV	10^{-7}
	–	
10^{15}	blue	10^{-6}
10^{15}	yellow	10^{-6}
10^{15}	red	10^{-6}
	–	
10^{14}	IR	10^{-5}
10^{10}	microwave	10^{-1}
10^{6}	AM	10^{3}

In regards with our discretization questions, in the electrodynamics setting, the situation is quite complicated, but a few things can be done.

Of particular difficulty is the problem of talking about radiation, that is, creation of electromagnetic waves by accelerating or decelerating charges, because this would suppose to have far more electromagnetic theory discretized, than the wave equation only.

Finally, in relation with this latter question, a first basic problem is that of discretizing electrostatics, and more specifically the Coulomb force. We will discuss such questions later, first in the next chapter for the Newton gravity force, which modulo a rescaling is the same as the Coulomb force for opposite charges, at least in the context of the 2-body problem, and then later, on various occasions, in more general contexts.

Regarding now the known basic results about the electromagnetic waves, these fall under the banner "basic optics", which features questions regarding polarization, decomposition of wave packets, and especially reflection and transmission questions.

4c. Quantum physics

Getting now to some truly exciting applications of light and spectroscopy, let us discuss the beginnings of the atomic theory. There is a long story here, involving many discoveries of many people, around 1890-1900, focusing on the spectral lines of hydrogen H.

We will present here things a bit retrospectively. First on our list is the following discovery, which actually came second, by Lyman in 1906:

FACT 4.12 (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, which is invisible to the naked eye. Due to this fact, this series, while theoretically being the most important, from a modern perspective, was discovered only second, and more precisely, as mentioned, in 1906.

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

FACT 4.14 (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
$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	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 4.15 (Rydberg, Ritz). *The spectral lines of the hydrogen atom are given by the Rydberg formula, depending on integer parameters* $n_1 < n_2$*,*

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

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
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

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.

From a theoretical physics viewpoint, the main result remains the middle assertion, called Ritz-Rydberg combination principle, which is something quite puzzling. But this combination principle reminds the formula $e_{n_1n_2}e_{n_2n_3} = e_{n_1n_3}$ for the usual matrix units $e_{ij} : e_j \to e_i$. Thus, we are in familiar territory here, and we can start dreaming of:

THOUGHT 4.16. *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.*

Now back to more concrete things, 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 for all this, let us agree on:

CLAIM 4.17. *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, as a natural question, 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 it

is so came from several clever experiments. Let us first recall 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 4.17.

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 ${}_{92}\mathrm{U}$, polonium ${}_{84}\mathrm{Po}$ and radium ${}_{88}\mathrm{Ra}$. 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 4.17, 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 4.17.

So, taking now Claim 4.17 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 4.17, but led to an improved version of it, as follows:

CLAIM 4.18. *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. And with this in mind, all the pieces of our puzzle start fitting together, and we are led to the following grand conclusion:

CLAIM 4.19 (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 mathematics in all the above can lead to a beautiful axiomatization of quantum mechanics, as a "matrix mechanics", along the lines of Thought 4.16.

In order to discuss now the work of Heisenberg and Schrödinger, which eventually led to a complete proof of Claim 4.19, let us go back to the Ritz-Rydberg combination principle from Conclusion 4.15. This key combination principle was as follows:

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

As explained in the above, in Thought 4.16, all this reminds the multiplication formula $e_{n_1n_2}e_{n_2n_3}=e_{n_1n_3}$ for the elementary matrices $e_{ij}:e_j\to e_i$, and suggests, as a general principle, that observables in quantum mechanics should be some sort of infinite matrices, and multiplying between them as the matrices do, as to produce further observables.

Before getting into this, Heisenberg matrix mechanics, let us hear as well the point of view of Schrödinger, which came a few years later. His idea was to forget about exact things, and try to investigate the hydrogen atom statistically. Let us start with:

QUESTION 4.20. *In the context of the hydrogen atom, assuming that the proton is fixed, what is the probability density $\varphi_t(x)$ of the position of the electron e, at time t,*

$$P_t(e\in V)=\int_V\varphi_t(x)dx$$

as function of an intial probability density $\varphi_0(x)$? Moreover, can the corresponding equation be solved, and will this prove the Bohr claims for hydrogen, statistically?

In order to get familiar with this question, let us first look at examples coming from classical mechanics. In the context of a particle whose position at time t is given by $x_0+\gamma(t)$, the evolution of the probability density will be given by:

$$\varphi_t(x)=\varphi_0(x)+\gamma(t)$$

However, such examples are somewhat trivial, of course not in relation with the computation of γ, usually a difficult question, but in relation with our questions, and do not apply to the electron. The point indeed is that, in what regards the electron, we have:

FACT 4.21. *In respect with various simple interference experiments:*

(1) *The electron is definitely not a particle in the usual sense.*
(2) *But in most situations it behaves exactly like a wave.*
(3) *But in other situations it behaves like a particle.*

Getting back now to the Schrödinger question, all this suggests to use, as for the waves, an amplitude function $\psi_t(x) \in \mathbb{C}$, related to the density $\varphi_t(x) > 0$ by the formula $\varphi_t(x) = |\psi_t(x)|^2$. Not that a big deal, you would say, because the two are related by simple formulae as follows, with $\theta_t(x)$ being an arbitrary phase function:

$$\varphi_t(x) = |\psi_t(x)|^2 \quad , \quad \psi_t(x) = e^{i\theta_t(x)}\sqrt{\varphi_t(x)}$$

However, such manipulations can be crucial, raising for instance the possibility that the amplitude function satisfies some simple equation, while the density itself, maybe not. And this is what happens indeed. Schrödinger was led in this way to:

CLAIM 4.22 (Schrödinger). *In the context of the hydrogen atom, the amplitude function of the electron $\psi = \psi_t(x)$ is subject to the Schrödinger equation*

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

m being the mass, $h = h_0/2\pi$ the reduced Planck constant, and V the Coulomb potential of the proton. The same holds for movements of the electron under any potential V.

Observe the similarity with the wave equation $\ddot{\varphi} = v^2\Delta\varphi$, and with the heat equation $\dot{\varphi} = \alpha\Delta\varphi$ too. Many things can be said here. Following now Heisenberg and Schrödinger, and then especially Dirac, who did the axiomatization work, we have:

DEFINITION 4.23. *In quantum mechanics the states of the system are vectors of a Hilbert space H, and the observables of the system are linear operators*

$$T : H \to H$$

which can be densely defined, and are taken self-adjoint, $T = T^*$. The average value of such an observable T, evaluated on a state $\xi \in H$, is given by:

$$< T >=< T\xi, \xi >$$

In the context of the Schrödinger mechanics of the hydrogen atom, the Hilbert space is the space $H = L^2(\mathbb{R}^3)$ where the wave function ψ lives, and we have

$$< T >= \int_{\mathbb{R}^3} T(\psi) \cdot \bar{\psi}\, dx$$

which is called "sandwiching" formula, with the operators

$$x \quad , \quad -\frac{ih}{m}\nabla \quad , \quad -ih\nabla \quad , \quad -\frac{h^2\Delta}{2m} \quad , \quad -\frac{h^2\Delta}{2m} + V$$

representing the position, speed, momentum, kinetic energy, and total energy.

In other words, we are doing here two things. First, we are declaring by axiom that various "sandwiching" formulae found before by Heisenberg, involving the operators at the end, that we will not get into in detail here, hold true. And second, we are raising

the possibility for other quantum mechanical systems, more complicated, to be described as well by the mathematics of the operators on a certain Hilbert space H, as above.

4d. Max Planck

Consider a black body, that is to say, a body at thermal equilibrium, assumed to be at temperature T. This body radiates heat, and we are interested in computing the energy density of the radiation $\mathcal{E}(\nu, T)$, around a given frequency ν of this radiation. Quite surprisingly, the intuitive and honest modelling of the problem, and the subsequent math, done honestly too, lead to a spectacularly wrong result, as follows:

THEOREM 4.24. *We have the Rayleigh-Jeans formula for the energy density*

$$\mathcal{E}(\nu, T) = \frac{8\pi bT}{c^3}\,\nu^2$$

where b is the Boltzmann constant, leading globally to the divergent integral

$$\mathcal{E} = \frac{8\pi bTV}{c^3}\int_0^\infty \nu^2\,d\nu$$

over a volume V, with this divergence phenomenon being called UV catastrophe.

PROOF. This is arguably the most famous wrong result in the history of physics, so we will spend some time in trying to understand its proof. And with the comment that this will be no waste of time, because the fix, found later by Max Planck, uses exactly the same ideas and computations, but with an unexpected twist at the end.

(1) Our starting point are the equations for the electromagnetic radiation, that we will now regard as heat, as formulated before, namely:

$$E = Re(\widetilde{E}) \quad : \quad \widetilde{E} = e_n\, e^{i(<k_n,x>-w_nt)}$$

$$B = Re(\widetilde{B}) \quad : \quad \widetilde{B} = b_n\, e^{i(<k_n,x>-w_nt)}$$

Here n is a certain parameter, that will appear later on, and that we can for the moment ignore. Now inserting this data into the Maxwell equations gives the following formulae, connecting the parameters, that we will use several times in what follows:

$$k_n \times b_n + \frac{w_n}{c}\,e_n = 0$$

$$k_n \times e_n - \frac{w_n}{c}\,b_n = 0$$

$$< k_n, e_n >=< k_n, b_n >= 0$$

(2) Let us compute the electromagnetic energy in a finite volume $V = L^3$. We will use here the well-known fact, coming from classical electrodynamics, that the energy density in radiation is $(||E||^2 + ||B||^2)/8\pi$. Thus, the energy we are looking for is given by:

$$\mathcal{E} = \frac{1}{8\pi}\int_V ||E||^2 + ||B||^2$$

(3) In order to compute this integral, let us better model our question. Due to obvious periodicity reasons, the wave number k and the angular frequency w must be of the following form, with $n \in \mathbb{Z}^3$ being a vector with integer components:

$$k_n = \frac{2\pi}{L} \cdot n \quad , \quad w_n = c||k_n||$$

Thus, the electric and magnetic fields in our enclosure $V = L^3$ appear as linear combinations as follows, for certain vectors $e_n, b_n \perp n$, related by the formulae in (1):

$$E = Re(\widetilde{E}) \quad : \quad \widetilde{E} = \sum_n e_n \, e^{i(<k_n,x>-w_nt)}$$

$$B = Re(\widetilde{B}) \quad : \quad \widetilde{B} = \sum_n b_n \, e^{i(<k_n,x>-w_nt)}$$

(4) According to the above formula of E, we have:

$$\begin{aligned}
||E||^2 &= ||Re(\widetilde{E})||^2 \\
&= \frac{1}{4}\left|\left|\sum_n e_n \, e^{i(<k_n,x>-w_nt)} + \bar{e}_n \, e^{-i(<k_n,x>-w_nt)}\right|\right|^2 \\
&= \frac{1}{4}\sum_{nm} < e_n, e_m > e^{i(<k_n-k_m,x>-(w_n-w_m)t)} \\
&+ \frac{1}{4}\sum_{nm} < e_n, \bar{e}_m > e^{i(<k_n+k_m,x>-(w_n+w_m)t)} \\
&+ \frac{1}{4}\sum_{nm} < \bar{e}_n, e_m > e^{i(-<k_n+k_m,x>+(w_n+w_m)t)} \\
&+ \frac{1}{4}\sum_{nm} < \bar{e}_n, \bar{e}_m > e^{i(-<k_n-k_m,x>+(w_n-w_m)t)}
\end{aligned}$$

(5) Now by integrating, we obtain the following formula:

$$\begin{aligned}
\frac{1}{V}\int_V ||E||^2 &= \frac{1}{4}\sum_n < e_n, e_n > + \frac{1}{4}\sum_n < e_n, \bar{e}_{-n} > e^{-2iw_nt} \\
&+ \frac{1}{4}\sum_n < \bar{e}_n, e_{-n} > e^{2iw_nt} + \frac{1}{4}\sum_n < \bar{e}_n, \bar{e}_n >
\end{aligned}$$

(6) Similarly, according to the above formula of B, we have:

$$\begin{aligned}
\frac{1}{V}\int_V ||B||^2 &= \frac{1}{4}\sum_n < b_n, b_n > + \frac{1}{4}\sum_n < b_n, \bar{b}_{-n} > e^{-2iw_nt} \\
&+ \frac{1}{4}\sum_n < \bar{b}_n, b_{-n} > e^{2iw_nt} + \frac{1}{4}\sum_n < \bar{b}_n, \bar{b}_n >
\end{aligned}$$

(7) Before summing the integrals that we found, let us use the formulae connecting the parameters k_n, e_n, b_n found in (1) above, namely:

$$k_n \times b_n + \frac{w_n}{c}\, e_n = 0$$

$$k_n \times e_n - \frac{w_n}{c}\, b_n = 0$$

$$< k_n, e_n >=< k_n, b_n >= 0$$

By using these formulae, we first obtain the following identity:

$$\begin{aligned} < b_n, b_n > &= \frac{c^2}{w_n^2} < k_n \times e_n, k_n \times e_n > \\ &= \frac{c^2 ||k_n||^2}{w_n^2} < e_n, e_n > \\ &= < e_n, e_n > \end{aligned}$$

Similarly, we have we well the following identity:

$$\begin{aligned} < b_n, \bar{b}_{-n} > &= \frac{c^2}{w_n^2} < k_n \times e_n, k_{-n} \times \bar{e}_n > \\ &= -\frac{c^2 ||k_n||^2}{w_n^2} < e_n, \bar{e}_{-n} > \\ &= - < e_n, \bar{e}_{-n} > \end{aligned}$$

Also similarly, we have as well the following identity:

$$\begin{aligned} < \bar{b}_n, b_{-n} > &= \frac{c^2}{w_n^2} < k_n \times \bar{e}_n, k_{-n} \times e_n > \\ &= -\frac{c^2 ||k_n||^2}{w_n^2} < \bar{e}_n, e_{-n} > \\ &= - < \bar{e}_n, e_{-n} > \end{aligned}$$

Finally, we have as well the following identity:

$$\begin{aligned} < \bar{b}_n, \bar{b}_n > &= \frac{c^2}{w_n^2} < k_n \times \bar{e}_n, k_n \times \bar{e}_n > \\ &= \frac{c^2 ||k_n||^2}{w_n^2} < \bar{e}_n, \bar{e}_n > \\ &= < \bar{e}_n, \bar{e}_n > \end{aligned}$$

(8) We conclude that when summing the integrals computed in (5) and (6), all the terms involving phases will cancel, and we obtain the following formula:

$$\frac{1}{V} \int_V ||E||^2 + ||B||^2 = \frac{1}{2} \sum_n < e_n, e_n > + \frac{1}{2} \sum_n < \bar{e}_n, \bar{e}_n >$$

Now by multiplying everything by $V/8\pi$, as explained in (2), we obtain:

$$\mathcal{E} = \frac{V}{16\pi} \sum_n \left(< e_n, e_n > + < \bar{e}_n, \bar{e}_n > \right)$$

(9) The point now is that, by computing this sum, we are led to the Rayleigh-Jeans formula in the statement for the corresponding radiation energy density, namely:

$$\mathcal{E}(\nu, T) = \frac{8\pi bT}{c^3} \nu^2$$

(10) And this is certainly wrong, because the total energy which is radiated by our black body, all over the frequency spectrum, follows to be:

$$\mathcal{E} = \frac{8\pi bTV}{c^3} \int_0^\infty \nu^2 \, d\nu = \infty$$

More precisely, the Rayleigh-Jeans formula works quite well all across the frequency spectrum, in particular fitting well with the known data, expect for the UV range, where things diverge. And with this phenomenon being called "UV catastrophe". □

Fortunately, the solution to the UV catastrophe, and to the black body problem in general, was found a few years later by Max Plank, his result being as follows:

THEOREM 4.25. *The correct formula for the black body radiation, obtained by assuming that energy is quantized, is the Planck formula*

$$\mathcal{E}(\nu, T) \, d\nu = \frac{8\pi h}{c^3} \cdot \frac{\nu^3 d\nu}{e^{h\nu/bt} - 1}$$

with h being a new constant, called Planck constant, given by

$$h = 6.626\ 070\ 15 \times 10^{-34}$$

as per the latest SI regulations. The Planck formula fits with all known data, fits as well with the Rayleigh-Jeans formula outside the UV range, and globally leads to

$$\mathcal{E} = \int_0^\infty \mathcal{E}(\nu, T) \, d\nu = aT^4$$

with the radiation energy constant on the right being given by:

$$a = \frac{16\pi^8 b^4}{15 h^3 c^3}$$

PROOF. This is something quite technical, obtained along the lines of the proof of Theorem 4.24, by counting in a new way, by assuming that energy is quantized. □

Regarding applications, a very interesting continuation of Planck's work concerns the black body radiation of the early universe, with the microwave part of it, via a Doppler shift, still permeating the space that we live in. And with this phenomenon, called "cosmic

microwave background", being at the origin of all modern cosmology. We will be back to this later in this book, towards the end, when discussing cosmology.

4e. Exercises

Exercises:

Exercise 4.26.

Exercise 4.27.

Exercise 4.28.

Exercise 4.29.

Exercise 4.30.

Exercise 4.31.

Bonus exercise.

Part II

The Ising model

I wish I was a messenger
And all the news was good
I wish I was the full moon shining
Off a Camaro's hood

CHAPTER 5

Statistical mechanics

5a. Thermodynamics

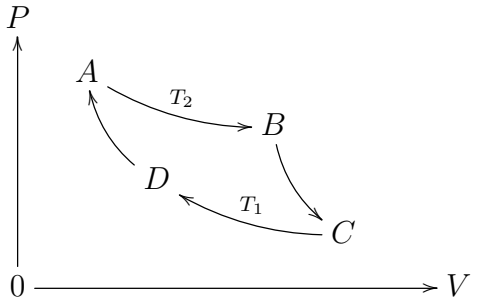
Getting to more advanced thermodynamics, we will need Carnot engines. These are somewhat mathematical objects, constructed as follows:

DEFINITION 5.1. *A Carnot engine is a 4-cycle engine, functioning between temperatures $T_1 < T_2$, consisting of:*

(1) *An isothermal expansion at T_2, absorbing heat.*
(2) *An adiabatic expansion, down to temperature T_1.*
(3) *An isothermal compression at T_1, expelling heat.*
(4) *An adiabatic compression, back to temperature T_2.*

So, what kind of engine is this. In a nutshell, these are the best possible engines, and quite fascinating objects, no one really knowing to which area of science they belong. We will see that, from the perspective of physics, such things rather belong to mathematics. From the perspective of mathematics, such things look more as engineering. And from the perspective of engineering, such things look more like theoretical physics.

In order to get familiar with such engines, let us discuss their functioning. We can represent a Carnot engine on a (P, V) diagram, as follows, with the horizontals AB, CD being the isothermals, and the verticals BC, DA being the adiabatics:



Regarding now the engineering, the Carnot engines can be implemented as 1-cylinder, 4-stroke engines, with the functioning of one full cycle being as follows:

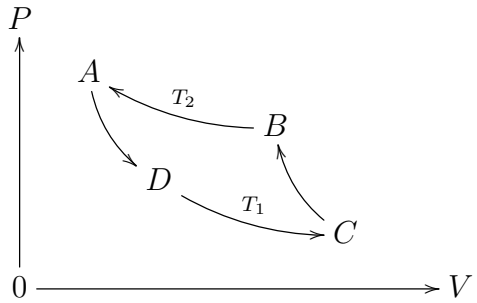
$$\begin{array}{cccccc} & A & B & C & D & A' \\ V_C & & & \perp & & \\ V_B & & \perp & & & \\ V_D & & & & \perp & \\ V_A & \perp & & & & \perp \\ & & & & & \\ & T_2 & T_2 & \equiv & T_1 & \equiv \end{array}$$

To be more precise, pictured here is the height of the piston $\perp$ at every step, with the lower data corresponding to the transformations made, T_1, T_2 standing for the temperatures of the isothermals, and $\equiv$ standing for an insulator, needed for the adiabatics.

Obviously, a Carnot engine is an engine as we know them, converting heat into work. To be more precise, an amount of heat Q_2 is absorbed during the upper isotherm AB, and is transformed into an amount of work W during the lower isotherm CD, with an amount of heat Q_1 expelled as well during this lower isotherm CD. And with the adiabatics being there for adjusting the machine, as to make it work as a cycle, in the $[T_1, T_2]$ regime.

As an interesting theoretical remark about the Carnot engines, that we will need in what follows, for theoretical purposes, but which is of practical use as well, we have:

REMARK 5.2. *When running a Carnot engine in the other sense, namely*



what we have is a refrigerator, absorbing heat Q_1 on the isotherm DC, helped by some input work W, and expelling it as heat Q_2 during the isotherm BA.

Observe here in passing that due to $Q_2 > Q_1$, using a refrigerator for cooling the room during the Summer, with the door wide open, is not a good idea.

Getting to work now, we must first talk about entropy:

THEOREM 5.3. *Given a gas, with states denoted $\sigma = (P, V, T) \in \mathbb{R}^3$, subject to the equation of state $f(\sigma) = 0$, define the entropy of a state by the formula*

$$S(\sigma) = \int_{\sigma_0}^{\sigma} \frac{dQ}{T}$$

where σ_0 is a chosen state, and the integral is over a reversible transformation from σ_0 to σ. Then S is well-defined up to an additive scalar, and we have the inequality

$$S(\sigma_1) \leq S(\sigma_2)$$

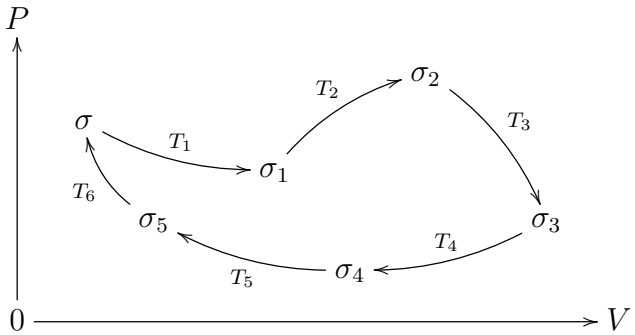
for any transformation $\sigma_1 \to \sigma_2$, with equality when this transformation is reversible.

PROOF. This comes as a continuation of basic thermodynamics, and we will use the same methods as there, namely calculus, and Carnot engines:

(1) Our claim, which will basically prove the result, via some simple mathematical reasoning afterwards, is that for any cycle $\circ : \sigma \to \sigma$ we have:

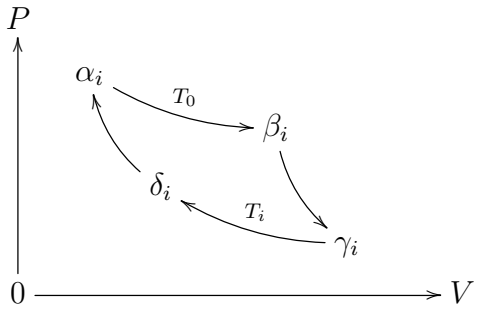
$$\int_{\circ} \frac{dQ}{T} \leq 0$$

(2) In order to prove this claim, up to some simple mathematics, we approximate the cycle $\circ$ by a sequence of isothermals $\gamma_1, \ldots, \gamma_k$, at temperatures $T_1, \ldots, T_k$:



(3) Now let us fix $T_0 > T_i$, and for any i consider a Carnot engine C_i, refrigerator or engine, which functions in the range $[T_i, T_0]$, by absorbing at T_i the positive or negative heat dQ_i emitted by our gas during γ_i. This engine is as follows, with the horizontals

$\alpha_i\beta_i, \gamma_i\delta_i$ being the isothermals, and the verticals $\beta_i\gamma_i, \delta_i\alpha_i$ being the adiabatics:



(4) Let us see now how our machinery, consisting of our gas and of the above Carnot engines $C_1, \ldots, C_k$, works, over the full cycle $\circ$:

– At each T_i nothing happens, because what is absorbed or expelled by the gas is expelled or absorbed by C_i, due to our conventions above.

– So, if there is something that our complex machinery does, that happens at the temperature $T_0 > T_i$, introduced in the above.

(5) But at T_0 the gas does nothing, and the system of our Carnot engines $C_1, \ldots, C_k$ extracts or absorbs work, of total quantity, positive or negative, given by:

$$W = T_0 \int_\circ \frac{dQ}{T}$$

(6) On the other hand, we know from the second law of thermodynamics, in Lord Kelvin's formulation, that you cannot extract work from a given temperature T_0. Thus, our system of Carnot engines $C_1, \ldots, C_k$ works overall as a refrigerator:

$$W \leq 0$$

(7) The point now is that, together with the formula in (5), this gives the following estimate, as desired, proving our claim in (1):

$$\int_\circ \frac{dQ}{T} \leq 0$$

(8) Now with this claim in (1) proved, all the rest is trivial, just by gluing paths. Indeed, by using (1), we first conclude that $S(\sigma)$ is not dependent on the choice of a reversible path $\sigma_0 \to \sigma$. Moreover, also by using (1), we conclude afterwards that $S(\sigma)$ is well-defined up to an additive scalar, independently of the choice of σ_0. And finally, again by using (1), we conclude that the equality in the statement $S(\sigma_1) \leq S(\sigma_2)$ holds indeed, with equality when the transformation $\sigma_1 \to \sigma_2$ is reversible. □

Let us do now some more mathematics for the entropy, for a gas, or whatever other system, represented on a (P, V) diagram. We know from chapter 3 that the amount of heat dQ received by the system during an infinitesimal transformation is given by:

$$dQ = \left(\frac{dE}{dT}\right)_V dT + \left[\left(\frac{dE}{dV}\right)_T + P\right] dV$$

Now since $dS = dQ/T$, we obtain from this, by dividing by T:

$$dS = \frac{1}{T}\left(\frac{dE}{dT}\right)_V dT + \frac{1}{T}\left[\left(\frac{dE}{dV}\right)_T + P\right] dV$$

As an illustration for this, let us consider an ideal gas. Here we have:

PROPOSITION 5.4. *For an ideal gas, $PV = kT$, the entropy is given by*

$$S = C_P \log T + k \log k - k \log T + a$$

with $a \in \mathbb{R}$ being a constant, whose indeterminacy comes from our definition of S.

PROOF. For an ideal gas, $PV = kT$, we have the following formula:

$$\begin{aligned} dQ &= C_V dT + P dV \\ &= C_V dT + \frac{kT}{V} dV \end{aligned}$$

By using this, we obtain as well the following formula:

$$\begin{aligned} dS &= \frac{d}{T}\left(C_V dT + \frac{kT}{V} dV\right) \\ &= \frac{C_V}{T} dT + \frac{k}{V} dV \end{aligned}$$

Now by integrating, we obtain from this a formula as follows, with $a \in \mathbb{R}$ being a constant, whose indeterminacy comes from our definition of the entropy:

$$S = C_V \log T + k \log V + a$$

Moreover, by using the formula $C_P = C_V + k$, that we know from chapter 3, and the equation of state $PV = kT$, this further reads:

$$\begin{aligned} S &= (C_P - k) \log T + k \log \frac{kT}{P} + a \\ &= (C_P - k) \log T + k(\log k + \log T - \log P) + a \\ &= C_P \log T + k \log k - k \log T + a \end{aligned}$$

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

Now back to the general case, we have here the following result:

THEOREM 5.5. *We have the following formula,*

$$\left(\frac{dE}{dV}\right)_T = T\left(\frac{dP}{dT}\right)_V - P$$

for any system represented on a (P, V) diagram.

PROOF. We recall from the above that we have:

$$dS = \frac{1}{T}\left(\frac{dE}{dT}\right)_V dT + \frac{1}{T}\left[\left(\frac{dE}{dV}\right)_T + P\right] dV$$

But this expression is a perfect differential with respect to T, V, in a multivariable calculus sense, and by taking derivatives, we obtain from this:

$$\frac{d}{dV}\left(\frac{1}{T}\cdot\frac{dE}{dT}\right) = \frac{d}{dT}\left[\frac{1}{T}\left(\frac{dE}{dV}\right) + P\right]$$

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

Following now Boltzmann, we have the following key result:

THEOREM 5.6 (Boltzmann). *The kinetic theory of gases, taking into account the collisions between molecules, leads to the formula for entropy*

$$\mathfrak{E}(S) = -b\int P(S)\log P(S)dS$$

with P being the probability on the state space, around our given state S.

PROOF. This is something quite complicated, and we refer to the literature. Observe that $\mathfrak{E}$ as computed above is given by an exact formula, not depending on an additive constant, the point here being that, in the context of Theorem 5.1, we can formally choose the state S_0 there to be the one, which is unique, at temperature $T = 0$. □

Moving ahead now, as a second key result, also due to Boltzmann, we have:

THEOREM 5.7 (Boltzmann). *Given a gas with intial molecular speed distribution P, the collisions between molecules, leading to equilibrium, will work such as the quantity*

$$H = \int P(v)\log P(v)dv$$

decreases over time. The final distribution reached, over time, which is the one at equilibrium, is precisely the one which minimizes H, given by the formula

$$P(v) = \left(\frac{m}{2\pi bT}\right)^{3/2}\exp\left(-\frac{m||v||^2}{2bT}\right)$$

which is the Maxwell-Boltzmann distribution, that we knew before from Maxwell.

PROOF. Again, this is something at the same time a bit complicated, but of utter beauty, and for the proof, we again refer to any solid statistical mechanics book. □

Observe the obvious similarity between Theorem 5.6 and Theorem 5.7, and also the important theoretical consequences of the above, in view of the comments that we previously made in connection with the Maxwell-Boltzmann distribution.

In order to further clarify all the above, we will explain now a key result of Gibbs, called H theorem. This is something quite abstract and general, as follows:

THEOREM 5.8. *For a large system, wih states denoted S, the quantity*

$$H = \int P(S) \log P(S) dS$$

can only decrease over time, as to reach a minumum, where we have

$$P(S) \sim \exp\left(\sum_i \lambda_i E_i \right)$$

where E_i are the conserved quantities, and $\lambda_i \in \mathbb{R}$ are scalars.

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

(1) In order to prove this result, let us define quantities $A(S, T)$, depending on two states S, T, close to each other, such that the rate at which the system transitions $S \to T$ equals $A(S, T)dT$. With this convention, we have the following formula:

$$\frac{dP(S)dS}{dt} = \int \Big[P(T)A(T, S)dS - P(S)A(S, T)dS \Big] dT$$

By cancelling the differentials dS, this gives the following formula:

$$\frac{dP(S)}{dt} = \int \Big[P(T)A(T, S) - P(S)A(S, T) \Big] dT$$

(2) On the other hand, consider the quantity in the statement, namely:

$$H = \int P(S) \log P(S) dS$$

By using the formula found in (1), we obtain the following formula:

$$\frac{dH}{dt} = \int \int \Big[P(T)A(T, S) - P(S)A(S, T) \Big] (\log P(S) + 1) dS dT$$

Now by interchanging S, T in the second integral, this gives:

$$\frac{dH}{dt} = \int \int P(T)A(T, S) \log \frac{P(S)}{P(T)} \, dS dT$$

(3) We use now the following standard inequality, whose proof is elementary:

$$x, y \geq 0 \implies y \log \frac{x}{y} \leq x - y$$

By applying this to the last formula found in (2), that formula gives:

$$\frac{dH}{dt} \leq \int\int (P(S) - P(T))A(T,S)dSdT$$

Now by interchanging again S, T, this inequality can be written as follows:

$$\frac{dH}{dt} \leq \int\int P(T)(A(S,T) - A(T,S))dSdT$$

(4) Normally this latter estimate allows us to conclude that H decreases, as claimed in the statement, because since the laws of physics must be invariant under time reversal, we should have $A(S,T) = A(T,S)$, and so our inequality simply reads:

$$\frac{dH}{dt} \leq 0$$

However, our argument has a flaw, in the context of advanced quantum mechanics, but as good news, the same advanced quantum mechanics tells us that we have:

$$\int (A(S,T) - A(T,S))dS = 0$$

Thus, either way, we obtain $dH/dt \leq 0$, and so that H decreases, as claimed.

(5) In order to prove now the second assertion, regarding what happens when H is minimal, observe that for an infinitesimal change $\Delta P(S)$, we have:

$$\Delta H = \int \Delta P(S)(\log P(S) + 1)dS$$

Now assuming that H is minimal, we must have $\Delta H = 0$ for any allowed infinitesimal change $\Delta P(S)$, with this meaning that the change $\Delta P(S)$ must preserve the conserved quantities of the system, as for instance the energy $E(S)$, and also be such that:

$$\int P(S)dS = 1$$

But this can only happen when $P(S)$ is of the following form, where E_i are the conserved quantities of the system, and $\lambda_i \in \mathbb{R}$ are certain scalars:

$$P(S) \sim \exp\left(\sum_i \lambda_i E_i\right)$$

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

In practice now, assuming that the energy E is the only conserved quantity, Theorem 5.8 leads to the following distribution, called "canonical ensemble":

$$P(S) = \exp\left[C - \frac{E(S)}{\theta}\right]$$

To be more precise, here θ is a certain constant, and C is another constant, uniquely determined by $\int P(S)dS = 1$. But, in practice, we have $\theta = kT$, which gives:

$$P(S) = \exp\left[C - \frac{E(S)}{kT}\right]$$

More generally now, in case our system has other conserved quantities N_i, Theorem 5.8 leads to the following distribution, called "grand canonical ensemble":

$$P(S) = \exp\left[C - \frac{E(S) - \sum_i \lambda_i N_i(S)}{kT}\right]$$

We will be back to canonical ensembles later. For the moment, let us record the following result, coming as a complement to Theorem 5.8:

THEOREM 5.9. *We have the formula*

$$\mathfrak{E}(S) = -kH$$

with $\mathfrak{E}$ being the thermodynamical entropy.

PROOF. In the context of our equations above, adding some heat ΔQ to the system gives the following formula, for the corresponding change in the quantity H:

$$\begin{aligned}\Delta H &= \int \Delta P(S)(\log P(S) + 1)dS\\ &= -\frac{1}{kT}\int \Delta P(S)E(S)dS\\ &= -\frac{\Delta Q}{kT}\end{aligned}$$

Now since the thermodynamical entropy satisfies by definition $d\mathfrak{E} = dQ/kT$, we conclude from the above equation that we have $\mathfrak{E}(S) = -kH$, as claimed. □

5b. Ferromagnetism

Ferromagnetism.

5c. Lattice models

Lattice models.

5d. Lattice gases

Lattice gases.

5e. Exercises

Exercises:

Exercise 5.10.

Exercise 5.11.

Exercise 5.12.

Exercise 5.13.

Exercise 5.14.

Exercise 5.15.

Bonus exercise.

CHAPTER 6

1D Ising model

6a. Partition function

Partition function.

6b. Transfer matrix

Transfer matrix.

6c. Critical behavior

Critical behavior.

6d. Generalizations

Generalizations.

6e. Exercises

Exercises:

Exercise 6.1.

Exercise 6.2.

Exercise 6.3.

Exercise 6.4.

Exercise 6.5.

Exercise 6.6.

Bonus exercise.

CHAPTER 7

2D Ising model

7a. Transfer matrices

Transfer matrices.

7b. Symmetry, eigenvalues

Symmetry, eigenvalues.

7c. Critical behavior

Critical behavior.

7d. Further results

Further results.

7e. Exercises

Exercises:

Exercise 7.1.

Exercise 7.2.

Exercise 7.3.

Exercise 7.4.

Exercise 7.5.

Exercise 7.6.

Bonus exercise.

CHAPTER 8

Potts model

8a. Ice-type models

Ice-type models.

8b. Eight-vertex model

Eight-vertex model.

8c. Graphs, Potts model

Graphs, Potts model.

8d. Critical exponents

Critical exponents.

8e. Exercises

Exercises:

EXERCISE 8.1.

EXERCISE 8.2.

EXERCISE 8.3.

EXERCISE 8.4.

EXERCISE 8.5.

EXERCISE 8.6.

Bonus exercise.

Part III

Algebraic aspects

Stand and fight, live by your heart
Always one more try, I'm not afraid to die
Stand and fight, say what you feel
Born with a heart of steel

CHAPTER 9

Quantum groups

9a. Quantum spaces

We have seen in Part II that the partition function of many statistical mechanical models, of Ising type, has a lot of interesting algebra and combinatorics, coming from the Yang-Baxter equation, and from the symmetry features of the original data.

We discuss here the mathematical aspects of all this, in relation with quantum algebra, at large. There are many possible approaches, and in what follows we will focus on a core idea, which is common to most of these approaches, namely that of constructing a "quantum group" whose representation theory computes the partition function.

Our first approach, which is the most obvious one, is by focusing on the Yang-Baxter equation. The idea indeed is that, algebrically speaking, this equation appears as a generalization of the usual commutation relations $ab = ba$. And the point is that abstract algebra, and more specifically quantum algebra, allows us to construct "quantum groups", which are objects similar to the usual compact Lie groups $G \subset U_N$, but with the commutation relation $ab = ba$ between standard coordinates $u_{ij} : G \to \mathbb{C}$ being replaced precisely by the Yang-Baxter equation. And, once the quantum group constructed, its representation theory can be used in order to study the statistical mechanical model.

Such ideas go back to the work of Faddeev and the Leningrad School from the late 70s [31], [32], [33], and then to the papers of Drinfeld [27] and Jimbo [54]. There are many books explaining this, a particularly enjoyable one being Chari-Pressley [21]. In what concerns us, we will discuss all this more in detail in chapter 11 below.

We will be interested in what follows in a second approach to the partition function question, which still uses quantum groups, but is different from the Yang-Baxter one. The theory here goes back to the work of Jones from the 80s and 90s, originally in terms of subfactors and planar algebras, and then to some subsequent work, reformulating some of Jones' results directly in terms of quantum groups, the story here being as follows:

(1) Jones made two big discoveries in the 80s, on one hand in relation with the index of subfactors [55], and on the other hand with knot invariants [56], [57]. These discoveries are related, both featuring a key statistical mechanical object, namely the algebra

of Temperley-Lieb [80]. Later, Jones wrote an influential paper on knot invariants and statistical mechanics [58], explaining the link between his newly constructed knot invariants, and work in general, and statistical mechanics. And then, all over the 90s, Jones built out of this a unified theory, mixing subfactors, knots and statistical mechanics, with the central notion being that of a planar algebra [59], [60], [61], [62], [63].

(2) At about the same time with Jones writing his paper [58], and laying the foundations of his future planar algebra theory, deeply rooted in the 2D nature of his knot polynomial, Witten wrote an equally famous paper, [94], going in an opposite direction, by proposing there a 3D approach to the same Jones polynomial. This eventually evolved into a connection with the Drinfeld-Jimbo quantum groups [27], [54], but the story did not end here, because Jones' subsequent notion of planar algebra [61] proved to be more general than the Drinfeld-Jimbo quantum groups anyway. In short, thanks to Jones and Witten, all this appears now as a subtle mixture of 2D and 3D mathematics.

(3) As a last piece of story, the compact quantum group formalism of Woronowicz [95], [96] and the subsequent discovery by Wang of the free quantum groups [83], [84] raised the possibility of reformulating some of Jones' work directly in terms of quantum groups, of quite simple type. This is something that I personally got involved into, as a young researcher in the late 90s and early 00s, with a number of theoretical papers, doing this in relation with subfactors and planar algebras. I came back later to this, first in the mid 00s, and then in the mid 10s and afterwards, with a number of more concrete papers, notably those with Bichon [13], Collins [14], Curran [15] and Nechita [16].

This was for the story, and we will briefly discuss all this in this chapter, at the theoretical level, and then later in this book, on various occasions, at the level of concrete models and applications. We will be quite brief, and although most of the mathematics that we will discuss here will be well-known, or at least folklore material, for the most going back to the late 90s, some of the physical applications will be new.

Getting started now, we will need to talk about compact quantum groups in the sense of Woronowicz [95], [96], then free quantum groups in the sense of Wang [83], [84], then matrix models for such quantum groups, following [10], [13], then basic examples of such models, following [14], [16] and other papers, and finally about the relation of this with statistical mechanics, and with Jones' work on subfactors and planar algebras.

Before that, however, we need to know what a quantum space is. According to the work of Heisenberg, then von Neumann and many others, a reasonable definition for the quantum spaces is that of being the abstract duals of the operator algebras. So, it is about the algebras of linear operators $T : H \to H$ on a complex Hilbert space H that we must talk about, first. Let us start with the following definition:

DEFINITION 9.1. *An operator algebra is an algebra of bounded operators $A \subset B(H)$ which contains the unit, is closed under taking adjoints,*

$$T \in A \implies T^* \in A$$

and is closed as well under the norm.

This definition is of course one of the many possible ones, with the choice here being a matter of knowledge of mathematics, and physics, and taste. But more on this later. Getting now to where we wanted to get, with this, we can formulate some tough results, inspired by the usual linear algebra, that of the algebra $M_N(\mathbb{C})$, as follows:

THEOREM 9.2. *The following happen:*

(1) *Any self-adjoint operator, $T = T^*$, is diagonalizable.*
(2) *More generally, any normal operator, $TT^* = T^*T$, is diagonalizable.*
(3) *In fact, any family $\{T_i\}$ of commuting normal operators is diagonalizable.*

Thus, any commutative operator algebra is of the form $A = C(X)$, with X compact space.

PROOF. This is certainly a tough theorem, with (1,2,3) coming by generalizing the Spectral Theorem, in its various incarnations, for the usual matrices $M \in M_N(\mathbb{C})$. As for the final conclusion, this follows from (3), because if we write $A = span(T_i)$, then the family $\{T_i\}$ consists of commuting normal operators, and this leads to the conclusion $A = C(X)$, with X being a certain compact space associated to the family $\{T_i\}$. □

In relation with the above result, there are some good news and some bad news. The good news first, we can, eventually, talk about quantum spaces, as follows:

DEFINITION 9.3. *We can think of any operator algebra $A \subset B(H)$ as being of the form*

$$A = C(X)$$

with X compact quantum space. When A is commutative, X is a usual compact space.

As for the bad news, all this is based on Theorem 9.2, which remains something terribly complicated, and that we would rather like to avoid, when building foundations. Also, there is a problem with functoriality, because a morphism a quantum spaces $X \to Y$ should normally come from a morphism of algebras $C(Y) \to C(X)$, but shall we ask or not something in relation with the embeddings $C(X) \subset B(H)$ and $C(Y) \subset B(K)$. And finally, we have a philosophical problem too, the Hilbert spaces are certainly nice objects, but do we really need them for talking about basic things like quantum spaces.

In addition, as a somewhat dumb argument, needing a Hilbert space for talking about the circle is a bit ridiculous. To be more precise, the circle $\mathbb{T}$ is the simplest compact space that we know, and this since childhood. However, in order to view it as a quantum space, as in Definition 9.3, we need something of type $C(\mathbb{T}) \subset B(L^2(\mathbb{T}))$. No way.

Summarizing, Definition 9.1, Theorem 9.2 and Definition 9.3 are not good, and we must invent something else. And here is the magic trick, due to Gelfand:

DEFINITION 9.4. *An abstract operator algebra, or C^*-algebra, is a complex algebra A having a norm $||.||$ and an involution $*$, subject to the following conditions:*

(1) *A is closed with respect to the norm.*
(2) *We have $||aa^*|| = ||a||^2$, for any $a \in A$.*

In other words, what we did here is to axiomatize the abstract properties of the operator algebras $A \subset B(H)$, without any reference to the Hilbert space H. We will see in a moment that our axiomatization is indeed complete, in the sense that any C^*-algebra appears as an operator algebra, $A \subset B(H)$. Thus, getting back now to our quantum space questions, we will be able to recycle Defintion 9.3, simply by replacing there "operator algebra" by C^*-algebra, and everything, or almost, will be fine.

Getting to work now, let us develop the theory of C^*-algebras. We first have:

PROPOSITION 9.5. *Given an element $a \in A$ of a C^*-algebra, define its spectrum as:*

$$\sigma(a) = \left\{\lambda \in \mathbb{C}\middle| a - \lambda \notin A^{-1}\right\}$$

The following spectral theory results hold, exactly as in the $A = B(H)$ case:

(1) *We have $\sigma(ab) \cup \{0\} = \sigma(ba) \cup \{0\}$.*
(2) *We have $\sigma(f(a)) = f(\sigma(a))$, for any $f \in \mathbb{C}(X)$ having poles outside $\sigma(a)$.*
(3) *The spectrum $\sigma(a)$ is compact, non-empty, and contained in $D_0(||a||)$.*
(4) *The spectra of unitaries ($u^* = u^{-1}$) and self-adjoints ($a = a^*$) are on $\mathbb{T}, \mathbb{R}$.*
(5) *The spectral radius of normal elements ($aa^* = a^*a$) is given by $\rho(a) = ||a||$.*

In addition, assuming $a \in A \subset B$, the spectra of a with respect to A and to B coincide.

PROOF. Here the assertions (1-5), which are formulated a bit informally, are well-known for the full operator algebra $A = B(H)$, and the proof in general is similar:

(1) Assuming that $1 - ab$ is invertible, with inverse c, we have $abc = cab = c - 1$, and it follows that $1 - ba$ is invertible too, with inverse $1 + bca$. Thus $\sigma(ab), \sigma(ba)$ agree on $1 \in \mathbb{C}$, and by linearity, it follows that $\sigma(ab), \sigma(ba)$ agree on any point $\lambda \in \mathbb{C}^*$.

(2) The formula $\sigma(f(a)) = f(\sigma(a))$ is clear for polynomials, $f \in \mathbb{C}[X]$, by factorizing $f - \lambda$, with $\lambda \in \mathbb{C}$. Then, the extension to the rational functions is straightforward, because $P(a)/Q(a) - \lambda$ is invertible precisely when $P(a) - \lambda Q(a)$ is.

(3) By using $1/(1-b) = 1 + b + b^2 + \ldots$ for $||b|| < 1$ we obtain that $a - \lambda$ is invertible for $|\lambda| > ||a||$, and so $\sigma(a) \subset D_0(||a||)$. It is also clear that $\sigma(a)$ is closed, so what we have is a compact set. Finally, assuming $\sigma(a) = \emptyset$ the function $f(\lambda) = \varphi((a - \lambda)^{-1})$ is well-defined, for any $\varphi \in A^*$, and by Liouville we get $f = 0$, contradiction.

(4) Assuming $u^* = u^{-1}$ we have $||u|| = 1$, and so $\sigma(u) \subset D_0(1)$. But with $f(z) = z^{-1}$ we obtain via (2) that we have as well $\sigma(u) \subset f(D_0(1))$, and this gives $\sigma(u) \subset \mathbb{T}$. As for the result regarding the self-adjoints, this can be obtained from the result for the unitaries, by using (2) with functions of type $f(z) = (z + it)/(z - it)$, with $t \in \mathbb{R}$.

(5) It is routine to check, by integrating quantities of type $z^n/(z - a)$ over circles centered at the origin, and estimating, that the spectral radius is given by $\rho(a) = \lim ||a^n||^{1/n}$. But in the self-adjoint case, $a = a^*$, this gives $\rho(a) = ||a||$, by using exponents of type $n = 2^k$, and then the extension to the general normal case is straightforward.

(6) Regarding now the last assertion, the inclusion $\sigma_B(a) \subset \sigma_A(a)$ is clear. For the converse, assume $a - \lambda \in B^{-1}$, and set $b = (a - \lambda)^*(a - \lambda)$. We have then:

$$\sigma_A(b) - \sigma_B(b) = \left\{\mu \in \mathbb{C} - \sigma_B(b) \middle| (b - \mu)^{-1} \in B - A\right\}$$

Thus this difference in an open subset of $\mathbb{C}$. On the other hand b being self-adjoint, its two spectra are both real, and so is their difference. Thus the two spectra of b are equal, and in particular b is invertible in A, and so $a - \lambda \in A^{-1}$, as desired. □

With these ingredients, we can now a prove a key result, as follows:

THEOREM 9.6 (Gelfand). *If X is a compact space, the algebra $C(X)$ of continuous functions on it $f : X \to \mathbb{C}$ is a C^*-algebra, with usual norm and involution, namely:*

$$||f|| = \sup_{x \in X} |f(x)| \quad , \quad f^*(x) = \overline{f(x)}$$

Conversely, any commutative C^*-algebra is of this form, $A = C(X)$, with

$$X = \left\{\chi : A \to \mathbb{C} \ , \ \text{normed algebra character}\right\}$$

with topology making continuous the evaluation maps $ev_a : \chi \to \chi(a)$.

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

(1) The first assertion is clear from definitions. Observe that we have indeed:

$$||ff^*|| = \sup_{x \in X} |f(x)|^2 = ||f||^2$$

Observe also that the algebra $C(X)$ is commutative, because $fg = gf$.

(2) Conversely, given a commutative C^*-algebra A, let us define X as in the statement. Then X is compact, and $a \to ev_a$ is a morphism of algebras, as follows:

$$ev : A \to C(X)$$

(3) We first prove that ev is involutive. We use the following formula, which is similar to the $z = Re(z) + iIm(z)$ decomposition formula for usual complex numbers:

$$a = \frac{a + a^*}{2} + i \cdot \frac{a - a^*}{2i}$$

Thus it is enough to prove $ev_{a^*} = ev_a^*$ for the self-adjoint elements a. But this is the same as proving that $a = a^*$ implies that ev_a is a real function, which is in turn true, by Proposition 9.5, because $ev_a(\chi) = \chi(a)$ is an element of $\sigma(a)$, contained in $\mathbb{R}$.

(4) Since A is commutative, each element is normal, so ev is isometric:

$$||ev_a|| = \rho(a) = ||a||$$

It remains to prove that ev is surjective. But this follows from the Stone-Weierstrass theorem, because $ev(A)$ is a closed subalgebra of $C(X)$, which separates the points. □

The above result is something truly remarkable, and we can now formulate:

DEFINITION 9.7. *Given an arbitrary C^*-algebra A, we write it as*

$$A = C(X)$$

with X compact quantum space. When A is commutative, X is a usual compact space.

Observe the similarity with Definition 9.3, which is now to be forgotten. Indeed, our theory based on C^*-algebras is much better, not using Hilbert spaces, and free as a bird, and all the issues mentioned after Definition 9.3 simply dissapear.

Of course, what we have is still just a beginning of something, and we will soon see, once we will be more advanced, that there are in fact some bugs with Definition 9.7 too. But more on this later, for the moment let us enjoy what we have. A quick comparison between Theorem 9.2 and Theorem 9.6 suggests that operator algebra and C^*-algebra might be actually the same thing. And this is indeed the case, the result being:

THEOREM 9.8. *Any C^*-algebra appears as an operator algebra:*

$$A \subset B(H)$$

Moreover, when A is separable, which is usually the case, H can be taken separable.

PROOF. This result, called GNS representation theorem after Gelfand-Naimark-Segal, comes as a continuation of the Gelfand theorem, the idea being as follows:

(1) Let us first prove that the result holds in the commutative case, $A = C(X)$. Here, we can pick a positive measure on X, and construct our embedding as follows:

$$C(X) \subset B(L^2(X)) \quad , \quad f \to [g \to fg]$$

(2) In general the proof is similar, the idea being that given a C^*-algebra A we can construct a Hilbert space $H = L^2(A)$, and then an embedding as above:

$$A \subset B(L^2(A)) \quad , \quad a \to [b \to ab]$$

(3) Finally, the last assertion is clear, because when A is separable, meaning that it has a countable algebraic basis, so does the associated Hilbert space $H = L^2(A)$. □

All this is very nice, and getting back to our original motivations, we have now a beautiful notion of compact quantum space, coming from Definition 9.4, Theorem 9.6 and Definition 9.7. Also, as a bonus, we have as well some spectral theory tools for the study of such spaces, coming from Proposition 9.5, and even a theorem allowing us to pull out of a hat a Hilbert space, in case we ever get lost, namely Theorem 9.8.

9b. Quantum groups

Regarding now quantum groups, the formalism of Woronowicz [95], suitably modified as to fit with our purposes here, is something very simple, the starting point being:

DEFINITION 9.9. *A Woronowicz algebra is a C^*-algebra A, given with a unitary matrix $u \in M_N(A)$ whose coefficients generate A, such that the formulae*

$$\Delta(u_{ij}) = \sum_k u_{ik} \otimes u_{kj}$$

$$\varepsilon(u_{ij}) = \delta_{ij}$$

$$S(u_{ij}) = u_{ji}^*$$

define morphisms of C^*-algebras $\Delta : A \to A \otimes A$, $\varepsilon : A \to \mathbb{C}$, $S : A \to A^{opp}$.

The morphisms Δ, ε, S are called comultiplication, counit and antipode. We say that A is cocommutative when $\Sigma\Delta = \Delta$, where $\Sigma(a \otimes b) = b \otimes a$ is the flip. We have the following result, which justifies the terminology and axioms:

PROPOSITION 9.10. *The following are Woronowicz algebras:*

(1) *$C(G)$, with $G \subset U_N$ compact Lie group. Here the structural maps are:*

$$\Delta(\varphi) = (g, h) \to \varphi(gh)$$

$$\varepsilon(\varphi) = \varphi(1)$$

$$S(\varphi) = g \to \varphi(g^{-1})$$

(2) *$C^*(\Gamma)$, with $F_N \to \Gamma$ finitely generated group. Here the structural maps are:*

$$\Delta(g) = g \otimes g$$

$$\varepsilon(g) = 1$$

$$S(g) = g^{-1}$$

Moreover, we obtain in this way all the commutative/cocommutative algebras.

PROOF. This is something very standard, the idea being as follows:

(1) Given $G \subset U_N$, we can set $A = C(G)$, which is a Woronowicz algebra, together with the matrix $u = (u_{ij})$ formed by coordinates of G, given by:

$$g = \begin{pmatrix} u_{11}(g) & \ldots & u_{1N}(g) \\ \vdots & & \vdots \\ u_{N1}(g) & \ldots & u_{NN}(g) \end{pmatrix}$$

Conversely, if (A, u) is a commutative Woronowicz algebra, by using the Gelfand theorem we can write $A = C(X)$, with X being a certain compact space. The coordinates u_{ij} give then an embedding $X \subset M_N(\mathbb{C})$, and since the matrix $u = (u_{ij})$ is unitary we actually obtain an embedding $X \subset U_N$, and finally by using the maps Δ, ε, S we conclude that our compact subspace $X \subset U_N$ is in fact a compact Lie group, as desired.

(2) Consider a finitely generated group $F_N \to \Gamma$. We can set $A = C^*(\Gamma)$, which is by definition the completion of the complex group algebra $\mathbb{C}[\Gamma]$, with involution given by $g^* = g^{-1}$, for any $g \in \Gamma$, with respect to the biggest C^*-norm, and we obtain a Woronowicz algebra, together with the diagonal matrix formed by the generators of Γ:

$$u = \begin{pmatrix} g_1 & & 0 \\ & \ddots & \\ 0 & & g_N \end{pmatrix}$$

Conversely, if (A, u) is a cocommutative Woronowicz algebra, the Peter-Weyl theory of Woronowicz, to be explained below, shows that the irreducible corepresentations of A are all 1-dimensional, and form a group Γ, and so we have $A = C^*(\Gamma)$, as desired. □

In relation with the above, in order to avoid amenability issues, we make the convention that, in the context of Definition 9.9, we write $(A, u) = (B, v)$ when there is a $*$-algebra isomorphism as follows, mapping standard coordinates to standard coordinates:

$$< u_{ij} > \simeq < v_{ij} > \quad , \quad u_{ij} \to v_{ij}$$

In general now, the structural maps Δ, ε, S have the following properties:

PROPOSITION 9.11. *Let (A, u) be a Woronowicz algebra.*

(1) *Δ, ε satisfy the usual axioms for a comultiplication and a counit, namely:*

$$\begin{aligned} (\Delta \otimes id)\Delta &= (id \otimes \Delta)\Delta \\ (\varepsilon \otimes id)\Delta &= (id \otimes \varepsilon)\Delta = id \end{aligned}$$

(2) *S satisfies the antipode axiom, on the $*$-subalgebra generated by entries of u:*

$$m(S \otimes id)\Delta = m(id \otimes S)\Delta = \varepsilon(.)1$$

(3) *In addition, the square of the antipode is the identity, $S^2 = id$.*

PROOF. The two comultiplication axioms follow from:

$$\begin{aligned}(\Delta\otimes id)\Delta(u_{ij}) &= (id\otimes\Delta)\Delta(u_{ij})=\sum_{kl}u_{ik}\otimes u_{kl}\otimes u_{lj}\\ (\varepsilon\otimes id)\Delta(u_{ij}) &= (id\otimes\varepsilon)\Delta(u_{ij})=u_{ij}\end{aligned}$$

As for the antipode formulae, the verification here is similar. □

Summarizing, the Woronowicz algebras appear to have nice properties. In view of Proposition 9.10 and Proposition 9.11, we can formulate the following definition:

DEFINITION 9.12. *Given a Woronowicz algebra A, we formally write*

$$A=C(G)=C^*(\Gamma)$$

and call G compact quantum group, and Γ discrete quantum group.

When A is both commutative and cocommutative, G is a compact abelian group, Γ is a discrete abelian group, and these groups are dual to each other, $G=\widehat{\Gamma},\Gamma=\widehat{G}$. In general, we still agree to write, but in a formal sense:

$$G=\widehat{\Gamma}\quad,\quad\Gamma=\widehat{G}$$

With this in mind, let us call now corepresentation of A any unitary matrix $v\in M_n(A)$ satisfying the same conditions as those satisfied by u, namely:

$$\Delta(v_{ij})=\sum_k v_{ik}\otimes v_{kj}\quad,\quad\varepsilon(v_{ij})=\delta_{ij}\quad,\quad S(v_{ij})=v_{ji}^*$$

These corepresentations can be thought of as corresponding representations of the underlying compact quantum group G. Following Woronowicz [95], we have:

THEOREM 9.13. *Any Woronowicz algebra has a unique Haar integration functional,*

$$\left(\int_G\otimes id\right)\Delta=\left(id\otimes\int_G\right)\Delta=\int_G(.)1$$

which can be constructed by starting with any faithful positive form $\varphi\in A^*$, and setting

$$\int_G=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\varphi^{*k}$$

where $\phi*\psi=(\phi\otimes\psi)\Delta$. Moreover, for any corepresentation $v\in M_n(\mathbb{C})\otimes A$ we have

$$\left(id\otimes\int_G\right)v=P$$

where P is the orthogonal projection onto $Fix(v)=\{\xi\in\mathbb{C}^n|v\xi=\xi\}$.

PROOF. Following [95], this can be done in 3 steps, as follows:

(1) Given $\varphi \in A^*$, our claim is that the following limit converges, for any $a \in A$:

$$\int_\varphi a = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \varphi^{*k}(a)$$

Indeed, by linearity we can assume that a is the coefficient of corepresentation, $a = (\tau \otimes id)v$. But in this case, an elementary computation shows that we have the following formula, where P_φ is the orthogonal projection onto the 1-eigenspace of $(id \otimes \varphi)v$:

$$\left(id \otimes \int_\varphi\right) v = P_\varphi$$

(2) Since $v\xi = \xi$ implies $[(id \otimes \varphi)v]\xi = \xi$, we have $P_\varphi \geq P$, where P is the orthogonal projection onto the space $Fix(v) = \{\xi \in \mathbb{C}^n | v\xi = \xi\}$. The point now is that when $\varphi \in A^*$ is faithful, by using a positivity trick, one can prove that we have $P_\varphi = P$. Thus our linear form $\int_\varphi$ is independent of φ, and is given on coefficients $a = (\tau \otimes id)v$ by:

$$\left(id \otimes \int_\varphi\right) v = P$$

(3) With the above formula in hand, the left and right invariance of $\int_G = \int_\varphi$ is clear on coefficients, and so in general, and this gives all the assertions. See [95]. □

Consider the dense $*$-subalgebra $\mathcal{A} \subset A$ generated by the coefficients of the fundamental corepresentation u, and endow it with the following scalar product:

$$< a, b >= \int_G ab^*$$

We have then the following result, also due to Woronowicz [95]:

THEOREM 9.14. *We have the following Peter-Weyl type results:*

(1) *Any corepresentation decomposes as a sum of irreducible corepresentations.*
(2) *Each irreducible corepresentation appears inside a certain $u^{\otimes k}$.*
(3) $\mathcal{A} = \bigoplus_{v \in Irr(A)} M_{\dim(v)}(\mathbb{C})$*, the summands being pairwise orthogonal.*
(4) *The characters of irreducible corepresentations form an orthonormal system.*

PROOF. All these results are from [95], the idea being as follows:

(1) Given $v \in M_n(A)$, its intertwiner algebra $End(v) = \{T \in M_n(\mathbb{C}) | Tv = vT\}$ is a finite dimensional C^*-algebra, and so decomposes as $End(v) = M_{n_1}(\mathbb{C}) \oplus \ldots \oplus M_{n_r}(\mathbb{C})$. But this gives a decomposition of type $v = v_1 + \ldots + v_r$, as desired.

(2) Consider indeed the Peter-Weyl corepresentations, $u^{\otimes k}$ with k colored integer, defined by $u^{\otimes \emptyset} = 1$, $u^{\otimes \circ} = u$, $u^{\otimes \bullet} = \bar{u}$ and multiplicativity. The coefficients of these corepresentations span the dense algebra $\mathcal{A}$, and by using (1), this gives the result.

(3) Here the direct sum decomposition, which is technically a $*$-coalgebra isomorphism, follows from (2). As for the second assertion, this follows from the fact that $(id\otimes\int_G)v$ is the orthogonal projection P_v onto the space $Fix(v)$, for any corepresentation v.

(4) Let us define indeed the character of $v\in M_n(A)$ to be the matrix trace, $\chi_v=Tr(v)$. Since this character is a coefficient of v, the orthogonality assertion follows from (3). As for the norm 1 claim, this follows once again from $(id\otimes\int_G)v=P_v$. □

As a basic application, observe that in the cocommutative case, we obtain from (4) that the irreducible corepresentations must be all 1-dimensional, and so that we must have $A=C^*(\Gamma)$ for some discrete group Γ, as mentioned in Proposition 9.10.

9c. Basic examples

Moving ahead now, let us talk about some truly interesting examples of compact quantum groups in the sense of Woronowicz, namely the free quantum groups in the sense of Wang [83], [84]. There is a long story here, and we first have, following [83]:

THEOREM 9.15. *The following universal algebras are Woronowicz algebras,*

$$C(O_N^+)=C^*\left((u_{ij})_{i,j=1,\ldots,N}\middle|u=\bar{u},u^t=u^{-1}\right)$$

$$C(U_N^+)=C^*\left((u_{ij})_{i,j=1,\ldots,N}\middle|u^*=u^{-1},u^t=\bar{u}^{-1}\right)$$

so the underlying quantum spaces O_N^+,U_N^+ are compact quantum groups.

PROOF. This follows from the elementary fact that if a matrix $u=(u_{ij})$ is orthogonal or biunitary, then so must be the following matrices:

$$u_{ij}^{\Delta}=\sum_k u_{ik}\otimes u_{kj}\quad,\quad u_{ij}^{\varepsilon}=\delta_{ij}\quad,\quad u_{ij}^{S}=u_{ji}^*$$

Thus, we can indeed define morphisms Δ,ε,S as in Definition 9.9, by using the universal properties of $C(O_N^+)$, $C(U_N^+)$, and this gives the result. □

Still following Wang [84], let us discuss now the construction and basic properties of the quantum permutation group S_N^+. Let us first look at S_N. We have:

PROPOSITION 9.16. *Consider the symmetric group S_N, viewed as permutation group of the N coordinate axes of $\mathbb{R}^N$. The coordinate functions on $S_N\subset O_N$ are given by*

$$u_{ij}=\chi\left(\sigma\in G\middle|\sigma(j)=i\right)$$

and the matrix $u=(u_{ij})$ that these functions form is magic, in the sense that its entries are projections ($p^2=p^*=p$), summing up to 1 on each row and each column.

PROOF. The action of S_N on the standard basis $e_1,\ldots,e_N\in\mathbb{R}^N$ being given by $\sigma:e_j\to e_{\sigma(j)}$, this gives the formula of u_{ij} in the statement. As for the fact that the matrix $u=(u_{ij})$ that these functions form is magic, this is clear. □

With a bit more effort, we obtain the following nice characterization of S_N:

PROPOSITION 9.17. *The algebra of functions on S_N has the following presentation,*

$$C(S_N) = C^*_{comm}\left((u_{ij})_{i,j=1,\ldots,N}\middle|u = \text{magic}\right)$$

and the multiplication, unit and inversion map of S_N appear from the maps

$$\Delta(u_{ij}) = \sum_k u_{ik} \otimes u_{kj} \quad , \quad \varepsilon(u_{ij}) = \delta_{ij} \quad , \quad S(u_{ij}) = u_{ji}$$

defined at the algebraic level, of functions on S_N, by transposing.

PROOF. The universal algebra A in the statement being commutative, by the Gelfand theorem it must be of the form $A = C(X)$, with X being a certain compact space. Now since we have coordinates $u_{ij} : X \to \mathbb{R}$, we have an embedding $X \subset M_N(\mathbb{R})$. Also, since we know that these coordinates form a magic matrix, the elements $g \in X$ must be 0-1 matrices, having exactly one 1 entry on each row and each column, and so $X = S_N$. Thus we have proved the first assertion, and the second assertion is clear as well. □

We can now liberate S_N, still following Wang's paper [84], as follows:

THEOREM 9.18. *The following universal C^*-algebra, with magic meaning as usual formed by projections ($p^2 = p^* = p$), summing up to 1 on each row and each column,*

$$C(S_N^+) = C^*\left((u_{ij})_{i,j=1,\ldots,N}\middle|u = \text{magic}\right)$$

is a Woronowicz algebra, with comultiplication, counit and antipode given by:

$$\Delta(u_{ij}) = \sum_k u_{ik} \otimes u_{kj} \quad , \quad \varepsilon(u_{ij}) = \delta_{ij} \quad , \quad S(u_{ij}) = u_{ji}$$

Thus the space S_N^+ is a compact quantum group, called quantum permutation group.

PROOF. In order to prove that we have a Woronowicz algebra, we must construct maps Δ, ε, S given by the formulae in the statement. Consider the following matrices:

$$u^\Delta_{ij} = \sum_k u_{ik} \otimes u_{kj} \quad , \quad u^\varepsilon_{ij} = \delta_{ij} \quad , \quad u^S_{ij} = u_{ji}$$

Our claim is that, since u is magic, so are these three matrices. Indeed, regarding u^Δ, its entries are idempotents, as shown by the following computation:

$$(u^\Delta_{ij})^2 = \sum_{kl} u_{ik}u_{il} \otimes u_{kj}u_{lj} = \sum_{kl} \delta_{kl}u_{ik} \otimes \delta_{kl}u_{kj} = u^\Delta_{ij}$$

These elements are self-adjoint as well, as shown by the following computation:

$$(u^\Delta_{ij})^* = \sum_k u^*_{ik} \otimes u^*_{kj} = \sum_k u_{ik} \otimes u_{kj} = u^\Delta_{ij}$$

The row and column sums for the matrix u^Δ can be computed as follows:

$$\sum_j u_{ij}^\Delta = \sum_{jk} u_{ik} \otimes u_{kj} = \sum_k u_{ik} \otimes 1 = 1$$

$$\sum_i u_{ij}^\Delta = \sum_{ik} u_{ik} \otimes u_{kj} = \sum_k 1 \otimes u_{kj} = 1$$

Thus, u^Δ is magic. Regarding now u^ε, u^S, these matrices are magic too, and this for obvious reasons. Thus, all our three matrices $u^\Delta, u^\varepsilon, u^S$ are magic, so we can define Δ, ε, S by the formulae in the statement, by using the universality property of $C(S_N^+)$. □

Our first task now is to make sure that Theorem 9.18 produces indeed a new quantum group, which does not collapse to S_N. Following Wang [84], we have:

THEOREM 9.19. *We have an embedding $S_N \subset S_N^+$, given at the algebra level by:*

$$u_{ij} \to \chi\left(\sigma \in S_N \middle| \sigma(j) = i\right)$$

This is an isomorphism at $N \leq 3$, but not at $N \geq 4$, where S_N^+ is not classical, nor finite.

PROOF. The fact that we have an embedding as above follows from Proposition 9.17. Observe that in fact more is true, because Proposition 9.17 and Theorem 9.18 give:

$$C(S_N) = C(S_N^+) \Big/ \Big\langle ab = ba \Big\rangle$$

Thus, the inclusion $S_N \subset S_N^+$ is a "liberation", in the sense that S_N is the classical version of S_N^+. We will often use this basic fact, in what follows. Regarding now the second assertion, we can prove this in four steps, as follows:

Case $N = 2$. The fact that S_2^+ is indeed classical, and hence collapses to S_2, is trivial, because the 2×2 magic matrices are as follows, with p being a projection:

$$U = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}$$

Indeed, this shows that the entries of U commute. Thus $C(S_2^+)$ is commutative, and so equals its biggest commutative quotient, which is $C(S_2)$. Thus, $S_2^+ = S_2$.

Case $N = 3$. By using the same argument as in the $N = 2$ case, and the symmetries of the problem, it is enough to check that u_{11}, u_{22} commute. But this follows from:

$$\begin{aligned} u_{11}u_{22} &= u_{11}u_{22}(u_{11} + u_{12} + u_{13}) \\ &= u_{11}u_{22}u_{11} + u_{11}u_{22}u_{13} \\ &= u_{11}u_{22}u_{11} + u_{11}(1 - u_{21} - u_{23})u_{13} \\ &= u_{11}u_{22}u_{11} \end{aligned}$$

Indeed, by applying the involution to this formula, we obtain that we have as well $u_{22}u_{11} = u_{11}u_{22}u_{11}$. Thus, we obtain $u_{11}u_{22} = u_{22}u_{11}$, as desired.

Case $N = 4$. Consider the following matrix, with p, q being projections:

$$U = \begin{pmatrix} p & 1-p & 0 & 0 \\ 1-p & p & 0 & 0 \\ 0 & 0 & q & 1-q \\ 0 & 0 & 1-q & q \end{pmatrix}$$

This matrix is magic, and we can choose $p, q \in B(H)$ as for the algebra $< p, q >$ to be noncommutative and infinite dimensional. We conclude that $C(S_4^+)$ is noncommutative and infinite dimensional as well, and so S_4^+ is non-classical and infinite, as claimed.

Case $N \geq 5$. Here we can use the standard embedding $S_4^+ \subset S_N^+$, obtained at the level of the corresponding magic matrices in the following way:

$$u \to \begin{pmatrix} u & 0 \\ 0 & 1_{N-4} \end{pmatrix}$$

Indeed, with this in hand, the fact that S_4^+ is a non-classical, infinite compact quantum group implies that S_N^+ with $N \geq 5$ has these two properties as well. □

The above result is quite surprising. As a first observation, as a matter of doublechecking our findings, we are not wrong with our formalism, because we have as well:

THEOREM 9.20. *The quantum permutation group S_N^+ acts on the set $X = \{1, \ldots, N\}$, the corresponding coaction map $\Phi : C(X) \to C(X) \otimes C(S_N^+)$ being given by:*

$$\Phi(e_i) = \sum_j e_j \otimes u_{ji}$$

In fact, S_N^+ is the biggest compact quantum group acting on X, by leaving the counting measure invariant, in the sense that $(tr \otimes id)\Phi = tr(.)1$, where $tr(e_i) = \frac{1}{N}, \forall i$.

PROOF. Our claim is that given a compact matrix quantum group G, the following formula defines a morphism of algebras, which is a coaction map, leaving the trace invariant, precisely when the matrix $u = (u_{ij})$ is a magic corepresentation of $C(G)$:

$$\Phi(e_i) = \sum_j e_j \otimes u_{ji}$$

Indeed, let us first determine when Φ is multiplicative. We have:

$$\Phi(e_i)\Phi(e_k) = \sum_{jl} e_j e_l \otimes u_{ji} u_{lk} = \sum_j e_j \otimes u_{ji} u_{jk}$$

$$\Phi(e_i e_k) = \delta_{ik}\Phi(e_i) = \delta_{ik} \sum_j e_j \otimes u_{ji}$$

We conclude that the multiplicativity of Φ is equivalent to the following conditions:

$$u_{ji} u_{jk} = \delta_{ik} u_{ji} \quad , \quad \forall i, j, k$$

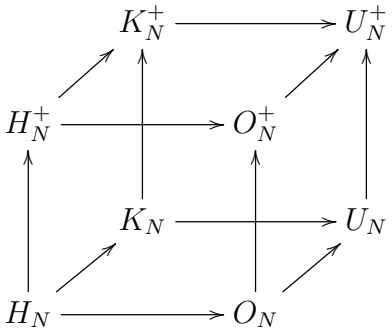
Similarly, Φ is unital when $\sum_i u_{ji} = 1$, $\forall j$. Finally, the fact that Φ is a $*$-morphism translates into $u_{ij} = u_{ij}^*$, $\forall i, j$. Summing up, in order for $\Phi(e_i) = \sum_j e_j \otimes u_{ji}$ to be a morphism of C^*-algebras, the elements u_{ij} must be projections, summing up to 1 on each row of u. Regarding now the preservation of the trace, observe that we have:

$$(tr \otimes id)\Phi(e_i) = \frac{1}{N}\sum_j u_{ji}$$

Thus the trace is preserved precisely when the elements u_{ij} sum up to 1 on each of the columns of u. We conclude from this that $\Phi(e_i) = \sum_j e_j \otimes u_{ji}$ is a morphism of C^*-algebras preserving the trace precisely when u is magic, and since the coaction conditions on Φ are equivalent to the fact that u must be a corepresentation, this finishes the proof of our claim. But this claim proves all the assertions in the statement. □

More generally now, the above constructions $O_N \to O_N^+$, $U_N \to U_N^+$, $S_N \to S_N^+$ can be suitably adapted to other compact Lie groups, and we have the following result:

THEOREM 9.21. *The classical and free, real and complex quantum rotation groups can be complemented with quantum reflection groups, as follows,*



with $H_N = \mathbb{Z}_2 \wr S_N$ and $K_N = \mathbb{T} \wr S_N$ being the hyperoctahedral group and the full complex reflection group, and $H_N^+ = \mathbb{Z}_2 \wr_* S_N^+$ and $K_N^+ = \mathbb{T} \wr_* S_N^+$ being their free versions.

PROOF. Consider the group $H_N^s \subset U_N$ consisting of permutation-like matrices having as entries the s-th roots of unity. This group decomposes as follows:

$$H_N^s = \mathbb{Z}_s \wr S_N$$

It is straightforward then to construct a free analogue $H_N^{s+} \subset U_N^+$ of this group, for instance by formulating a definition as follows, with $\wr_*$ being a free wreath product:

$$H_N^{s+} = \mathbb{Z}_s \wr_* S_N^+$$

In order to finish, besides the case $s = 1$, of particular interest are the cases $s = 2, \infty$. Here the corresponding reflection groups are as follows:

$$H_N = \mathbb{Z}_2 \wr S_N \quad , \quad K_N = \mathbb{T} \wr S_N$$

As for the corresponding quantum groups, these are denoted as follows:

$$H_N^+ = \mathbb{Z}_2 \wr_* S_N^+ \quad , \quad K_N^+ = \mathbb{T} \wr_* S_N^+$$

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

9d. Matrix models

Getting now the point where we wanted to get, namely matrix models for the quantum groups, the story here goes back to the paper [10], where this construction was first axiomatixed. We will follow here the modern approach from [13]. Let us start with:

DEFINITION 9.22. *A matrix model for $G \subset U_N^+$ is a morphism of C^*-algebras*

$$\pi : C(G) \to M_K(C(T))$$

where $K \geq 1$ is an integer, and T is a compact space.

The simplest situation is when π is faithful in the usual sense. Here π obviously reminds G. However, this is something quite restrictive, because in this case the algebra $C(G)$ must be quite small, admitting an embedding as follows:

$$\pi : C(G) \subset M_K(C(T))$$

Technically, this means that $C(G)$ must be of type I, as an operator algebra, and we will discuss this in a moment, with the comment that this is indeed something quite restrictive. However, there are many interesting examples here, and all this is worth a detailed look. First, we have the following result, providing us with basic examples:

PROPOSITION 9.23. *The following closed subgroups $G \subset U_N^+$ have faithful models:*

(1) *The compact Lie groups $G \subset U_N$.*
(2) *The finite quantum groups $G \subset U_N^+$.*

In both cases, we can arrange for $\int_G$ to be restriction of the random matrix trace.

PROOF. These assertions are all elementary, the proofs being as follows:

(1) This is clear, because we can simply use here the identity map:

$$id : C(G) \to M_1(C(G))$$

(2) Here we can use the left regular representation $\lambda : C(G) \to M_{|G|}(\mathbb{C})$. Indeed, let us endow the linear space $H = C(G)$ with the scalar product $< a, b >= \int_G ab^*$. We have then a representation of $*$-algebras, as follows:

$$\lambda : C(G) \to B(H) \quad , \quad a \to [b \to ab]$$

Now since we have $H \simeq \mathbb{C}^{|G|}$, we can view λ as a matrix model map, as above.

(3) Finally, our claim is that we can choose our model as for the following formula to hold, where $\int_T$ is the integration with respect to a given probability measure on T:

$$\int_G = \left(tr \otimes \int_T \right) \pi$$

But this is clear for the model in (1), by definition, and is clear as well for the model in (2), by using the basic properties of the left regular representation. □

In the above result, the last assertion is quite interesting, and suggests formulating the following definition, somewhat independently on the notion of faithfulness:

DEFINITION 9.24. *A matrix model $\pi : C(G) \to M_K(C(T))$ is called stationary when*

$$\int_G = \left(tr \otimes \int_T \right) \pi$$

where $\int_T$ is the integration with respect to a given probability measure on T.

We will see in a moment that stationarity implies faithfulness, so that stationarity can be regarded as being a useful, pragmatic version of faithfulness. In order to discuss this, we will need the following result, due to Murray-von Neumann and Connes:

THEOREM 9.25. *Given a von Neumann algebra $A \subset B(H)$, if we write its center as*

$$Z(A) = L^\infty(X)$$

then we have a decomposition as follows, with the fibers A_x having trivial center:

$$A = \int_X A_x \, dx$$

Moreover, the factors, $Z(A) = \mathbb{C}$, can be basically classified in terms of the II_1 factors, which are those satisfying $\dim A = \infty$, and having a faithful trace $tr : A \to \mathbb{C}$.

PROOF. This is something that we know to hold in finite dimensions, and in the commutative case too. In general, this is something heavy, the idea being as follows:

(1) The first assertion, regarding the decomposition into factors, is von Neumann's reduction theory main result, which is actually one of the heaviest results in fundamental mathematics, and whose proof uses advanced functional analysis techniques.

(2) The classification of factors, due to Murray-von Neumann and Connes, is again something heavy, the idea being that the II_1 factors are the "building blocks", with other factors basically appearing from them via crossed product type constructions. □

Back now to matrix models, as a first general result, which is something which is not exactly trivial, and whose proof requires some functional analysis, we have:

THEOREM 9.26. *Assuming that a closed subgroup $G \subset U_N^+$ has a stationary model*

$$\pi : C(G) \to M_K(C(T))$$

it follows that G must be coamenable, and that the model is faithful. Moreover, π extends into an embedding of von Neumann algebras, as follows,

$$L^\infty(G) \subset M_K(L^\infty(T))$$

which commutes with the canonical integration functionals.

PROOF. Assume that we have a stationary model, as in the statement. By performing the GNS construction with respect to $\int_G$, we obtain a factorization as follows, which commutes with the respective canonical integration functionals:

$$\pi : C(G) \to C(G)_{red} \subset M_K(C(T))$$

Thus, in what regards the coamenability question, we can assume that π is faithful. With this assumption made, we have an embedding as follows:

$$C(G) \subset M_K(C(T))$$

Now observe that the GNS construction gives a better embedding, as follows:

$$L^\infty(G) \subset M_K(L^\infty(T))$$

Now since the von Neumann algebra on the right is of type I, so must be its subalgebra $A = L^\infty(G)$. But this means that, when writing the center of this latter algebra as $Z(A) = L^\infty(X)$, the whole algebra decomposes over X, as an integral of type I factors:

$$L^\infty(G) = \int_X M_{K_x}(\mathbb{C})\, dx$$

In particular, we can see from this that $C(G) \subset L^\infty(G)$ has a unique C^*-norm, and so G is coamenable. Thus we have proved our first assertion, and the second assertion follows as well, because our factorization of π consists of the identity, and of an inclusion. □

Let us discuss now the general, non-coamenable case, with the aim of finding a weaker notion of faithfulness, which still does the job, namely that of "reminding" the quantum group. The idea comes by looking at the group duals $G = \widehat{\Gamma}$. Consider indeed a general model for the associated group algebra, which can be written as follows:

$$\pi : C^*(\Gamma) \to M_K(C(T))$$

The point is that such a representation of the group algebra must come by linearization from a unitary group representation, as follows:

$$\rho : \Gamma \to C(T, U_K)$$

Now observe that when this group representation ρ is faithful, the representation π is in general not faithful, for instance because when $T = \{.\}$ its target algebra is finite dimensional. On the other hand, this representation "reminds" Γ, so can be used in order

to fully understand Γ. Thus, we have an idea here, basically saying that, for practical purposes, the faithfuless property can be replaced with something much weaker.

This weaker notion, which will be of great interest for us, is called "inner faithfulness". The general theory here, from [13], starts with the following definition:

DEFINITION 9.27. *Let $\pi : C(G) \to M_K(C(T))$ be a matrix model.*

(1) *The Hopf image of π is the smallest quotient Hopf C^*-algebra $C(G) \to C(H)$ producing a factorization as follows:*

$$\pi : C(G) \to C(H) \to M_K(C(T))$$

(2) *When the inclusion $H \subset G$ is an isomorphism, i.e. when there is no non-trivial factorization as above, we say that π is inner faithful.*

The above notions are quite tricky, and having them well understood will take us some time. As a first example, motivated by the above discussion, in the case where $G = \widehat{\Gamma}$ is a group dual, π must come from a group representation, as follows:

$$\rho : \Gamma \to C(T, U_K)$$

Thus the minimal factorization in (1) is obtained by taking the image:

$$\rho : \Gamma \to \Lambda \subset C(T, U_K)$$

Thus, as a conclusion, in this case π is inner faithful precisely when we have:

$$\Gamma \subset C(T, U_K)$$

Dually now, given a compact Lie group G, and elements $g_1, \ldots, g_K \in G$, we have a diagonal representation $\pi : C(G) \to M_K(\mathbb{C})$, appearing as follows:

$$f \to \begin{pmatrix} f(g_1) & & \\ & \ddots & \\ & & f(g_K) \end{pmatrix}$$

The minimal factorization of this representation π, as in Definition 9.27 (1), is then via the algebra $C(H)$, with H being the following closed subgroup of G:

$$H = \overline{< g_1, \ldots, g_K >}$$

Thus, as a conclusion, π is inner faithful precisely when we have:

$$G = H$$

There are many other examples of inner faithful representations, which are however far more technically advanced, and we will discuss them later.

Back to general theory now, in the framework of Definition 9.27, the existence and uniqueness of the Hopf image come by dividing $C(G)$ by a suitable ideal, with this being something standard. Alternatively, in Tannakian terms, as explained in [13], we have:

THEOREM 9.28. *Assuming $G \subset U_N^+$, with fundamental corepresentation $u = (u_{ij})$, the Hopf image of a model $\pi : C(G) \to M_K(C(T))$ comes from the Tannakian category*

$$C_{kl} = Hom(U^{\otimes k}, U^{\otimes l})$$

where $U_{ij} = \pi(u_{ij})$, and where the spaces on the right are taken in a formal sense.

PROOF. Since morphisms increase the intertwining spaces, when defined either in a representation theory sense, or just formally, we have inclusions as follows:

$$Hom(u^{\otimes k}, u^{\otimes l}) \subset Hom(U^{\otimes k}, U^{\otimes l})$$

More generally, we have such inclusions when replacing (G, u) with any pair producing a factorization of π. Thus, by Tannakian duality, the Hopf image must be given by the fact that the intertwining spaces must be the biggest, subject to the above inclusions. On the other hand, since u is biunitary, so is U, and it follows that the spaces on the right form a Tannakian category. Thus, we have a quantum group (H, v) given by:

$$Hom(v^{\otimes k}, v^{\otimes l}) = Hom(U^{\otimes k}, U^{\otimes l})$$

By the above discussion, $C(H)$ follows to be the Hopf image of π, as claimed. □

Regarding now the study of the inner faithful models, a key problem is that of computing the Haar integration functional. The result here is as follows:

THEOREM 9.29. *Given an inner faithful model $\pi : C(G) \to M_K(C(T))$, we have*

$$\int_G = \lim_{k\to\infty} \frac{1}{k} \sum_{r=1}^{k} \int_G^r$$

with the truncations of the integration on the right being given by

$$\int_G^r = (\varphi \circ \pi)^{*r}$$

with $\phi * \psi = (\phi \otimes \psi)\Delta$, and with $\varphi = tr \otimes \int_T$ being the random matrix trace.

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

(1) As a first observation, there is an obvious similarity here with the Woronowicz construction of the Haar measure, explained before. In fact, the above result holds more generally for any model $\pi : C(G) \to B$, with $\varphi \in B^*$ being a faithful trace.

(2) In order to prove now the result, we can proceed as before. If we denote by $\int_G'$ the limit in the statement, we must prove that this limit converges, and that:

$$\int_G' = \int_G$$

It is enough to check this on the coefficients of the Peter-Weyl corepresentations, and if we let $v = u^{\otimes k}$ be one of these corepresentations, we must prove that we have:

$$\left(id \otimes \int_G'\right) v = \left(id \otimes \int_G\right) v$$

(3) In order to prove this, we already know, from the Haar measure theory discussed before, that the matrix on the right is the orthogonal projection onto $Fix(v)$:

$$\left(id \otimes \int_G\right) v = Proj\Big[Fix(v)\Big]$$

Regarding now the matrix on the left, the trick in [95] applied to the linear form $\varphi\pi$ tells us that this is the orthogonal projection onto the 1-eigenspace of $(id \otimes \varphi\pi)v$:

$$\left(id \otimes \int_G'\right) v = Proj\Big[1 \in (id \otimes \varphi\pi)v\Big]$$

(4) Now observe that, if we set $V_{ij} = \pi(v_{ij})$, we have the following formula:

$$(id \otimes \varphi\pi)v = (id \otimes \varphi)V$$

Thus, we can apply the trick in [95], and we conclude that the 1-eigenspace that we are interested in equals $Fix(V)$. But, according to Theorem 9.28, we have:

$$Fix(V) = Fix(v)$$

Thus, we have proved that we have $\int_G' = \int_G$, as desired. □

Interestingly, the above results regarding inner faithfulness have applications as well to the notion of stationarity introduced before, clarifying among others the use of the word "stationary". To be more precise, in order to detect the stationary models, we have the following useful criterion, mixing linear algebra and analysis:

THEOREM 9.30. *For a model $\pi : C(G) \to M_K(C(T))$, the following are equivalent:*

(1) *$Im(\pi)$ is a Hopf algebra, and the Haar integration on it is:*

$$\psi = \left(tr \otimes \int_T\right) \pi$$

(2) *The linear form $\psi = (tr \otimes \int_T)\pi$ satisfies the idempotent state property:*

$$\psi * \psi = \psi$$

(3) *We have $T_e^2 = T_e$, $\forall p \in \mathbb{N}$, $\forall e \in \{1, *\}^p$, where:*

$$(T_e)_{i_1 \ldots i_p, j_1 \ldots j_p} = \left(tr \otimes \int_T\right) (U_{i_1 j_1}^{e_1} \ldots U_{i_p j_p}^{e_p})$$

If these conditions are satisfied, we say that π is stationary on its image.

PROOF. Given a matrix model $\pi : C(G) \to M_K(C(T))$ as in the statement, we can factorize it via its Hopf image, as in Definition 9.27:

$$\pi : C(G) \to C(H) \to M_K(C(T))$$

Now observe that (1,2,3) above depend only on the factorized representation:

$$\nu : C(H) \to M_K(C(T))$$

Thus, we can assume in practice that we have $G = H$, which means that we can assume that π is inner faithful. With this assumption made, the formula in Theorem 9.29 applies to our situation, and the proof of the equivalences goes as follows:

(1) $\implies$ (2) This is clear from definitions, because the Haar integration on any compact quantum group satisfies the idempotent state equation:

$$\psi * \psi = \psi$$

(2) $\implies$ (1) Assuming $\psi * \psi = \psi$, we have $\psi^{*r} = \psi$ for any $r \in \mathbb{N}$, and Theorem 9.29 gives $\int_G = \psi$. By using now Theorem 9.26, we obtain the result.

In order to establish now (2) $\iff$ (3), we use the following elementary formula, which comes from the definition of the convolution operation:

$$\psi^{*r}(u^{e_1}_{i_1j_1} \ldots u^{e_p}_{i_pj_p}) = (T_e^r)_{i_1 \ldots i_p, j_1 \ldots j_p}$$

(2) $\implies$ (3) Assuming $\psi * \psi = \psi$, by using the above formula at $r = 1, 2$ we obtain that the matrices T_e and T_e^2 have the same coefficients, and so they are equal.

(3) $\implies$ (2) Assuming $T_e^2 = T_e$, by using the above formula at $r = 1, 2$ we obtain that the linear forms ψ and $\psi * \psi$ coincide on any product of coefficients $u^{e_1}_{i_1j_1} \ldots u^{e_p}_{i_pj_p}$. Now since these coefficients span a dense subalgebra of $C(G)$, this gives the result. □

Following [14], [16], there are many interesting examples of stationary models, related to the Pauli matrices, and to the Weyl matrices. We will discuss them later.

9e. Exercises

Exercises:

EXERCISE 9.31.

EXERCISE 9.32.

EXERCISE 9.33.

EXERCISE 9.34.

EXERCISE 9.35.

EXERCISE 9.36.

Bonus exercise.

CHAPTER 10

Ising and Potts

10a. Ising model

Ising model.

10b. Partition function

Partition function.

10c. Critical behavior

Critical behavior.

10d. Potts model

Potts model.

10e. Exercises

Exercises:

Exercise 10.1.

Exercise 10.2.

Exercise 10.3.

Exercise 10.4.

Exercise 10.5.

Exercise 10.6.

Bonus exercise.

CHAPTER 11

Planar algebras

11a. Subfactor theory

In order to discuss subfactors, we will need some basic von Neumann algebra theory, coming as a complement to the operator algebra theory from chapter 9, as follows:

THEOREM 11.1. *The von Neumann algebras, which are the $*$-algebras of operators*

$$A \subset B(H)$$

closed under the weak operator topology, making each $T \to Tx$ continuous, are as follows:

(1) *They are exactly the $*$-algebras of operators $A \subset B(H)$ which are equal to their bicommutant, $A = A''$.*
(2) *In the commutative case, these are the algebras $A = L^\infty(X)$, with X measured space, represented on $H = L^2(X)$, up to a multiplicity.*
(3) *If we write the center as $Z(A) = L^\infty(X)$, then we have a decomposition of type $A = \int_X A_x\, dx$, with the fibers A_x having trivial center, $Z(A_x) = \mathbb{C}$.*
(4) *The factors, $Z(A) = \mathbb{C}$, can be fully classified in terms of II_1 factors, which are those satisfying $\dim A = \infty$, and having a faithful trace $tr : A \to \mathbb{C}$.*
(5) *The II_1 factors enjoy the "continuous dimension geometry" property, in the sense that the traces of their projections can take any values in $[0, 1]$.*
(6) *Among the II_1 factors, the most important one is the Murray-von Neumann hyperfinite factor R, obtained as an inductive limit of matrix algebras.*

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

(1) This is von Neumann's bicommutant theorem, which is well-known in finite dimensions, and whose proof in general is not that complicated, either.

(2) It is clear, via basic measure theory, that $L^\infty(X)$ is indeed a von Neumann algebra on $H = L^2(X)$. The converse can be proved as well, by using spectral theory.

(3) This is von Neumann's reduction theory main result, whose statement is already quite hard to understand, and whose proof uses advanced functional analysis.

(4) This is something heavy, due to Murray-von Neumann and Connes, the idea being that the other factors can be basically obtained via crossed product constructions.

(5) This is a gem of functional analysis, with the rational traces being relatively easy to obtain, and with the irrational ones coming from limiting arguments.

(6) Once again, heavy results, by Murray-von Neumann and Connes, the idea being that any finite dimensional construction always leads to the same factor, called R. □

Now given a II_1 factor $A \subset B(H)$, some standard functional analysis study shows that H must appear as an "inflated" version of $L^2(A)$. The corresponding inflation constant is a certain number, that we can call coupling constant, as follows:

DEFINITION 11.2. *Given a representation of a* II_1 *factor* $A \subset B(H)$*, we can talk about the corresponding coupling constant, as being the number*

$$\dim_A H \in (0, \infty]$$

constructed as follows, with $u : H \to L^2(A) \otimes l^2(\mathbb{N})$ *isometry satisfying* $ux = (x \otimes 1)u$*:*

$$\dim_A H = tr(uu^*)$$

For the standard form, where $H = L^2(A)$*, this coupling constant takes the value* 1*.*

This definition might seem a bit complicated, but things here are quite non-trivial, and there is no way of doing something substantially simpler. Alternatively, we can define the coupling constant via the following formula, after proving first that the number on the right is indeed independent of the choice on a nonzero vector $x \in H$:

$$\dim_A H = \frac{tr_A(P_{A'x})}{tr_{A'}(P_{Ax})}$$

This latter formula was in fact the original definition of the coupling constant, by Murray and von Neumann. However, technically speaking, things are easier when using the approach in Definition 11.2. For details here, check any operator algebra book.

Given an inclusion of II_1 factors $A_0 \subset A_1$, a first question is that of defining its index, measuring how big is A_1, when compared to A_0. This can be done as follows:

THEOREM 11.3. *Given an inclusion of* II_1 *factors* $A_0 \subset A_1$*, the number*

$$N = \frac{\dim_{A_0} H}{\dim_{A_1} H}$$

is independent of the ambient Hilbert space H*, and is called index.*

PROOF. This is standard, with the fact that the index as defined by the above formula is indeed independent of the ambient Hilbert space H coming from the various basic properties of the coupling constant, established by Murray and von Neumann. □

There are many examples of subfactors coming from groups, and every time we obtain the intuitive index. In general now, following Jones [55], let us start with:

PROPOSITION 11.4. *Given a subfactor $A_0 \subset A_1$, there is a unique linear map*

$$E : A_1 \to A_0$$

which is positive, unital, trace-preserving and which is such that, for any $a_1, a_2 \in A_0$:

$$E(a_1 b a_2) = a_1 E(b) a_2$$

This map is called conditional expectation from A_1 onto A_0.

PROOF. We make use of the standard representation of the II_1 factor A_1, with respect to its unique trace $tr : A_1 \to \mathbb{C}$, namely:

$$A_1 \subset L^2(A_1)$$

If we denote by Ω the standard cyclic and separating vector of $L^2(A_1)$, we have an identification of vector spaces $A_0\Omega = L^2(A_0)$. Consider now the following projection:

$$e : L^2(A_1) \to L^2(A_0)$$

It follows from definitions that we have an inclusion $e(A_1\Omega) \subset A_0\Omega$. Thus the above projection e induces by restriction a certain linear map, as follows:

$$E : A_1 \to A_0$$

This linear map E and the orthogonal projection e are related by:

$$exe = E(x)e$$

But this shows that the linear map E satisfies the various conditions in the statement, namely positivity, unitality, trace preservation and bimodule property. As for the uniqueness assertion, this follows by using the same argument, applied backwards, the idea being that a map E as in the statement must come from a projection e. □

Following Jones [55], we will be interested in what follows in the orthogonal projection $e : L^2(A_1) \to L^2(A_0)$ producing the expectation $E : A_1 \to A_0$, rather than in E itself:

DEFINITION 11.5. *Associated to any subfactor $A_0 \subset A_1$ is the orthogonal projection*

$$e : L^2(A_1) \to L^2(A_0)$$

producing the conditional expectation $E : A_1 \to A_0$ via the following formula:

$$exe = E(x)e$$

This projection is called Jones projection for the subfactor $A_0 \subset A_1$.

Quite remarkably, the subfactor $A_0 \subset A_1$, as well as its commutant, can be recovered from the knowledge of this projection, in the following way:

PROPOSITION 11.6. *Given a subfactor $A_0 \subset A_1$, with Jones projection e, we have*

$$A_0 = A_1 \cap \{e\}' \quad , \quad A_0' = (A_1' \cap \{e\})''$$

as equalities of von Neumann algebras, acting on the space $L^2(A_1)$.

PROOF. The above two formulae both follow from $exe = E(x)e$, via some elementary computations, and for details here, we refer to Jones' paper [55]. □

We are now ready to formulate a key definition, as follows:

DEFINITION 11.7. *Associated to any subfactor $A_0 \subset A_1$ is the basic construction*

$$A_0 \subset_e A_1 \subset A_2$$

with $A_2 =< A_1, e >$ being the algebra generated by A_1 and by the Jones projection

$$e : L^2(A_1) \to L^2(A_0)$$

acting on the Hilbert space $L^2(A_1)$.

The idea now, following as before Jones [55], will be that the inclusion $A_1 \subset A_2$ appears as a kind of "reflection" of the original inclusion $A_0 \subset A_1$, and also that the basic construction can be iterated, and with all this leading to non-trivial results.

Let us start by further studying the basic construction. We have here:

THEOREM 11.8. *Given a subfactor $A_0 \subset A_1$ having finite index,*

$$[A_1 : A_0] < \infty$$

the basic construction $A_0 \subset_e A_1 \subset A_2$ has the following properties:

(1) $A_2 = JA_0'J$.
(2) $A_2 = \overline{A_1 + A_1eb}$.
(3) A_2 *is a* II_1 *factor.*
(4) $[A_2 : A_1] = [A_1 : A_0]$.
(5) $eA_2e = A_0e$.
(6) $tr(e) = [A_1 : A_0]^{-1}$.
(7) $tr(xe) = tr(x)[A_1 : A_0]^{-1}$, *for any* $x \in A_1$.

PROOF. All this is standard, via some elementary computations, and for details here, in the spirit of those before, we refer again to Jones' paper [55]. □

The above result is quite interesting, potentially leading to some interesting mathematics, so let us perform now twice the basic construction, and see what we get. The result here, which is something more technical, at least at the first glance, is as follows:

PROPOSITION 11.9. *Associated to $A_0 \subset A_1$ is the double basic construction*

$$A_0 \subset_e A_1 \subset_f A_2 \subset A_3$$

with $e : L^2(A_1) \to L^2(A_0)$ and $f : L^2(A_2) \to L^2(A_1)$ having the following properties,

$$fef = [A_1 : A_0]^{-1}f \quad , \quad efe = [A_1 : A_0]^{-1}e$$

in addition to their usual properties, from their corresponding basic constructions.

PROOF. Again, all this is standard, via some elementary computations, in the spirit of those before, and for details here, we refer again to Jones' paper [55]. □

We can in fact perform the basic construction by recurrence, and we obtain:

THEOREM 11.10. *Associated to any subfactor $A_0 \subset A_1$ is the Jones tower*

$$A_0 \subset_{e_1} A_1 \subset_{e_2} A_2 \subset_{e_3} A_3 \subset \ldots\ldots$$

with the Jones projections having the following properties:

(1) $e_i^2 = e_i^* = e_i$.
(2) $e_i e_j = e_j e_i$ *for* $|i - j| \geq 2$.
(3) $e_i e_{i\pm 1} e_i = [A_1 : A_0]^{-1} e_i$.
(4) $tr(we_{n+1}) = [A_1 : A_0]^{-1} tr(w)$, *for any word* $w \in < e_1, \ldots, e_n >$.

PROOF. This follows from Theorem 11.8 and Proposition 11.9, because the triple basic construction does not need in fact any further study. See [55]. □

The relations found in Theorem 11.10 are in fact well-known, from the standard theory of the Temperley-Lieb algebra. This algebra, discovered by Temperley and Lieb in the context of statistical mechanics [80], has a very simple definition, as follows:

DEFINITION 11.11. *The Temperley-Lieb algebra of index $N \in [1, \infty)$ is defined as*

$$TL_N(k) = span(NC_2(k, k))$$

with product given by vertical concatenation, with the rule

$$\bigcirc = N$$

for the closed circles that might appear when concatenating.

In other words, the algebra $TL_N(k)$, depending on parameters $k \in \mathbb{N}$ and $N \in [1, \infty)$, is the linear span of the pairings $\pi \in NC_2(k, k)$. The product operation is obtained by linearity, for the pairings which span $TL_N(k)$ this being the usual vertical concatenation, with the conventions that things go "from top to bottom", and that each circle that might appear when concatenating is replaced by a scalar factor, equal to N.

In what concerns us, we will just need some elementary results. First, we have:

PROPOSITION 11.12. *The Temperley-Lieb algebra $TL_N(k)$ is generated by the diagrams*

$$\varepsilon_1 = \begin{smallmatrix} \cup \\ \cap \end{smallmatrix} \quad , \quad \varepsilon_2 = | \begin{smallmatrix} \cup \\ \cap \end{smallmatrix} \quad , \quad \varepsilon_3 = || \begin{smallmatrix} \cup \\ \cap \end{smallmatrix} \quad , \quad \ldots$$

which are all multiples of projections, in the sense that their rescaled versions

$$e_i = N^{-1} \varepsilon_i$$

satisfy the abstract projection relations $e_i^2 = e_i^* = e_i$.

PROOF. We have two assertions here, the idea being as follows:

(1) The fact that the Temperley-Lieb algebra $TL_N(k)$ is indeed generated by the sequence $\varepsilon_1, \varepsilon_2, \ldots$ follows by drawing pictures, and more specifically by decomposing each basis element $\pi \in NC_2(k, k)$ as a product of such elements ε_i.

(2) Regarding now the projection assertion, when composing ε_i with itself we obtain ε_i itself, times a circle. Thus, according to our multiplication convention, we have:

$$\varepsilon_i^2 = N\varepsilon_i$$

Also, when turning upside-down ε_i, we obtain ε_i itself. Thus, according to our involution convention for the Temperley-Lieb algebra, we have the following formula:

$$\varepsilon_i^* = \varepsilon_i$$

We conclude that the rescalings $e_i = N^{-1}\varepsilon_i$ satisfy $e_i^2 = e_i^* = e_i$, as desired. □

As a second result now, making the link with Theorem 11.10, we have:

PROPOSITION 11.13. *The standard generators $e_i = N^{-1}\varepsilon_i$ of the Temperley-Lieb algebra $TL_N(k)$ have the following properties, where tr is the trace obtained by closing:*

(1) $e_ie_j = e_je_i$ *for* $|i - j| \geq 2$.
(2) $e_ie_{i\pm1}e_i = N^{-1}e_i$.
(3) $tr(we_{n+1}) = N^{-1}tr(w)$, *for any word* $w \in < e_1, \ldots, e_n >$.

PROOF. This follows indeed by doing some elementary computations with diagrams, in the spirit of those performed in the proof of Proposition 11.12. □

With the above results in hand, and still following Jones' paper [55], we can now reformulate Theorem 11.10 into something more conceptual, as follows:

THEOREM 11.14. *Given a subfactor $A_0 \subset A_1$, construct its the Jones tower:*

$$A_0 \subset_{e_1} A_1 \subset_{e_2} A_2 \subset_{e_3} A_3 \subset \ldots\ldots$$

The rescaled sequence of projections $e_1, e_2, e_3, \ldots \in B(H)$ produces then a representation

$$TL_N \subset B(H)$$

of the Temperley-Lieb algebra of index $N = [A_1 : A_0]$.

PROOF. We know from Theorem 11.10 that the rescaled sequence of Jones projections $e_1, e_2, e_3, \ldots \in B(H)$ behaves algebrically exactly as the following TL_N diagrams:

$$\varepsilon_1 = \cap^{\cup} \quad , \quad \varepsilon_2 = |\cap^{\cup} \quad , \quad \varepsilon_3 = ||\cap^{\cup} \quad , \quad \ldots$$

But these diagrams generate TL_N, and so we have an embedding $TL_N \subset B(H)$, where H is the Hilbert space where our subfactor $A_0 \subset A_1$ lives, as claimed. □

Let us make the following key observation, also from [55]:

THEOREM 11.15. *Given a finite index subfactor $A_0 \subset A_1$, the graded algebra $P = (P_k)$ formed by the sequence of higher relative commutants*

$$P_k = A_0' \cap A_k$$

contains the copy of the Temperley-Lieb algebra constructed above, $TL_N \subset P$. This graded algebra $P = (P_k)$ is called "planar algebra" of the subfactor.

PROOF. As a first observation, since the Jones projection $e_1 : A_1 \to A_0$ commutes with A_0, we have $e_1 \in P_2$. By translation we obtain, for any $k \in \mathbb{N}$:

$$e_1, \ldots, e_{k-1} \in P_k$$

Thus we have indeed an inclusion of graded algebras $TL_N \subset P$, as claimed. □

As an interesting consequence of the above results, also from [55], we have:

THEOREM 11.16. *The index of subfactors $A \subset B$ is "quantized" in the $[1, 4]$ range,*

$$N \in \left\{ 4\cos^2\left(\frac{\pi}{n}\right) \middle| n \geq 3 \right\} \cup [4, \infty]$$

with the obstruction coming from the existence of the representation $TL_N \subset B(H)$.

PROOF. This comes from the basic construction, and more specifically from the combinatorics of the Jones projections $e_1, e_2, e_3, \ldots$, the idea being as folows:

(1) In order to best comment on what happens, when iterating the basic construction, let us record the first few values of the numbers in the statement:

$$4\cos^2\left(\frac{\pi}{3}\right) = 1 \quad , \quad 4\cos^2\left(\frac{\pi}{4}\right) = 2$$

$$4\cos^2\left(\frac{\pi}{5}\right) = \frac{3+\sqrt{5}}{2} \quad , \quad 4\cos^2\left(\frac{\pi}{6}\right) = 3$$

$$\ldots$$

(2) When performing a basic construction, we obtain, by trace manipulations on e_1:

$$N \notin (1, 2)$$

With a double basic construction, we obtain, by trace manipulations on $< e_1, e_2 >$:

$$N \notin \left(2, \frac{3+\sqrt{5}}{2}\right)$$

With a triple basic construction, we obtain, by trace manipulations on $< e_1, e_2, e_3 >$:

$$N \notin \left(\frac{3+\sqrt{5}}{2}, 3\right)$$

Thus, we are led to the conclusion in the statement, by a kind of recurrence, involving a certain family of orthogonal polynomials.

(3) In practice now, the most elegant way of proving the result is by using the fundamental fact, explained in Theorem 11.15, that that sequence of Jones projections $e_1, e_2, e_3, \ldots \subset B(H)$ generate a copy of the Temperley-Lieb algebra of index N:

$$TL_N \subset B(H)$$

With this result in hand, we must prove that such a representation cannot exist in index $N < 4$, unless we are in the following special situation:

$$N = 4\cos^2\left(\frac{\pi}{n}\right)$$

But this can be proved by using some suitable trace and positivity manipulations on TL_N, as in (2) above. For full details here, we refer to [55]. □

11b. Planar algebras

Quite remarkably, the planar algebra structure of TL_N, taken in an intuitive sense, of composing diagrams, extends to a planar algebra structure on P. In order to discuss this, let us start with axioms for the planar algebras. Following Jones [61], we have:

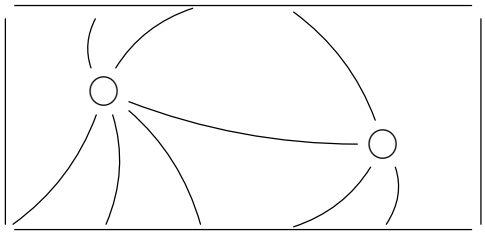
DEFINITION 11.17. *The planar algebras are defined as follows:*

(1) *We consider rectangles in the plane, with the sides parallel to the coordinate axes, and taken up to planar isotopy, and we call such rectangles boxes.*
(2) *A labeled box is a box with $2n$ marked points on its boundary, n on its upper side, and n on its lower side, for some integer $n \in \mathbb{N}$.*
(3) *A tangle is labeled box, containing a number of labelled boxes, with all marked points, on the big and small boxes, being connected by noncrossing strings.*
(4) *A planar algebra is a sequence of finite dimensional vector spaces $P = (P_n)$, together with linear maps $P_{n_1} \otimes \ldots \otimes P_{n_k} \to P_n$, one for each tangle, such that the gluing of tangles corresponds to the composition of linear maps.*

In this definition we are using rectangles, but everything being up to isotopy, we could have used instead circles with marked points, as in [61]. Our choice for using rectangles comes from the main examples that we have in mind, to be discussed below, where the planar algebra structure is best viewed by using rectangles, as above.

This being said, when convenient, we agree to use circles with marked points for the outer box, or for the inner boxes, or for both, with the convention that the marked point is the lower left corner of the rectangle. Here is a planar tangle, drawn in this way, with

the marked points on both circles being by definition those at East:



And, exercise for you to see what this tangle becomes, in rectangular notation.

Let us also mention that Definition 11.17 is something quite simplified, based on [61]. As explained in [61], in order for subfactors to produce planar algebras and vice versa, there are quite a number of supplementary axioms that must be added, and in view of this, it is perhaps better to start with something stronger than Definition 11.17, as basic axioms. However, as before with rectangles vs circles, our axiomatic choices here are mainly motivated by the concrete examples that we have in mind. More on this later.

As a basic example of a planar algebra, we have the Temperley-Lieb algebra:

THEOREM 11.18. *The Temperley-Lieb algebra TL_N, viewed as graded algebra*

$$TL_N = (TL_N(n))_{n\in\mathbb{N}}$$

is a planar algebra, with the corresponding linear maps associated to the planar tangles

$$TL_N(n_1) \otimes \ldots \otimes TL_N(n_k) \to TL_N(n)$$

appearing by putting the various $TL_N(n_i)$ diagrams into the small boxes of the given tangle, which produces a $TL_N(n)$ diagram.

PROOF. This is something trivial, which follows from definitions:

(1) Assume indeed that we are given a planar tangle π, as in Definition 11.17, consisting of a box having $2n$ marked points on its boundary, and containing k small boxes, having respectively $2n_1, \ldots, 2n_k$ marked points on their boundaries, and then a total of $n + \Sigma n_i$ noncrossing strings, connecting the various $2n + \Sigma 2n_i$ marked points.

(2) We want to associate to this tangle π a linear map as follows:

$$T_\pi : TL_N(n_1) \otimes \ldots \otimes TL_N(n_k) \to TL_N(n)$$

For this purpose, by linearity, it is enough to construct elements as follows, for any choice of Temperley-Lieb diagrams $\sigma_i \in TL_N(n_i)$, with $i = 1, \ldots, k$:

$$T_\pi(\sigma_1 \otimes \ldots \otimes \sigma_k) \in TL_N(n)$$

(3) But constructing such an element is obvious, just by putting the various diagrams $\sigma_i \in TL_N(n_i)$ into the small boxes the given tangle π. Indeed, this procedure produces a certain diagram in $TL_N(n)$, that we can call $T_\pi(\sigma_1 \otimes \ldots \otimes \sigma_k)$, as above.

(4) Finally, we have to check that everything is well-defined up to planar isotopy, and that the gluing of tangles corresponds to the composition of linear maps. But both these checks are trivial, coming from the definition of TL_N, and we are done. □

As a conclusion to all this, $P = TL_N$ is indeed a planar algebra, but of somewhat "trivial" type, with the triviality coming from the fact that, in this case, the elements of P are planar diagrams themselves, and so the planar structure appears trivially.

The Temperley-Lieb planar algebra TL_N is however an important planar algebra, because it is the "smallest" one, appearing inside the planar algebra of any subfactor. But more on this later, when talking about planar algebras and subfactors.

Moving ahead now, here is our second basic example of a planar algebra, which is also "trivial" in the above sense, with the elements of the planar algebra being planar diagrams themselves, but which appears in a bit more complicated way:

THEOREM 11.19. *The Fuss-Catalan algebra $FC_{N,M}$, which appears by coloring the Temperley-Lieb diagrams with black/white colors, clockwise, as follows*

$$\circ\bullet\bullet\circ\circ\bullet\bullet\circ\ldots\ldots\ldots\circ\bullet\bullet\circ$$

and keeping those diagrams whose strings connect either $\circ-\circ$ or $\bullet-\bullet$, is a planar algebra, with again the corresponding linear maps associated to the planar tangles

$$FC_{N,M}(n_1) \otimes \ldots \otimes FC_{N,M}(n_k) \to FC_{N,M}(n)$$

appearing by putting the various $FC_{N,M}(n_i)$ diagrams into the small boxes of the given tangle, which produces a $FC_{N,M}(n)$ diagram.

PROOF. The proof here is nearly identical to the proof of Theorem 11.18, with the only change appearing at the level of the colors. To be more precise:

(1) Forgetting about upper and lower sequences of points, which must be joined by strings, a Temperley-Lieb diagram can be thought of as being a collection of strings, say black strings, which compose in the obvious way, with the rule that the value of the circle, which is now a black circle, is N. And it is this obvious composition rule that gives the planar algebra structure, as explained in the proof of Theorem 11.18.

(2) Similarly, forgetting about points, a Fuss-Catalan diagram can be thought of as being a collection of strings, which come now in two colors, black and white. These Fuss-Catalan diagrams compose then in the obvious way, with the rule that the value of the black circle is N, and the value of the white circle is M. And it is this obvious composition rule that gives the planar algebra structure, as before for TL_N. □

Getting back now to generalities, and to Definition 11.17, that of a general planar algebra, we have so far two illustrations for it, which, while both important, are both "trivial", with the planar structure simply coming from the fact that, in both these cases, the elements of the planar algebra are planar diagrams themselves.

In general, the planar algebras can be more complicated than this, and we will see some further examples in a moment. However, the idea is very simple, namely:

PRINCIPLE 11.20. *The elements of a planar algebra are not necessarily diagrams, but they behave like diagrams.*

In relation now with subfactors, the result, which extends Theorem 11.15, and which was found by Jones in [61], almost 20 years after [55], is as follows:

THEOREM 11.21. *Given a subfactor $A_0 \subset A_1$, the collection $P = (P_n)$ of linear spaces*

$$P_n = A_0' \cap A_n$$

has a planar algebra structure, extending the planar algebra structure of TL_N.

PROOF. We know from Theorem 11.15 that we have an inclusion as follows, coming from the basic construction, and with TL_N itself being a planar algebra:

$$TL_N \subset P$$

Thus, the whole point is that of proving that the trivial planar algebra structure of TL_N extends into a planar algebra structure of P. But this can be done via a long algebraic study, and for the full computation here, we refer to Jones' paper [61]. □

As a first illustration for the above result, we have:

THEOREM 11.22. *We have the following universality results:*

(1) *The Temperley-Lieb algebra TL_N appears inside the planar algebra of any subfactor $A_0 \subset A_1$ having index N.*
(2) *The Fuss-Catalan algebra $FC_{N,M}$ appears inside the planar algebra of any subfactor $A_0 \subset A_1$, in the presence of an intermediate subfactor $A_0 \subset B \subset A_1$.*

PROOF. Here the first assertion is something that we already know, from Theorem 11.15, and the second assertion is something quite standard as well, by carefully working out the basic construction for $A_0 \subset A_1$, in the presence of an intermediate subfactor $A_0 \subset B \subset A_1$. For details here, we refer to Bisch and Jones. □

The above results raise the question on whether any planar algebra produces a subfactor. The answer here is yes, but with many subtleties, and in order to talk about this, we first need to introduce a certain distinguished II_1 factor R, as follows:

DEFINITION 11.23. *The Murray-von Neumann hyperfinite* II_1 *factor is*

$$R=\overline{\bigcup_i M_{n_i}(\mathbb{C})}^{w}$$

independently of the choice of the algebras $M_{n_i}(\mathbb{C})$*, and of the embeddings between them.*

Getting back now to subfactors, and to our questions regarding the correspondence between subfactors and planar algebras, these are difficult questions too, and the various answers to these questions can be summarized, a bit informally, as follows:

THEOREM 11.24. *We have the following results:*

(1) *Any planar algebra with positivity produces a subfactor.*
(2) *In particular, we have* TL *and* FC *type subfactors.*
(3) *In the amenable case, and with* $A_1=R$*, the correspondence is bijective.*
(4) *In general, we must take* $A_1=L(F_\infty)$*, and we do not have bijectivity.*
(5) *The axiomatization of* P*, in the case* $A_1=R$*, is not known.*

PROOF. All this is quite heavy, basically coming from the work of Popa in the 90s, using heavy functional analysis, the idea being as follows:

(1) As already mentioned in the comments after Definition 11.17, our planar algebra axioms here are something quite simplified, based on [60]. However, when getting back to Theorem 11.20, the conclusion is that the subfactor planar algebras there satisfy a number of supplementary "positivity" conditions, basically coming from the positivity of the II_1 factor trace. And the point is that, with these positivity conditions axiomatized, we reach to something which is equivalent to Popa's axiomatization of the lattice of higher relative commutants $A_i'\cap A_j$ of the finite index subfactors, obtained in the 90s via heavy functional analysis. For the full story here, and details, we refer to Jones' paper [61].

(2) The existence of the TL_N subfactors, also known as "A_∞ subfactors", is something which was known for some time, since some early work of Popa on the subject. As for the existence of the $FC_{N,M}$ subfactors, this can be shown by using the intermediate subfactor picture, $A_0\subset B\subset A_1$, by composing two A_∞ subfactors of suitable indices, $A_0\subset B$ and $B\subset A_1$. For the full story here, we refer as before to Jones [61].

(3) This is something fairly heavy, as it is always the case with operator algebra results regarding hyperfiniteness and amenability, due to Popa. For the story here, see [61].

(4) This is something a bit more recent, obtained by further building on the above-mentioned constructions of Popa. Again, we refer here to [61] and related work.

(5) This is the big open question in subfactors. The story here goes back to Jones' original paper [55], which contains at the end the question, due to Connes, of finding the possible values of the index for the irreducible subfactors of R. This question, which certainly looks much easier than (5) in the statement, is in fact still open, now 40 years after its formulation, and with no one having any valuable idea in dealing with it. □

In relation now with quantum groups, we have the following result:

THEOREM 11.25. *Let G be a compact quantum group, and $G \to Aut(P)$ be a minimal action on a II_1 factor. Consider a Markov inclusion of finite dimensional algebras*

$$B_0 \subset B_1$$

and let $G \to Aut(B_1)$ be an action which leaves invariant B_0 and which is such that its restrictions to the centers of B_0 and B_1 are ergodic. We have then a subfactor

$$(B_0 \otimes P)^G \subset (B_1 \otimes P)^G$$

of index $N = [B_1 : B_0]$, called generalized Wassermann subfactor, whose Jones tower is

$$(B_1 \otimes P)^G \subset (B_2 \otimes P)^G \subset (B_3 \otimes P)^G \subset \ldots$$

where $\{B_i\}_{i \geq 1}$ are the algebras in the Jones tower for $B_0 \subset B_1$, with the canonical actions of G coming from the action $G \to Aut(B_1)$, and whose planar algebra is given by:

$$P_k = (B_0' \cap B_k)^G$$

These subfactors generalize the Jones, Ocneanu, Wassermann and Popa subfactors.

PROOF. This is something quite standard, and for details here, which are quite technical, we refer to [10] and subsequent papers. Let us also mention, as a complement to what was said above, that the corresponding planar algebra appears as the fixed point algebra, under the action of G, of the bipartite planar algebra associated to the inclusion $B_0 \subset B_1$. For more on this, we refer as before to [10] and subsequent papers. □

As a conclusion, the quantum groups that we met in chapter 9 are closely related to the subfactors and planar algebras that we met in this chapter, via various correspondences. The subfactors and planar algebras remain of course more general objects, but for basic statistical mechanics purposes, the quantum groups from chapter 9 do the job.

11c. Knot invariants

Let us discuss now the relation with knot theory. Knots are something very familiar, from the real life, and mathematically, it is most convenient to define them as follows:

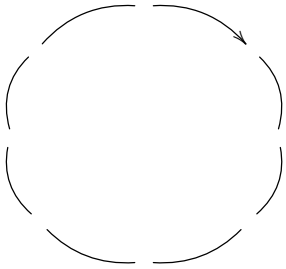
DEFINITION 11.26. *A knot is a smooth closed curve in $\mathbb{R}^3$,*

$$\gamma : \mathbb{T} \to \mathbb{R}^3$$

regarded modulo smooth transformations of $\mathbb{R}^3$.

Observe that our knots are by definition oriented. The reverse knot $z \to \gamma(z^{-1})$ can be isomorphic or not to the original knot $z \to \gamma(z)$, and we will discuss this in a moment.

At the level of examples, we first have the unknot, represented as follows:



The unknot is already a quite interesting mathematical object, suggesting a lot of exciting mathematical questions, as follows:

(1) Given a closed curve in $\mathbb{R}^3$, say given via its algebraic equations, can we decide if it is tied or not? But obviously, this looks like a quite difficult question.

(2) Perhaps simpler, given the 2D picture of a knot, can we decide if it is tied or not? But again, this looks like a quite difficult question.

(3) Experience with cables and ropes shows that a random closed curve is usually tied. But can we really prove this? Once again, difficult question.

As you can see, knot theory is not an easy thing, but do not worry, we will manage to find our way through this jungle, and even come up with some mathematics for it. Going ahead now with the examples, as the simplest possible true knot, meaning tied knot, we have the trefoil knot. And in relation with it, we have the following result:

PROPOSITION 11.27. *The trefoil knot and its opposite knot are not isomorphic. Thus, the trefoil knot comes in fact in two flavors, left-handed, and right-handed.*

PROOF. This is quite clear, by thinking a bit in $\mathbb{R}^3$, and in what regards a formal proof, there is a variety of methods for proving this. We will leave this as an exercise, and we will come back soon with some methods, and knot invariants, doing the job. □

Getting back now to Definition 11.26 as stated, it is convenient to allow, in relation with certain mathematical questions, links in our discussion:

DEFINITION 11.28. *A link is a collection of disjoint knots in $\mathbb{R}^3$, taken as usual oriented, and regarded as usual up to isotopy.*

As before with the knots, which can be truly knotted or not, there is a discussion here with respect to the links, which can be truly linked or not, and with orientation involved too. Drawing some pictures here, with some basic examples, is very instructive, the idea being that two or several basic unknots can be linked in many possible ways.

Meditating a bit at all this, and leaving aside the obvious difficulty of the subject, we won't get scared by such things, obviously, we are here at the core of everything that can be called "geometry". That is, thinking a bit at how knots and links can be tied, in so many fascinating ways, we are led to the following philosophical conclusion:

CONCLUSION 11.29. *Knots and links are to geometry and topology what prime numbers are to number theory.*

Very nice all this, we are now certainly very motivated for studying the knots and links, and time for some mathematics. But the question is, how to get started?

And in view of the above, this is not an easy question. Fortunately, graph theory comes to the rescue, via to the following simple fact:

FACT 11.30. *The plane projection of a knot is something similar to a graph with* 4*-valent vertices, except for the fact that we have some extra data at the vertices, telling us, about the* 2 *strings crossing there, which goes up and which goes down.*

Based on this, let us try to construct some knot invariants. A natural idea is that of defining the invariant on the 2D picture of the knot, that is, on a plane projection of the knot, and then proving that the invariant is indeed independent on the chosen plane.

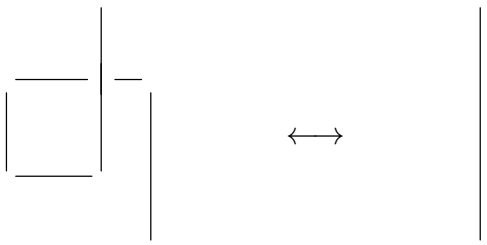
This method rests on the following technical result, which is well-known:

THEOREM 11.31. *Two pictures correspond to plane projections of the same knot precisely when they differ by a sequence of Reidemeister moves, namely:*

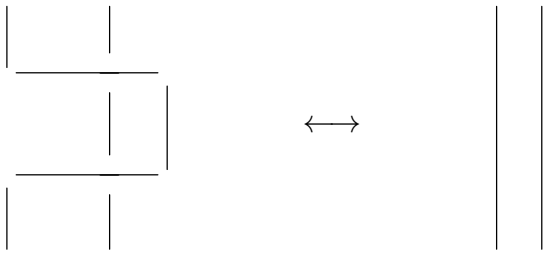
(1) *Moves of type I, given by* $\propto \leftrightarrow |$.
(2) *Moves of type II, given by* $)(\leftrightarrow)($.
(3) *Moves of type III, given by* $\triangle \leftrightarrow \triangledown$.

PROOF. This is something very standard, as follows:

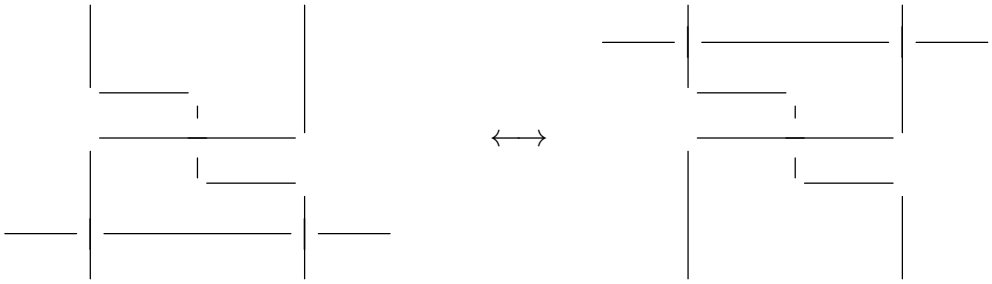
(1) To start with, the Reidemeister moves of type I are by definition as follows:



(2) Regarding the Reidemeister moves of type II, these are by definition as follows:



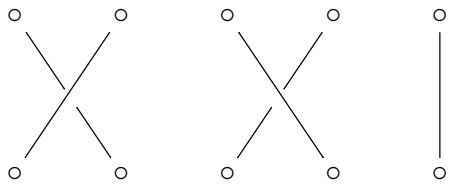
(3) As for the Reidemeister moves of type III, these are by definition as follows:



(4) This was for the precise statement of the theorem, and in what regards now the proof, this is somewhat clear from definitions, and in practice, this can be done by some sort of cut and paste procedure, or recurrence if you prefer, easy exercise for you. □

At a more advanced level now, we will need the following key observation, making the connection with group theory, and algebra in general, due to Alexander:

THEOREM 11.32. *Any knot or link can be thought of as being the closure of a braid,*

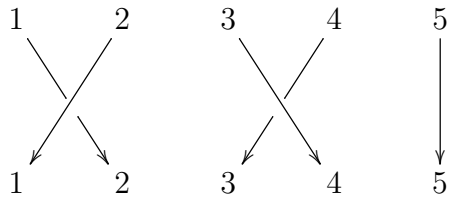


with the braids forming a group B_k, called braid group.

PROOF. Again, this is something quite self-explanatory, as follows:

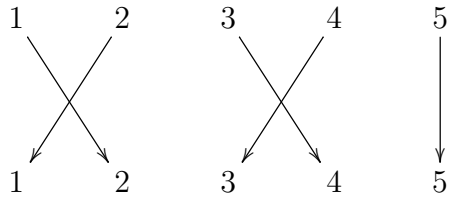
(1) Consider indeed the braids with k strings, with the convention that things go from up to down. With this convention, braids become some sort of permutations of $\{1, \ldots, k\}$, which are decorated at the level of crossings, with for instance the above braid

corresponding to the following permutation of $\{1,2,3,4,5\}$, with due decorations:



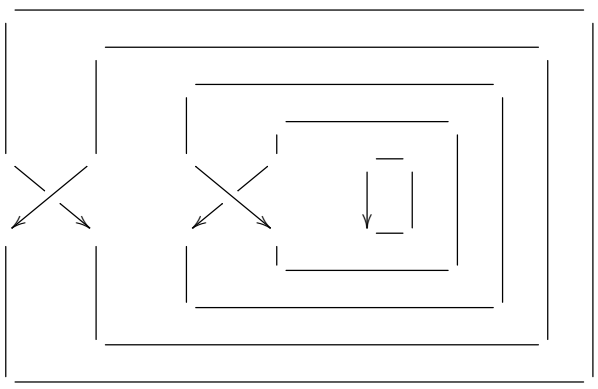
In any case, we can see in this picture that B_k is indeed a group, with composition law similar to that of the permutations in S_k, that is, going from up to down.

(2) Moreover, we can also see in this picture that we have a surjective group morphism $B_k \to S_k$, obtained by forgetting the decorations, at the level of crossings. For instance the braid pictured above is mapped in this way to the following permutation in S_5:



It is possible to do some more algebra here, in relation with the morphism $B_k \to S_k$, but we will not need this in what follows. We will keep in mind, from the above, the fact that "braids are not exactly permutations, but they compose like permutations".

(3) Regarding now the closure operation in the statement, this consists by definition in adding semicircles at right, which makes our braid into a certain oriented link. As an illustration, the closure of the braid pictured above is the following link:



(4) This was for the precise statement of the theorem, and in what regards now the proof, this can be done by some sort of cut and paste procedure, or recurrence if you prefer. As before with Theorem 11.31, we will leave this as an easy exercise for you. □

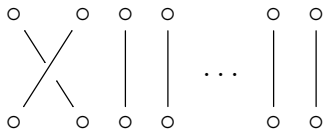
Many interesting things can be said about the braid group B_k, as for instance:

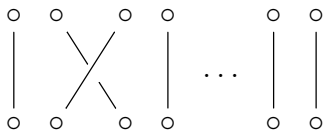
THEOREM 11.33. *The braid group B_k is generated by variables $g_1, \ldots, g_{k-1}$ satisfying*

$$g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1} \quad , \quad g_i g_j = g_j g_i \text{ for } |i-j| \geq 2$$

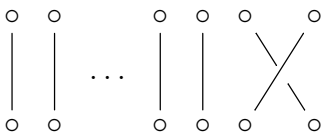
with these generators, called Artin generators, being the basic transpositions.

PROOF. Following Artin, consider indeed the following braids:

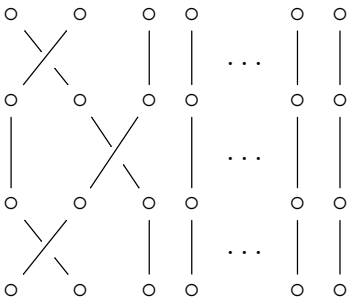
$g_1 =$ 

$g_2 =$ 

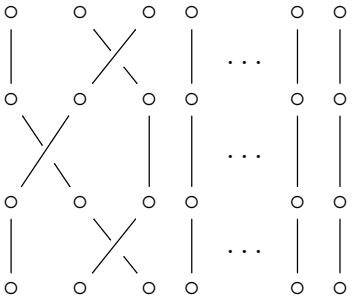
$\vdots$

$g_{k-1} =$ 

We have then $g_i g_j = g_j g_i$, for $|i-j| \geq 2$. As for the relation $g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1}$, by translation it is enough to check this at $i = 1$. And here, we first have:

$g_1 g_2 g_1 =$ 

On the other hand, we have as well the following computation:

$g_2 g_1 g_2 =$ 

Thus, the braid group B_k is indeed generated by elements $g_1, \ldots, g_{k-1}$ with the relations in the statement, and in what regards now the proof of universality, this can only be something quite routine, and we will leave this as an instructive exercise. □

In order to construct invariants, we will need one more theorem, due to Markov:

THEOREM 11.34. *Two elements of the full braid group, obtained as the increasing union of the various braid groups, with embeddings given by $\beta \to \beta\,|$,*

$$B_\infty = \bigsqcup_{k=1}^{\infty} B_k$$

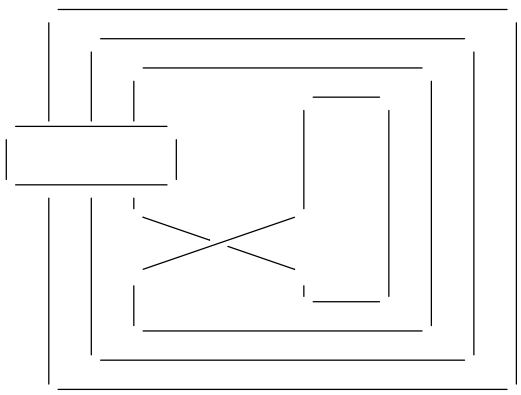
produce the same link, via closing, when one can pass from one to another via:

(1) *Conjugation: $\beta \to \alpha\beta\alpha^{-1}$.*
(2) *Markov move: $\beta \to g_k^{\pm 1}\beta$.*

PROOF. This is a version of the Reidemeister theorem, the idea being as follows:

(1) To start with, it is clear that conjugating a braid, $\beta \to \alpha\beta\alpha^{-1}$, will produce the same link after closing, because we can pull the α, α^{-1} to the right, in the obvious way, and there on the right, these α, α^{-1} will annihilate, according to $\alpha\alpha^{-1} = 1$.

(2) Regarding now the Markov move from the statement, with $\beta \in B_k \subset B_{k+1}$ and with $g_1, \ldots, g_k \in B_{k+1}$ being the standard Artin generators, from Theorem 11.33 and its proof, this is the tricky move, which is worth a proof. Taking $k = 3$ for an illustration, and representing $\beta \in B_3$ by a box, the link obtained by closing $g_4\beta$ is as follows, which is obviously the same link as the one obtained by closing β, and the same goes for $g_4^{-1}\beta$:



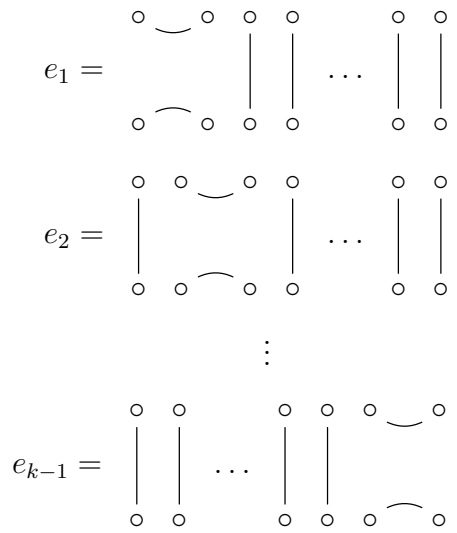
(3) Thus, the links produced by braids are indeed invariant under the two moves in the statement. As for the proof of the converse, this comes from the Reidemeister theorem, applied in the context of the Alexander theorem, or perhaps simpler, by reasoning directly, a bit as in the proof of the Reidemeister theorem. We will leave this as an exercise. □

In order to represent now the braid group, the idea is that of making a connection between braids and Temperley-Lieb diagrams. But this can be done as follows:

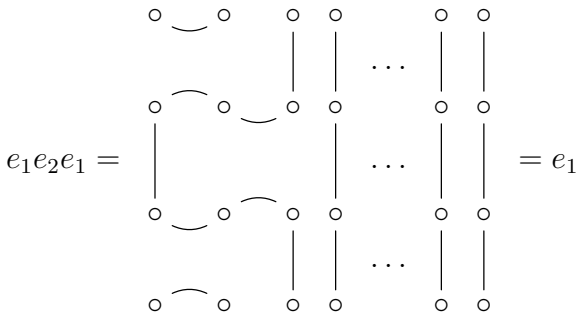
THEOREM 11.35. *The following happen:*

(1) *We have a braid group representation $B_k \to TL_N(k)$, mapping standard generators to standard generators.*
(2) *We have a trace $tr : TL_N(k) \to \mathbb{C}$, obtained by closing the diagrams, which is positive, and has a suitable Markov invariance property.*

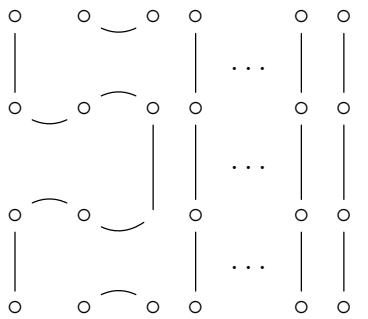
PROOF. The idea here is to map the Artin generators of the braid group to suitable modifications of the following Temperley-Lieb diagrams, called Jones projections:



As a first observation, these diagrams satisfy $e_i^2 = Ne_i$, with $N = \bigcirc$ being as usual the value of the circle, so it is rather the rescaled versions $f_i = e_i/N$ which are projections, but we will not bother with this, and use our terminology above. Next, our Jones projections certainly satisfy the Artin relations $e_ie_j = e_je_i$, for $|i - j| \geq 2$. Our claim now is that is that we have as well the formula $e_ie_{i\pm1}e_i = e_i$. Indeed, by translation it is enough to check $e_ie_{i+1}e_i = e_i$ at $i = 1$, and this follows from the following computation:



As for the verification of the relation $e_2e_1e_2 = e_2$, this is similar, as follows:

$e_2e_1e_2 =$  $= e_2$

Now with the relations $e_ie_{i\pm1}e_i = e_i$ in hand, let us try to reach to the Artin relations $g_ig_{i+1}g_i = g_{i+1}g_ig_{i+1}$. For this purpose, let us set $g_i = te_i - 1$. We have then:

$$\begin{aligned} g_ig_{i+1}g_i &= (te_i - 1)(te_{i+1} - 1)(te_i - 1)\\ &= t^3e_i - t^2(Ne_i + e_ie_{i+1} + e_{i+1}e_i) + t(2e_i + e_{i+1}) - 1\\ &= t(t^2 - Nt + 2)e_i + te_{i+1} - t^2(e_ie_{i+1} + e_{i+1}e_i) \end{aligned}$$

On the other hand, we have as well the following computation:

$$\begin{aligned} g_{i+1}g_ig_{i+1} &= (te_{i+1} - 1)(te_i - 1)(te_{i+1} - 1)\\ &= t^3e_{i+1} - t^2(Ne_{i+1} + e_ie_{i+1} + e_{i+1}e_i) + t(2e_{i+1} + e_i) - 1\\ &= t(t^2 - Nt + 2)e_{i+1} + te_i - t^2(e_ie_{i+1} + e_{i+1}e_i) \end{aligned}$$

Thus with $t^2 - Nt + 1 = 0$ we have a representation $B_k \to TL_N(k)$, as desired. As for the assertions regarding the trace, these can be established, via some combinatorics. □

We can now put everything together, and we obtain, following Jones:

THEOREM 11.36. *We can define the Jones polynomial of an oriented knot or link as being the image of the corresponding braid producing it via the map*

$$tr : B_k \to TL_N(k) \to \mathbb{C}$$

with the following change of variables:

$$N = q^{1/2} + q^{-1/2}$$

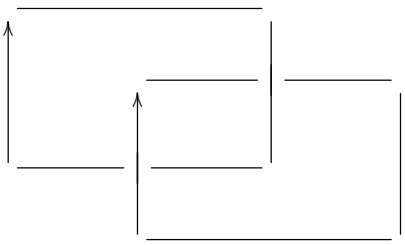
We obtain a Laurent polynomial in $q^{1/2}$, which is an invariant, up to planar isotopy.

PROOF. There is a long story here, the idea being as follows:

(1) To start with, the result follows indeed by combining the above ingredients, the idea being that the various algebraic properties of $tr : TL_N(k) \to \mathbb{C}$ are exactly what is needed for the above composition, up to a normalization, to be invariant under the Reidemeister moves of type I, II, III, and so to produce indeed a knot invariant.

(2) More specifically, the result follows from Theorem 11.34, combined with what we have in Theorem 11.35, and with the change of variables $N = q^{1/2} + q^{-1/2}$ coming from the equation $t^2 - Nt + 1 = 0$ that we found in the proof of Theorem 11.35.

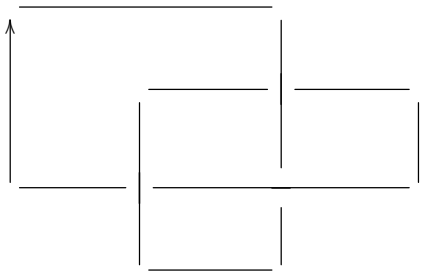
(3) As an illustration for this, let us first look at two linked unknots, namely:



For this link, the corresponding Jones polynomial is given by:

$$V = q^{1/2} + q^{5/2}$$

(4) Let us also look at the trefoil knot, which is as follows:



For this knot, the corresponding Jones polynomial is as follows:

$$V = q + q^3 - q^4$$

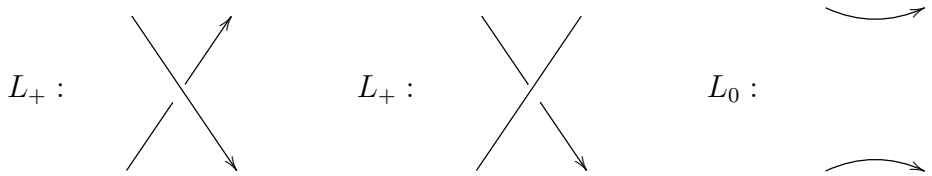
Observe that, as previously for the unknot, this is a Laurent polynomial in q. This is part of a more general phenomenon, the point being that for knots, or more generally for links having an odd number of components, we get a Laurent polynomial in q.

(5) In practice now, far more things can be said, about this. For instance the change of variables $N = q^{1/2} + q^{-1/2}$ in the statement is something well-known in planar algebras, and with all this being related to operator algebras and subfactor theory.

(6) From a purely topological perspective, however, nothing beats the skein relation interpretation of the Jones polynomial $V_L(q)$, which is as follows, with L_+, L_-, L_0 being knots, or rather links, differing at exactly 1 crossing, in the 3 possible ways:

$$q^{-1}V_{L_+} - qV_{L_-} = (q^{1/2} + q^{-1/2})V_{L_0}$$

To be more precise, here are the conventions for L_+, L_-, L_0, that you need to know, in order to play with the above formula, and compute Jones polynomials at wish:



As for the proof of the above formula, this comes from our definition of the Jones polynomial, because thinking well, "unclosing" links as to get braids, and then closing Temperley-Lieb diagrams as to get scalars, as required by the construction of $V_L(q)$, seemingly is some sort of identity operation, but the whole point comes from the fact that the Artin braids $g_1, \ldots, g_{k-1}$ and the Jones projections $e_1, \ldots, e_{k-1}$ differ precisely by a crossing being replaced by a non-crossing. Exercise for you, to figure out all this.

(7) In short, up to you to learn all this, in detail, and its generalizations too, with link polynomials defined more generally via relations of the following type:

$$xP_{L_+} + yP_{L_-} + zP_{L_0} = 0$$

Equivalently, we can define these more general invariants by using various versions of the Temperley-Lieb algebra. As usual, check here the papers of Jones [55], [56], [57].

(8) With the comment here that, among all these invariants, Jones polynomial included, the first came, historially, the Alexander polynomial. However, from a modern point of view, the Alexander polynomial is something more complicated than the Jones polynomial, which remains the central invariant of knots and links. □

11d. Three dimensions

We discuss now some interpretations of the knot invariants constructed in the previous section, and notably of the Jones polynomial, in relation with advanced theoretical physics, or if you prefer, in relation with advanced mathematics. The subject is deep and fascinating, lying at the core of what mathematics, physics, and life in general is, and there are two main discoveries to be discussed, due to Jones and Witten, as follows:

(1) After constructing his polynomial, as explained in the previous section, Jones further built on this, with a remarkable statistical mechanical interpretation of his invariant, and of other similar invariants. The idea is very simple, namely that on the 2D projection on the knot, "interactions happen at crossings", and it is these interactions which produce the knot invariant, as a kind of partition function. By the way, all this was originally motivated by the presence of the Temperley-Lieb algebra in the invariant constructions, because this algebra first appeared, guess where, in statistical mechanics.

(2) The Jones theory, namely construction of his invariant, of some related invariants, and subsequent statistical mechanics interpretation, can be completed with many other things, such as relation with quantum groups, planar algebras, subfactors and more, and is somehow a complete theory. However, this theory remains something 2D, somewhat missing the true 3D nature of the knot. Based on ideas from quantum field theory, Witten fixed this, by finding a 3D formula for the Jones polynomial, and other knot invariants. With this being, again, related to a lot of further modern mathematics and physics.

11e. Exercises

Exercises:

Exercise 11.37.

Exercise 11.38.

Exercise 11.39.

Exercise 11.40.

Exercise 11.41.

Exercise 11.42.

Bonus exercise.

CHAPTER 12

Further models

12a. Pauli models

Following [14], [16], let us discuss now some further examples of stationary models, related to the Pauli matrices, and Weyl matrices. We first discuss, following [14], an explicit model for the quantum group S_4^+, coming from the Pauli matrices, namely:

$$c_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad c_2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad , \quad c_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad , \quad c_4 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Observe that these matrices are elements of SU_2. In fact, SU_2 consists of linear combinations of Pauli matrices, with points on the real sphere S^3 as coefficients:

$$SU(2) = \left\{ \sum x_i c_i \,\middle|\, \sum |x_i|^2 = 1 \right\}$$

The Pauli matrices multiply according to the formulae for quaternions:

$$c_2^2 = c_3^2 = c_4^2 = -1$$

$$c_2 c_3 = -c_3 c_2 = c_4$$

$$c_3 c_4 = -c_4 c_3 = c_2$$

$$c_4 c_2 = -c_2 c_4 = c_3$$

The starting remark is that the Pauli matrices form an orthonormal basis of $M_2(\mathbb{C})$, with respect to the scalar product $< a, b >= tr(ab^*)$. Moreover, the same is true if we multiply them to the left or to the right by an element of SU_2. Thus, we have:

PROPOSITION 12.1. *For any $x \in SU_2$ the elements*

$$\xi_{ij} = c_i x c_j$$

form a magic basis of $M_2(\mathbb{C})$, with respect to the scalar product $< a, b >= tr(ab^*)$.

PROOF. This follows indeed from the above discussion. □

We fix a Hilbert space isomorphism $M_2(\mathbb{C}) \simeq \mathbb{C}^4$, and we use the corresponding identification of operator algebras $B(M_2(\mathbb{C})) \simeq M_4(\mathbb{C})$. Associated to each element $x \in SU_2$ is the representation of $C(S_4^+)$ mapping u_{ij} to the rank one projection on $c_i x c_j$:

$$\pi_x : C(S_4^+) \to M_4(\mathbb{C}) \quad , \quad u_{ij} \to Proj(c_i x c_j)$$

This representation depends on x. For getting a faithful representation, the idea is to regard all these representations as fibers of a single representation:

DEFINITION 12.2. *The Pauli representation of $C(S_4^+)$ is the map*

$$\pi : C(S_4^+) \to C(SU_2, M_4(\mathbb{C}))$$

mapping each standard coodinate u_{ij} to the function $x \to Proj(c_i x c_j)$.

Observe that we can replace here SU_2 by its quotient $PU_2 = SO_3$. For reasons that will become clear later on, we prefer to use SU_2. We will prove now that π is faithful, in the usual sense. Consider the standard linear form on the algebra on the right:

$$\int \varphi = \int_{SU_2} tr\left(\varphi(x)\right)\, dx$$

We use the following analytic formulation of faithfulness:

PROPOSITION 12.3. *The representation π is faithful provided that*

$$\int u_{i_1 j_1} \ldots u_{i_k j_k} = \int \pi_{i_1 j_1} \ldots \pi_{i_k j_k}$$

for any choice of k and of various i, j indices, where $\pi_{ij} = \pi(u_{ij})$.

PROOF. The condition in the statement tells us that, for any product a of generators u_{ij}, the following equality must happen:

$$\int a = \int \pi(a)$$

By linearity and density this formula holds on the whole algebra $C(S_4^+)$. In other words, the following diagram commutes:

$$\begin{array}{ccc} C(S_4^+) & \xrightarrow{\int} & \mathbb{C} \\ \pi \downarrow & & \uparrow tr \\ C\left(SU_2, M_4(\mathbb{C})\right) & \xrightarrow{\int} & M_4(\mathbb{C}) \end{array}$$

On the other hand, we know that the algebra $C(S_4^+)$ is amenable in the discrete quantum group sense, which means that its Haar functional is faithful:

$$a \neq 0 \implies \int aa^* > 0$$

Assume now that we have $\pi(a) = 0$. This implies $\pi(aa^*) = 0$, and the commutativity of the above diagram gives $\int aa^* = 0$. Thus $a = 0$, and we are done. □

Our first goal will be to compute the integral on the right in the above statement. For this purpose, consider the canonical action of SU_2 on the algebra $M_2(\mathbb{C})^{\otimes k}$, obtained as k-th tensor power of the adjoint action on $M_2(\mathbb{C})$, as follows:

$$\alpha_x(a_1 \otimes \ldots \otimes a_k) = xa_1x^* \otimes \ldots \otimes xa_kx^*$$

The following map will play an important role in what follows:

DEFINITION 12.4. *We define a linear map $R : M_2(\mathbb{C})^{\otimes k} \to M_2(\mathbb{C})^{\otimes k}$ by*

$$R(c_{i_1} \otimes \ldots \otimes c_{i_k}) = \frac{1}{2}(c_{i_1}c_{i_2}^* \otimes c_{i_2}c_{i_3}^* \otimes \ldots \otimes c_{i_k}c_{i_1}^*)$$

with the convention that at $k = 1$ we have $R(c_i) = c_ic_i^*/2 = 1/2$.

To any multi-index $i = (i_1, \ldots, i_k)$ we associate the following element:

$$c_i = c_{i_1} \otimes \ldots \otimes c_{i_k}$$

With these notations and conventions, we have the following result:

PROPOSITION 12.5. *We have the following formula,*

$$\int \pi_{i_1j_1} \ldots \pi_{i_kj_k} =< c_i, R^*ER(c_j) >$$

where E is the expectation under the canonical action of SU_2.

PROOF. We have indeed the following computation, where $Proj(\xi)$ denotes as usual the rank one projection onto a vector ξ:

$$\begin{aligned}\int \pi_{i_1j_1} \ldots \pi_{i_kj_k} &= \int \pi(u_{i_1j_1}) \ldots \pi(u_{i_kj_k}) \\ &= \int Proj(c_{i_1}xc_{j_1}) \ldots Proj(c_{i_k}xc_{j_k}) \\ &= \int_{SU_2} tr\left(Proj(c_{i_1}xc_{j_1}) \ldots Proj(c_{i_k}xc_{j_k})\right) dx\end{aligned}$$

We use now the following elementary formula, that we met many times in the above, valid for any sequence of norm one vectors $\xi_1, \ldots, \xi_k$ in a Hilbert space:

$$Tr\left(Proj(\xi_1) \ldots Proj(\xi_k)\right) =< \xi_1, \xi_k >< \xi_k, \xi_{k-1} > \ldots < \xi_2, \xi_1 >$$

In our situation these vectors are in fact matrices, and their scalar products are given by $<\xi,\eta>=tr(\xi\eta^*)$. This gives the following formula:

$$\begin{aligned}\int\pi_{i_1j_1}\ldots\pi_{i_kj_k} &= \frac{1}{4}\int_{SU_2}<c_{i_1}xc_{j_1},c_{i_k}xc_{j_k}>\ldots<c_{i_2}xc_{j_2},c_{i_1}xc_{j_1}>\,dx\\ &= \frac{1}{4}\int_{SU_2}tr(c_{i_1}xc_{j_1}c_{j_k}^*x^*c_{i_k}^*)\ldots tr(c_{i_2}xc_{j_2}c_{j_1}^*x^*c_{i_1}^*)\,dx\\ &= \frac{1}{4}\int_{SU_2}tr(c_{i_k}^*c_{i_1}xc_{j_1}c_{j_k}^*x^*)\ldots tr(c_{i_1}^*c_{i_2}xc_{j_2}c_{j_1}^*x^*)\,dx\end{aligned}$$

We use now the formula $c_s^*=\pm c_s$, valid for all the Pauli matrices c_s. The minus signs can be rearranged, and the computation can be continued as follows:

$$\begin{aligned}\int\pi_{i_1j_1}\ldots\pi_{i_kj_k} &= \frac{1}{4}\int_{SU_2}tr(c_{i_k}c_{i_1}^*xc_{j_1}c_{j_k}^*x^*)\ldots tr(c_{i_1}c_{i_2}^*xc_{j_2}c_{j_1}^*x^*)\,dx\\ &= \frac{1}{4}\int_{SU_2}tr\left(c_{i_1}c_{i_2}^*xc_{j_2}c_{j_1}^*x^*\otimes\ldots\otimes c_{i_k}c_{i_1}^*xc_{j_1}c_{j_k}^*x^*\right)\,dx\\ &= \int_{SU_2}tr\left(R(c_i)\alpha_x(R(c_j)^*)\right)\,dx\end{aligned}$$

We can interchange the trace and integral signs, and we obtain:

$$\begin{aligned}\int\pi_{i_1j_1}\ldots\pi_{i_kj_k} &= tr\left(\int_{SU_2}R(c_i)\alpha_x(R(c_j)^*)\,dx\right)\\ &= tr\left(R(c_i)\int_{SU_2}\alpha_x(R(c_j)^*)\,dx\right)\end{aligned}$$

Now since acting by group elements, then integrating, is the same as projecting onto fixed points, the computation can be continued as follows:

$$\begin{aligned}\int\pi_{i_1j_1}\ldots\pi_{i_kj_k} &= tr\left(R(c_i)ER(c_j)^*\right)\\ &= <R(c_i),ER(c_j)>\\ &= <c_i,R^*ER(c_j)>\end{aligned}$$

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

We know that the faithfulness of the Pauli representation is equivalent to a certain equality of integrals. Moreover, one of these integrals can be computed by using the Weingarten formula. As for the other integral, this can be computed as well, provided that we have enough information about the operator R from Definition 12.4.

We work out now a number of technical properties of this operator R. For this purpose, we first need to understand some aspects of the structure of the algebra of fixed points under the diagonal adjoint action of SU_2. We have the following result, to start with:

PROPOSITION 12.6. *The element*

$$f = \sum_{i=1}^{4} c_i \otimes c_i^*$$

is invariant under the action of SU_2.

PROOF. We have $c_i = -c_i^*$ for $i = 2, 3, 4$, so the element in the statement is:

$$f = 2(c_1 \otimes c_1) - \sum_{i=1}^{4} c_i \otimes c_i$$

Since $c_1 \otimes c_1$ is invariant under SU_2, what is left to prove is that the sum on the right is invariant as well. This sum, viewed as a matrix, is:

$$\sum_{i=1}^{4} c_i \otimes c_i = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

But this matrix is 4 times the projection onto the determinant subspace of $\mathbb{C}^2 \otimes \mathbb{C}^2$, which is invariant under the action of SU_2, and this concludes the proof. □

Given a noncrossing partition $p \in NC(k)$ and a multi-index $j = (j_1, \ldots, j_k)$, we can plug j into p in the obvious way, and we define a Kronecker type symbol $\delta_{pj} \in \{0, 1\}$ by setting $\delta_{pj} = 1$ if any block of p contains identical indices of j, and $\delta_{pj} = 0$ otherwise. To any $p \in NC(k)$ we associate an element of $M_2(\mathbb{C})^k$, as follows:

$$c_p = \sum_j \delta_{pj}\, c_j$$

Our next goal is to prove that $R(c_p)$ is invariant under the action of SU_2. We discuss first the case of the trivial partition, $0_k = \{\{1\}, \ldots, \{k\}\}$. We have:

PROPOSITION 12.7. *$R(c_{0_k})$ is invariant under the action of SU_2.*

PROOF. The partition $p = 0_k$ has the particular property that $\delta_{pj} = 1$ for any multi-index $j = (j_1, \ldots, j_k)$. This gives the following formula:

$$\begin{aligned} R(c_{0_k}) &= R\left(\sum_j c_j\right) \\ &= R\left(\sum_{j_1 \ldots j_k} c_{j_1} \otimes \ldots \otimes c_{j_k}\right) \\ &= \frac{1}{2} \sum_{j_1 \ldots j_k} c_{j_1} c_{j_2}^* \otimes c_{j_2} c_{j_3}^* \otimes \ldots \otimes c_{j_k} c_{j_1}^* \end{aligned}$$

In ordert to compute the above quantity, to any multi-index $j = (j_1, \ldots, j_k)$ we associate a multi-index $i = (i_1, \ldots, i_{k-1})$ in the following way:

– i_1 is such that $c_{i_1} = \pm c_{j_1} c_{j_2}^*$,

– then i_2 is such that $c_{i_2} = \pm c_{j_1} c_{j_3}^*$,

$\vdots$

– and so on up to i_{k-1}, which is such that $c_{i_{k-1}} = \pm c_{j_1} c_{j_k}^*$.

With this notation, we have the following formulae, where the possible dependences between the various $\pm$ signs are not taken into account:

$$\begin{aligned} c_{i_1} &= \pm c_{j_1} c_{j_2}^* \\ c_{i_1}^* c_{i_2} &= (\pm c_{j_1} c_{j_2}^*)^* (\pm c_{j_1} c_{j_3}^*) = \pm c_{j_2} c_{j_3}^* \\ c_{i_2}^* c_{i_3} &= (\pm c_{j_1} c_{j_3}^*)^* (\pm c_{j_1} c_{j_4}^*) = \pm c_{j_3} c_{j_4}^* \\ &\vdots \\ c_{i_{k-2}}^* c_{i_{k-1}} &= (\pm c_{j_1} c_{j_{k-1}}^*)^* (\pm c_{j_1} c_{j_k}^*) = \pm c_{j_{k-1}} c_{j_k}^* \\ c_{i_{k-1}}^* &= (\pm c_{j_1} c_{j_k}^*)^* = \pm c_{j_k} c_{j_1}^* \end{aligned}$$

By taking the tensor product of all these formulae, we get:

$$c_{i_1} \otimes c_{i_1}^* c_{i_2} \otimes \ldots \otimes c_{i_{k-2}}^* c_{i_{k-1}} \otimes c_{i_{k-1}}^* = \pm c_{j_1} c_{j_2}^* \otimes c_{j_2} c_{j_3}^* \otimes \ldots \otimes c_{j_k} c_{j_1}^*$$

By applying the linear map given by $a_1 \otimes \ldots \otimes a_k \to a_1 \ldots a_k$ to both sides we see that the sign on the right is actually $+$. That is, we have:

$$c_{i_1} \otimes c_{i_1}^* c_{i_2} \otimes \ldots \otimes c_{i_{k-2}}^* c_{i_{k-1}} \otimes c_{i_{k-1}}^* = c_{j_1} c_{j_2}^* \otimes c_{j_2} c_{j_3}^* \otimes \ldots \otimes c_{j_k} c_{j_1}^*$$

We recognize at right the basic summand in the formula of $R(c_{0_k})$. Now it follows from the definition of i that summing the right terms over all multi-indices $j = (j_1, \ldots, j_k)$ is the same as summing the left terms over all multi-indices $i = (i_1, \ldots, i_{k-1})$, then multiplying

by 4. Thus, we obtain the following formula:

$$\begin{aligned}
R(c_{0_k}) &= \frac{4}{2}\sum_{i_1\ldots i_{k-1}} c_{i_1}\otimes c_{i_1}^*c_{i_2}\otimes\ldots\otimes c_{i_{k-2}}^*c_{i_{k-1}}\otimes c_{i_{k-1}}^* \\
&= 2\sum_{i_1\ldots i_{k-1}} (c_{i_1}\otimes c_{i_1}^*)_{12}(c_{i_2}\otimes c_{i_2}^*)_{23}\ldots(c_{i_{k-1}}\otimes c_{i_{k-1}}^*)_{k-1,k} \\
&= 2\left(\sum_{i_1} c_{i_1}\otimes c_{i_1}^*\right)_{12}\left(\sum_{i_2} c_{i_2}\otimes c_{i_2}^*\right)_{23}\ldots\left(\sum_{i_{k-1}} c_{i_{k-1}}\otimes c_{i_{k-1}}^*\right)_{k-1,k} \\
&= 2f_{12}f_{23}\ldots f_{k-1,k}
\end{aligned}$$

Now f being invariant under SU_2, the same happens for each $f_{i,i+1}$. Since the invariants form an algebra, the above product is invariant, and this concludes the proof. □

We prove now that the previous result holds not only for c_0, but in fact for any c_p. We first introduce some standard combinatorial notations. The Kreweras complement of a partition $p\in NC(k)$ is constructed as follows. Consider the ordered set $\{1,\ldots,k\}$. At right of each index i we put an index i', as to get the following sequence of indices:

$$1,1',2,2',\ldots,k,k'$$

The Kreweras complement p^c is then the largest noncrossing partition of the new index set $\{1',\ldots,k'\}$, such that the union of p and p^c is noncrossing. We have:

DEFINITION 12.8. *For a noncrossing partition p we use the notation*

$$w(p)=2\prod_{i=1}^{l}\left(\prod_{p=1}^{k_i-1} f_{j_{i,p}j_{i,p+1}}\right)$$

where $\{j_{11}<\ldots<j_{1k_1}\},\ldots,\{j_{l1}<\ldots<j_{lk_l}\}$ are the blocks of p, with the convention that for $k_i=1$ the product on the right is by definition 1.

Here we use the element f introoduced above, and the leg-numbering notation. Observe that the element $w(p)$ is well-defined, because the products on the right pairwise commute. In particular in the case of the partition $1_k=\{\{1,\ldots,k\}\}$, we have:

$$w(1_k)=2\prod_{i=1}^{k-1} f_{i,i+1}$$

Since each f_{ij} is invariant under the adjoint action of SU_2 and since the invariants form an algebra, $w(p)$ is also invariant under the adjoint action of SU_2. We have:

PROPOSITION 12.9. *$R(c_p)$, with $p\in NC(k)$, is invariant under the action of SU_2.*

PROOF. The idea is to generalize the proof of the previous result. In particular we need to generalize the key formula there, namely:

$$R(c_{0_k}) = 2f_{12}f_{23}\ldots f_{k-1,k}$$

Our claim is that for any noncrossing partition p, we have:

$$R(c_p) = w(p^c)$$

As a first verification, for the trivial partition $p = 0_k$ we have $p^c = 1_k$, and the formula $R(c_p) = w(p^c)$ follows from the above identities. Also, for the rough partition $p = 1_k$ we have $p^c = 0_k$, and the claimed equality follows from:

$$\begin{aligned} R(c_{1_k}) &= \sum_i R(c_i \otimes \ldots \otimes c_i) \\ &= \frac{1}{2}\sum_i c_i c_i^* \otimes \ldots \otimes c_i c_i^* \\ &= \frac{1}{2}\,4(1 \otimes \ldots \otimes 1) \\ &= w(0_k) \end{aligned}$$

In the general case, we can use the same method as before. We have:

$$\begin{aligned} R(c_p) &= R\left(\sum_j \delta_{pj} c_j\right) \\ &= R\left(\sum_{j_1\ldots j_k} \delta_{pj} c_{j_1} \otimes \ldots \otimes c_{j_k}\right) \\ &= \frac{1}{2}\sum_{j_1\ldots j_k} \delta_{pj} c_{j_1} c_{j_2}^* \otimes c_{j_2} c_{j_3}^* \otimes \ldots \otimes c_{j_k} c_{j_1}^* \end{aligned}$$

As before, in the previous proof, to any multi-index $j = (j_1, \ldots, j_k)$ we associate a multi-index $i = (i_1, \ldots, i_k)$, in the following way:

– i_1 is such that $c_{i_1} = \pm c_{j_1} c_{j_2}^*$,

– i_2 is such that $c_{i_2} = \pm c_{j_1} c_{j_3}^*$,

$\vdots$

– and so on up to i_k, which is such that $c_{i_k} = \pm c_{j_k} c_{j_1}^*$.

With this notation, and with the observation that the product of all the above $\pm$ signs is actually $+$, the formula that we established before becomes:

$$R(c_p) = \frac{1}{2}\sum_j \delta_{pj} c_{i_1} \otimes \ldots \otimes c_{i_k}$$

Now let $l \in \{1, \ldots, k\}$ and assume that it is a last element of a block of p^c. Let $l_1 < \ldots < l_x = l$ be the ordered enumeration of the elements of this block. We have:

$$c_{j_{l_1}} \ldots c_{j_{l_x}} = 1$$

As an illustration here, if l and $l+1$ are in the same block of p, then $\{l\}$ is a one-element block in p^c and i_l will be such that $c_{i_l} = \pm c_{j_l} c_{j_{l+1}}^* = c_{j_l} c_{j_l}^* = 1$. The general case works by following the same argument as before. We obtain:

$$\sum_i \delta_{pi} c_{j_{l_1}} \otimes \ldots \otimes c_{j_{l_x}} = R(c_{0_x})$$

The point now is that the element $R(c_p)$ is the product of the above expressions, over the blocks of p^c. If we denote these blocks by $\{l_{11} < \ldots < l_{1x_1}\}$, $\ldots$, $\{l_{r1} < \ldots < l_{rx_r}\}$, we obtain, by using the leg-numbering notation, the following formula:

$$\begin{aligned} R(c_p) &= 2^{1-b} \sum_j \prod_{b=1}^{r} R(c_{0_{x_b}})_{l_{b1} \ldots l_{bx_b}} \\ &= 2^{1-b} \sum_j \prod_{b=1}^{r} \omega(1_{x_b})_{l_{b1} \ldots l_{bx_b}} \\ &= w(p^c) \end{aligned}$$

All this might seem a bit complicated, and to illustrate the above proof let us perform explicitly the computation in the case of the following partition:

$$p = \{\{1,5\},\{2\},\{3,4\},\{6\}\}$$

The Kreweras complement of this partition is then $p^c = \{\{1,2,4\},\{3\},\{5,6\}\}$, and the above method gives the following formula:

$$\begin{aligned} R(c_p) &= \frac{1}{2} \sum_{j_1 j_2 j_3 j_4} c_{j_1} c_{j_2^*} \otimes c_{j_2} c_{j_3}^* \otimes 1 \otimes c_{j_3} c_{j_1}^* \otimes c_{j_1} c_{j_4}^* \otimes c_{j_4} c_{j_1}^* \\ &= \frac{1}{2} 4 \sum_{i_1 i_2 i_3} c_{i_1} \otimes c_{i_1}^* c_{i_2} \otimes 1 \otimes c_{i_2}^* \otimes c_{i_3} \otimes c_{i_3}^* \\ &= 2 \sum_{i_1 i_2 i_3} (c_{i_1} \otimes c_{i_1}^* c_{i_2} \otimes 1 \otimes c_{i_2}^* \otimes 1 \otimes 1)(1 \otimes 1 \otimes 1 \otimes 1 \otimes c_{i_3} \otimes c_{i_3}^*) \\ &= 2(f_{12} f_{24}) f_{56} \\ &= w(p^c) \end{aligned}$$

Now back to the general case, the formula $R(c_p) = w(p^c)$ that we established shows that $R(c_p)$ is a product of certain elements f_{ij} obtained from f by acting on the various legs of $M_2(\mathbb{C})^{\otimes k}$, and we can conclude as in the proof of the previous result. □

We are now in position of proving our main technical result, as follows:

PROPOSITION 12.10. $R^*ER(c_p) = c_p$.

PROOF. This comes from a routine computation, as follows:

(1) To start with, the previous result gives the following formula:

$$\begin{aligned} < R^*ER(c_p), c_j > &= < R^*R(c_p), c_j > \\ &= < R(c_p), R(c_j) > \\ &= \sum_i \delta_{pi} < R(c_i), R(c_j) > \end{aligned}$$

(2) On the other hand, we have from definitions the following formula:

$$< c_p, c_j >= \delta_{pj}$$

Since the elements c_j span the ambient space, what is left to prove is:

$$\sum_i \delta_{pi} < R(c_i), R(c_j) >= \delta_{pj}$$

(3) But this can be checked by direct computation. Indeed, from the definition of the operator R, we get the following formula, for the above scalar products:

$$\begin{aligned} < R(c_i), R(c_j) > &= \frac{1}{4} < c_{i_1}c_{i_2}^* \otimes \ldots \otimes c_{i_k}c_{i_1}^*, c_{j_1}c_{j_2}^* \otimes \ldots \otimes c_{j_k}c_{j_1}^* > \\ &= \frac{1}{4} < c_{i_1}c_{i_2}^*, c_{j_1}c_{j_2}^* > \ldots < c_{i_k}c_{i_1}^*, c_{j_k}c_{j_1}^* > \end{aligned}$$

(4) In this latter formula all scalar products are 0, 1 and -1. Now assume that the indices $i_1, \ldots, i_k$ and j_1 are all fixed. If the first scalar product is ± 1 then j_2 is uniquely determined, then if the second scalar product is ± 1 then j_3 is uniquely determined as well, and so on. Thus for all scalar products to be ± 1, the multi-index j is uniquely determined by the multi-index i, up to a possible choice of the first index j_1.

(5) Moreover, each choice of the index j_1 leads to a multi-index $j = (j_1, \ldots, j_k)$, such that all the scalar products are ± 1. Indeed, once $j_2, \ldots, j_k$ are chosen as to satisfy $c_{i_r}c_{i_{r+1}}^* = \pm c_{j_r}c_{j_{r+1}}^*$ for $r = 1, \ldots, k-1$, by multiplying all these formulae we obtain $c_{i_1}c_{i_k}^* = \pm c_{j_1}c_{j_k}^*$, which shows that the last scalar product is ± 1 as well.

(6) Summarizing, given a multi-index $i = (i_1, \ldots, i_k)$ and a number $s \in \{1, 2, 3, 4\}$, there is a unique multi-index $j = (j_1, \ldots, j_k)$ with $j_1 = s$, such that all the above scalar products are ± 1. We use the notation $j = i \oplus s$, for this multi-index.

(7) In the situation $j = i \oplus s$ we have $< R(c_i), R(c_j) >= \pm 1$, and by applying the linear map given by $a_1 \otimes \ldots \otimes a_k \to a_1 \ldots a_k$ to the formula $c_{i_1}c_{i_2}^* \otimes \ldots \otimes c_{i_k}c_{i_1}^* = \pm c_{j_1}c_{j_2}^* \otimes \ldots \otimes c_{j_k}c_{j_1}^*$ we see that the sign is actually $+$. Thus we have the following formula:

$$< R(c_i), R(c_j) >= \begin{cases} 1/4 \text{ if } j = i \oplus s \text{ for some } s \in \{1, 2, 3, 4\} \\ 0 \text{ otherwise} \end{cases}$$

(8) We can come back now to the missing formula. We have:

$$\sum_i \delta_{pi} < R(c_i), R(c_j) >= \frac{1}{4}\sum_{s=1}^{4} \delta_{pi}$$

Our claim is that we have $\delta_{pi} = \delta_{pj}$, for any partition p. Indeed, this follows from the fact that for $r < s$ we have $i_r = i_s$ if and only if the product of $c_{i_t} c_{i_{t+1}}^*$ over $t = r, \ldots, s-1$ equals ± 1, and a similar statement holds for the multi-index j.

(9) We can now conclude the proof. By using $\delta_{pi} = \delta_{pj}$, we get:

$$\sum_i \delta_{pi} < R(c_i), R(c_j) >= \frac{1}{4}\sum_{s=1}^{4} \delta_{pj} = \delta_{pj}$$

But this is the formula that we wanted to prove, so we are done. □

Still following [14], we can now prove our main result, as follows:

THEOREM 12.11. *The Pauli representation of $C(S_4^+)$ is faithful.*

PROOF. We denote as usual by $c_1, \ldots, c_4$ the Pauli matrices, and we let $e_1, \ldots, e_4$ be the standard basis of $\mathbb{C}^4$. For a multi-index $i = (i_1, \ldots, i_k)$, we set:

$$e_i = e_{i_1} \otimes \ldots \otimes e_{i_k} \quad , \quad c_i = c_{i_1} \otimes \ldots \otimes c_{i_k}$$

Each partition $p \in NC(k)$ creates two tensors, in the following way:

$$e_p = \sum_i \delta_{pi}\, e_i \quad , \quad c_p = \sum_i \delta_{pi}\, c_i$$

Consider the following $4^k \times 4^k$ matrices, with entries labeled by multi-indices i, j:

$$P_{ij} = tr\left(\int \pi_{i_1 j_1} \ldots \pi_{i_k j_k}\right) \quad , \quad U_{ij} = \left(\int u_{i_1 j_1} \ldots u_{i_k j_k}\right)$$

According to Proposition 12.5, we have the following formula:

$$P_{ij} =< c_i, R^* E R(c_j) >$$

Now let Φ be the linear map $(\mathbb{C}^4)^{\otimes k} \to M_2(\mathbb{C})^{\otimes k}$ given by $\Phi(e_i) = c_i$. In terms of this map, the above equation can be rephrased as follows:

$$P_{ij} =< e_i, \Phi^* R^* E R \Phi(e_j) >$$

According to our results above, it is enough to prove that we have $P = U$. But U is the orthogonal projection of $(\mathbb{C}^4)^{\otimes k}$ onto the following space:

$$S_e = span\left\{e_p \middle| p \in NC(k)\right\}$$

Thus what we have to prove is that $\Phi^* R^* ER\Phi$ is the projection onto S_e. But this is equivalent to proving that R^*ER is the projection onto the following space:

$$S_c = span\left\{c_p\Big|p \in NC(k)\right\}$$

We know that $R^*ER(c_p) = c_p$, for all $p \in NC(k)$. But this implies that R^*ER restricted to S_c is the identity. This vector space has dimension the Catalan number C_k, and this is exactly the rank of the operator E. Therefore S_c has to be the image of R^*ER. Now since R^*ER is self-adjoint, and is the identity on its image, it is the orthogonal projection onto S_c, and this concludes the proof. □

12b. Weyl models

Following [16], we discuss now a generalization of the above models. We first have:

DEFINITION 12.12. *Given a finite abelian group H, the associated Weyl matrices are*

$$W_{ia} : e_b \to < i, b > e_{a+b}$$

where $i \in H$, $a, b \in \widehat{H}$, and where $(i, b) \to < i, b >$ is the Fourier coupling $H \times \widehat{H} \to \mathbb{T}$.

As a basic example, consider the cyclic group $H = \mathbb{Z}_2 = \{0, 1\}$. Here the Fourier coupling is given by $< i, b >= (-1)^{ib}$, and so the Weyl matrices act via:

$$W_{00} : e_b \to e_b \quad , \quad W_{10} : e_b \to (-1)^b e_b$$

$$W_{11} : e_b \to (-1)^b e_{b+1} \quad , \quad W_{01} : e_b \to e_{b+1}$$

Thus, we have the following formulae:

$$W_{00} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad W_{10} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$W_{11} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad , \quad W_{01} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

We recognize here, up to some multiplicative factors, the four Pauli matrices. Now back to the general case, we have the following well-known result:

PROPOSITION 12.13. *The Weyl matrices are unitaries, and satisfy:*

(1) $W_{ia}^* =< i, a > W_{-i,-a}$.
(2) $W_{ia}W_{jb} =< i, b > W_{i+j,a+b}$.
(3) $W_{ia}W_{jb}^* =< j - i, b > W_{i-j,a-b}$.
(4) $W_{ia}^*W_{jb} =< i, a - b > W_{j-i,b-a}$.

PROOF. The unitarity follows from (3,4), and the rest of the proof goes as follows:

(1) We have indeed the following computation:

$$\begin{aligned} W_{ia}^* &= \left(\sum_b < i, b > E_{a+b,b}\right)^* \\ &= \sum_b < -i, b > E_{b,a+b} \\ &= \sum_b < -i, b-a > E_{b-a,b} \\ &= < i, a > W_{-i,-a} \end{aligned}$$

(2) Here the verification goes as follows:

$$\begin{aligned} W_{ia}W_{jb} &= \left(\sum_d < i, b+d > E_{a+b+d,b+d}\right)\left(\sum_d < j, d > E_{b+d,d}\right) \\ &= \sum_d < i, b >< i+j, d > E_{a+b+d,d} \\ &= < i, b > W_{i+j,a+b} \end{aligned}$$

(3,4) By combining the above two formulae, we obtain:

$$\begin{aligned} W_{ia}W_{jb}^* &= < j, b > W_{ia}W_{-j,-b} \\ &= < j, b >< i, -b > W_{i-j,a-b} \end{aligned}$$

We obtain as well the following formula:

$$\begin{aligned} W_{ia}^*W_{jb} &= < i, a > W_{-i,-a}W_{jb} \\ &= < i, a >< -i, b > W_{j-i,b-a} \end{aligned}$$

But this gives the formulae in the statement, and we are done. □

With $n = |H|$, we can use an isomorphism $l^2(\widehat{H}) \simeq \mathbb{C}^n$ as to view each W_{ia} as a usual matrix, $W_{ia} \in M_n(\mathbb{C})$, and hence as a usual unitary, $W_{ia} \in U_n$. Also, given a vector ξ, we denote by $Proj(\xi)$ the orthogonal projection onto $\mathbb{C}\xi$. Following [16], we have:

PROPOSITION 12.14. *Given a closed subgroup $E \subset U_n$, we have a representation*

$$\pi_H : C(S_N^+) \to M_N(C(E))$$

$$w_{ia,jb} \to [U \to Proj(W_{ia}UW_{jb}^*)]$$

where $n = |H|, N = n^2$, and where W_{ia} are the Weyl matrices associated to H.

PROOF. The Weyl matrices being given by $W_{ia} : e_b \to < i, b > e_{a+b}$, we have:

$$tr(W_{ia}) = \begin{cases} 1 & \text{if } (i,a) = (0,0) \\ 0 & \text{if } (i,a) \neq (0,0) \end{cases}$$

Together with the formulae in Proposition 12.13, this shows that the Weyl matrices are pairwise orthogonal with respect to the following scalar product on $M_n(\mathbb{C})$:

$$< x, y >= tr(xy^*)$$

Thus, these matrices form an orthogonal basis of $M_n(\mathbb{C})$, consisting of unitaries:

$$W = \left\{ W_{ia} \middle| i \in H, a \in \widehat{H} \right\}$$

Thus, each row and each column of the matrix $\xi_{ia,jb} = W_{ia} U W_{jb}^*$ is an orthogonal basis of $M_n(\mathbb{C})$, and so the corresponding projections form a magic unitary, as claimed. □

We will need the following well-known result:

PROPOSITION 12.15. *With $T = Proj(x_1) \ldots Proj(x_p)$ and $||x_i|| = 1$ we have*

$$< T\xi, \eta >=< \xi, x_p >< x_p, x_{p-1} > \ldots < x_2, x_1 >< x_1, \eta >$$

for any ξ, η. In particular, we have:

$$Tr(T) =< x_1, x_p >< x_p, x_{p-1} > \ldots < x_2, x_1 >$$

PROOF. For $||x|| = 1$ we have $Proj(x)\xi =< \xi, x > x$. This gives:

$$\begin{aligned} T\xi &= Proj(x_1) \ldots Proj(x_p)\xi \\ &= Proj(x_1) \ldots Proj(x_{p-1}) < \xi, x_p > x_p \\ &= Proj(x_1) \ldots Proj(x_{p-2}) < \xi, x_p >< x_p, x_{p-1} > x_{p-1} \\ &= \ldots \\ &= < \xi, x_p >< x_p, x_{p-1} > \ldots < x_2, x_1 > x_1 \end{aligned}$$

Now by taking the scalar product with η, this gives the first assertion. As for the second assertion, this follows from the first assertion, by summing over $\xi = \eta = e_i$. □

Now back to the Weyl matrix models, let us first compute T_p. We have:

PROPOSITION 12.16. *We have the formula*

$$\begin{aligned} (T_p)_{ia,jb} &= \frac{1}{N} < i_1, a_1 - a_p > \ldots < i_p, a_p - a_{p-1} >< j_1, b_1 - b_2 > \ldots < j_p, b_p - b_1 > \\ &\int_E tr(W_{i_1 - i_2, a_1 - a_2} U W_{j_2 - j_1, b_2 - b_1} U^*) \ldots tr(W_{i_p - i_1, a_p - a_1} U W_{j_1 - j_p, b_1 - b_p} U^*) dU \end{aligned}$$

with all the indices varying in a cyclic way.

PROOF. By using the trace formula in Proposition 12.15, we obtain:

$$\begin{aligned}
&(T_p)_{ia,jb}\\
=\ &\left(tr\otimes\int_E\right)\left(Proj(W_{i_1a_1}UW_{j_1b_1}^*)\ldots Proj(W_{i_pa_p}UW_{j_pb_p}^*)\right)\\
=\ &\frac{1}{N}\int_E<W_{i_1a_1}UW_{j_1b_1}^*,W_{i_pa_p}UW_{j_pb_p}^*>\ldots<W_{i_2a_2}UW_{j_2b_2}^*,W_{i_1a_1}UW_{j_1b_1}^*>dU
\end{aligned}$$

In order to compute now the scalar products, observe that we have:

$$\begin{aligned}
<W_{ia}UW_{jb}^*,W_{kc}UW_{ld}^*> &= tr(W_{jb}U^*W_{ia}^*W_{kc}UW_{ld}^*)\\
&= tr(W_{ia}^*W_{kc}UW_{ld}^*W_{jb}U^*)\\
&= <i,a-c><l,d-b>tr(W_{k-i,c-a}UW_{j-l,b-d}U^*)
\end{aligned}$$

By plugging these quantities into the formula of T_p, we obtain the result. □

Consider now the Weyl group $W=\{W_{ia}\}\subset U_n$, that we already met in the proof of Proposition 12.14. We have the following result, from [16]:

THEOREM 12.17. *For any compact group $W\subset E\subset U_n$, the model*

$$\pi_H:C(S_N^+)\to M_N(C(E))$$

$$w_{ia,jb}\to[U\to Proj(W_{ia}UW_{jb}^*)]$$

constructed above is stationary on its image.

PROOF. We must prove that we have $T_p^2=T_p$. We have:

$$\begin{aligned}
&(T_p^2)_{ia,jb}\\
=\ &\sum_{kc}(T_p)_{ia,kc}(T_p)_{kc,jb}\\
=\ &\frac{1}{N^2}\sum_{kc}<i_1,a_1-a_p>\ldots<i_p,a_p-a_{p-1}><k_1,c_1-c_2>\ldots<k_p,c_p-c_1>\\
&<k_1,c_1-c_p>\ldots<k_p,c_p-c_{p-1}><j_1,b_1-b_2>\ldots<j_p,b_p-b_1>\\
&\int_E tr(W_{i_1-i_2,a_1-a_2}UW_{k_2-k_1,c_2-c_1}U^*)\ldots tr(W_{i_p-i_1,a_p-a_1}UW_{k_1-k_p,c_1-c_p}U^*)dU\\
&\int_E tr(W_{k_1-k_2,c_1-c_2}VW_{j_2-j_1,b_2-b_1}V^*)\ldots tr(W_{k_p-k_1,c_p-c_1}VW_{j_1-j_p,b_1-b_p}V^*)dV
\end{aligned}$$

By rearranging the terms, this formula becomes:

$$\begin{aligned}
&(T_p^2)_{ia,jb}\\
=\ &\frac{1}{N^2} < i_1, a_1 - a_p > \ldots < i_p, a_p - a_{p-1} >< j_1, b_1 - b_2 > \ldots < j_p, b_p - b_1 >\\
&\int_E \int_E \sum_{kc} < k_1 - k_p, c_1 - c_p > \ldots < k_p - k_{p-1}, c_p - c_{p-1} >\\
&tr(W_{i_1-i_2,a_1-a_2} U W_{k_2-k_1,c_2-c_1} U^*) tr(W_{k_1-k_2,c_1-c_2} V W_{j_2-j_1,b_2-b_1} V^*)\\
&\ldots\ldots\\
&tr(W_{i_p-i_1,a_p-a_1} U W_{k_1-k_p,c_1-c_p} U^*) tr(W_{k_p-k_1,c_p-c_1} V W_{j_1-j_p,b_1-b_p} V^*) dU dV
\end{aligned}$$

Let us denote by I the above double integral. By using $W_{kc}^* =< k, c > W_{-k,-c}$ for each of the couplings, and by moving as well all the U^* variables to the left, we obtain:

$$\begin{aligned}
I\ =\ &\int_E \int_E \sum_{kc} tr(U^* W_{i_1-i_2,a_1-a_2} U W_{k_2-k_1,c_2-c_1}) tr(W_{k_2-k_1,c_2-c_1}^* V W_{j_2-j_1,b_2-b_1} V^*)\\
&\ldots\ldots\\
&tr(U^* W_{i_p-i_1,a_p-a_1} U W_{k_1-k_p,c_1-c_p}) tr(W_{k_1-k_p,c_1-c_p}^* V W_{j_1-j_p,b_1-b_p} V^*) dU dV
\end{aligned}$$

In order to perform now the sums, we use the following formula:

$$\begin{aligned}
tr(AW_{kc}) tr(W_{kc}^* B) &= \frac{1}{N} \sum_{qrst} A_{qr} (W_{kc})_{rq} (W_{kc}^*)_{st} B_{ts}\\
&= \frac{1}{N} \sum_{qrst} A_{qr} < k, q > \delta_{r-q,c} < k, -s > \delta_{t-s,c} B_{ts}\\
&= \frac{1}{N} \sum_{qs} < k, q - s > A_{q,q+c} B_{s+c,s}
\end{aligned}$$

If we denote by A_x, B_x the variables which appear in the formula of I, we have:

$$\begin{aligned}
&I\\
=\ &\frac{1}{N^p} \int_E \int_E \sum_{kcqs} < k_2 - k_1, q_1 - s_1 > \ldots < k_1 - k_p, q_p - s_p >\\
&(A_1)_{q_1,q_1+c_2-c_1} (B_1)_{s_1+c_2-c_1,s_1} \ldots (A_p)_{q_p,q_p+c_1-c_p} (B_p)_{s_p+c_1-c_p,s_p}\\
=\ &\frac{1}{N^p} \int_E \int_E \sum_{kcqs} < k_1, q_p - s_p - q_1 + s_1 > \ldots < k_p, q_{p-1} - s_{p-1} - q_p + s_p >\\
&(A_1)_{q_1,q_1+c_2-c_1} (B_1)_{s_1+c_2-c_1,s_1} \ldots (A_p)_{q_p,q_p+c_1-c_p} (B_p)_{s_p+c_1-c_p,s_p}
\end{aligned}$$

Now observe that we can perform the sums over $k_1, \ldots, k_p$. We obtain in this way a multiplicative factor n^p, along with the condition:

$$q_1 - s_1 = \ldots = q_p - s_p$$

Thus we must have $q_x = s_x + a$ for a certain a, and the above formula becomes:

$$I = \frac{1}{n^p}\int_E\int_E \sum_{csa}(A_1)_{s_1+a,s_1+c_2-c_1+a}(B_1)_{s_1+c_2-c_1,s_1}\ldots(A_p)_{s_p+a,s_p+c_1-c_p+a}(B_p)_{s_p+c_1-c_p,s_p}$$

Consider now the variables $r_x = c_{x+1} - c_x$, which altogether range over the set Z of multi-indices having sum 0. By replacing the sum over c_x with the sum over r_x, which creates a multiplicative n factor, we obtain the following formula:

$$I = \frac{1}{n^{p-1}}\int_E\int_E \sum_{r\in Z}\sum_{sa}(A_1)_{s_1+a,s_1+r_1+a}(B_1)_{s_1+r_1,s_1}\ldots(A_p)_{s_p+a,s_p+r_p+a}(B_p)_{s_p+r_p,s_p}$$

For an arbitrary multi-index r, we have the following formula:

$$\delta_{\sum_i r_i,0} = \frac{1}{n}\sum_i < i, r_1 > \ldots < i, r_p >$$

Thus, we can replace the sum over $r \in Z$ by a full sum, as follows:

$$\begin{aligned} I &= \frac{1}{n^p}\int_E\int_E \sum_{rsia} < i, r_1 > (A_1)_{s_1+a,s_1+r_1+a}(B_1)_{s_1+r_1,s_1} \\ &\qquad\qquad \ldots\ldots \\ &\qquad\qquad < i, r_p > (A_p)_{s_p+a,s_p+r_p+a}(B_p)_{s_p+r_p,s_p} \end{aligned}$$

In order to "absorb" now the indices i, a, we can use the following formula:

$$\begin{aligned} & W_{ia}^*AW_{ia} \\ = & \left(\sum_b < i, -b > E_{b,a+b}\right)\left(\sum_{bc} E_{a+b,a+c}A_{a+b,a+c}\right)\left(\sum_c < i, c > E_{a+c,c}\right) \\ = & \sum_{bc} < i, c-b > E_{bc}A_{a+b,a+c} \end{aligned}$$

Thus we have the following formula:

$$(W_{ia}^*AW_{ia})_{bc} =< i, c-b > A_{a+b,a+c}$$

With this in hand, our formula becomes:

$$\begin{aligned} & I \\ = & \frac{1}{n^p}\int_E\int_E \sum_{rsia}(W_{ia}^*A_1W_{ia})_{s_1,s_1+r_1}(B_1)_{s_1+r_1,s_1}\ldots(W_{ia}^*A_pW_{ia})_{s_p,s_p+r_p}(B_p)_{s_p+r_p,s_p} \\ = & \int_E\int_E \sum_{ia} tr(W_{ia}^*A_1W_{ia}B_1)\ldots\ldots tr(W_{ia}^*A_pW_{ia}B_p) \end{aligned}$$

Now by replacing A_x, B_x with their respective values, we obtain:

$$\begin{aligned} I &= \int_E\int_E \sum_{ia} tr(W_{ia}^*U^*W_{i_1-i_2,a_1-a_2}UW_{ia}VW_{j_2-j_1,b_2-b_1}V^*) \\ &\quad\cdots\cdots \\ &\quad tr(W_{ia}^*U^*W_{i_p-i_1,a_p-a_1}UW_{ia}VW_{j_1-j_p,b_1-b_p}V^*)dUdV \end{aligned}$$

By moving the $W_{ia}^*U^*$ variables at right, we obtain, with $S_{ia} = UW_{ia}V$:

$$\begin{aligned} I &= \sum_{ia}\int_E\int_E tr(W_{i_1-i_2,a_1-a_2}S_{ia}W_{j_2-j_1,b_2-b_1}S_{ia}^*) \\ &\quad\cdots\cdots \\ &\quad tr(W_{i_p-i_1,a_p-a_1}S_{ia}W_{j_1-j_p,b_1-b_p}S_{ia}^*)dUdV \end{aligned}$$

Now since S_{ia} is Haar distributed when U, V are Haar distributed, we obtain:

$$I = N\int_E\int_E tr(W_{i_1-i_2,a_1-a_2}UW_{j_2-j_1,b_2-b_1}U^*)\ldots tr(W_{i_p-i_1,a_p-a_1}UW_{j_1-j_p,b_1-b_p}U^*)dU$$

But this is exactly N times the integral in the formula of $(T_p)_{ia,jb}$, from Proposition 12.16. Since the N factor cancels with one of the two N factors that we found in the beginning of the proof, when first computing $(T_p^2)_{ia,jb}$, we are done. □

As an illustration for the above result, going back to [14], we have:

THEOREM 12.18. *We have a stationary matrix model*

$$\pi : C(S_4^+) \subset M_4(C(SU_2))$$

given on the standard coordinates by the formula

$$\pi(u_{ij}) = [x \to Proj(c_i x c_j)]$$

where $x \in SU_2$, and c_1, c_2, c_3, c_4 are the Pauli matrices.

PROOF. As already explained in the comments following Definition 12.12, the Pauli matrices appear as particular cases of the Weyl matrices:

$$W_{00} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad W_{10} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$W_{11} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad , \quad W_{01} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Thus, Theorem 12.17 produces in this case the model in the statement. □

Observe that, since the projection $Proj(c_i x c_j)$ depends only on the image of x in the quotient $SU_2 \to SO_3$, we can replace the model space SU_2 by the smaller space SO_3.

This can be used in conjunction with the isomorphism $S_4^+ \simeq SO_3'$, and as explained in [13], our model becomes in this way something more conceptual, as follows:

$$\pi : C(SO_3') \subset M_4(C(SO_3))$$

In order to discuss the continuation of this, we need to know more about the quantum group $S_4^+ \simeq SO_3'$, and its subgroups. Let us start with the following definition:

DEFINITION 12.19. *We let $SO_3' \subset O_3'$ be the subgroup coming from the relation*

$$\sum_{\sigma\in S_3} u_{1\sigma(1)}u_{2\sigma(2)}u_{3\sigma(3)} = 1$$

called twisted determinant one condition.

Normally, we should prove here that $C(SO_3')$ is indeed a Woronowicz algebra. This is of course possible, directly, but follows as well from the following result:

THEOREM 12.20. *We have an isomorphism of compact quantum groups*

$$S_4^+ = SO_3'$$

given by the Fourier transform over the Klein group $K = \mathbb{Z}_2 \times \mathbb{Z}_2$.

PROOF. Consider the following matrix, coming from the standard action of SO_3':

$$u^+ = \begin{pmatrix} 1 & 0 \\ 0 & u \end{pmatrix}$$

We apply to this matrix the Fourier transform over the Klein group $K = \mathbb{Z}_2 \times \mathbb{Z}_2$:

$$v = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & u_{11} & u_{12} & u_{13} \\ 0 & u_{21} & u_{22} & u_{23} \\ 0 & u_{31} & u_{32} & u_{33} \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$$

This matrix is then magic, and vice versa, so the Fourier transform over K converts the relations in Definition 12.19 into the magic relations. But this gives the result. □

We know from the above that we have a stationary matrix model for the algebra $C(S_4^+) = C(SO_3')$. But this suggests the following conjecture:

CONJECTURE 12.21. *Given a quantum permutation group of 4 points,*

$$G \subset S_4^+ \simeq SO_3'$$

coming by twisting a usual ADE subgroup of the group SO_3,

$$H \subset SO_3$$

the restriction of the Pauli model for $C(S_4^+)$, with fibers coming from the elements of $H \subset SO_3$, has the algebra $C(G)$ as Hopf image.

To be more precise, we know that the conjecture holds for $G = S_4^+$ itself. Indeed, here we have $H = SO_3$, so the corresponding restriction of the Pauli model for $C(S_4^+)$ is the Pauli model itself, and the Hopf image is the algebra $C(S_4^+)$ itself, as stated.

12c. Fourier models

Fourier models.

12d. Hadamard models

Hadamard models.

12e. Exercises

Exercises:

EXERCISE 12.22.

EXERCISE 12.23.

EXERCISE 12.24.

EXERCISE 12.25.

EXERCISE 12.26.

EXERCISE 12.27.

Bonus exercise.

Part IV

Quantum fields

So lead me to the harbor
And float me on the waves
Sink me in the ocean
To sleep in a sailor's grave

CHAPTER 13

Quantum physics

13a. Dirac equation

As a starting point for more advanced quantum physics, we have:

FACT 13.1. *The master equation for free electromagnetic radiation, that is, for free photons, is the wave equation at speed $v = c$, namely:*

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

This equation can be reformulated in the more symmetric form

$$\left(\frac{1}{c^2} \cdot \frac{d^2}{dt^2} - \Delta \right) \varphi = 0$$

with the operator on the left being called the d'Alembertian.

In relation now with the electron, there is an obvious similarity here with the free Schrödinger equation, without potential V, which reads:

$$\left(i \frac{d}{dt} + \frac{h}{2m} \Delta \right) \psi = 0$$

This similarity suggests looking for a relativistic version of the Schrödinger equation, which is compatible with the wave equation at $v = c$. And coming up with such an equation is not very complicated, the straightforward answer being as follows:

DEFINITION 13.2. *The following abstract mathematical equation,*

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

on a function $\psi = \psi_t(x)$, is called the Klein-Gordon equation.

To be more precise, what we have here is some sort of a speculatory equation, formally obtained from the Schrödinger equation, via a few simple manipulations, as to make it relativistic. And with the relation with photons being something very simple, the thing being that at zero mass, $m = 0$, we obtain precisely the wave equation at $v = c$.

All this is very nice, looks like we have a beginning of theory here, both making the electrons relativistic, and unifying them with photons. And isn't this too beautiful to be true. Going ahead now with physics, the following question appears:

QUESTION 13.3. *What does the Klein-Gordon equation really describe?*

And here, unfortunately, bad news all the way. A closer look at the Klein-Gordon equation reveals all sorts of bugs, making it unusable for anything reasonable. And with the main bug, which is enough for disqualifying it, being that, unlike the Schrödinger equation which preserves probability amplitudes $|\psi|^2$, the Klein-Gordon equation does not have this property. Thus, even before trying to understand what the Klein-Gordon equation really describes, we are left with the conclusion that this equation cannot really describe anything reasonable, due to the formal nature of the function ψ involved.

This being said, the Klein-Gordon equation found later a number of interesting applications, and we will discuss all this, in what follows. To start with, Dirac came upon the idea of extracting the square root of the Klein-Gordon operator, as follows:

PROPOSITION 13.4. *We can extract the square root of the Klein-Gordon operator, via a formula as follows,*

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

by using matrices P,Q,R,S which anticommute, $AB=-BA$, and whose squares equal one, $A^2=1$.

PROOF. We have the following computation, valid for any matrices P,Q,R,S, with the notation $\{A,B\}=AB+BA$:

$$\begin{aligned}\left(\frac{i}{c}\cdot\frac{Pd}{dt}+\frac{Qd}{dx}+\frac{Rd}{dy}+\frac{Sd}{dz}\right)^2 &= -\frac{1}{c^2}\cdot\frac{P^2d^2}{dt^2}+\frac{Q^2d^2}{dx^2}+\frac{R^2d^2}{dy^2}+\frac{S^2d^2}{dz^2}\\ &+\frac{i}{c}\left(\frac{\{P,Q\}d^2}{dtdx}+\frac{\{P,R\}d^2}{dtdy}+\frac{\{P,S\}d^2}{dtdz}\right)\\ &+\frac{\{Q,R\}d^2}{dxdy}+\frac{\{Q,S\}d^2}{dxdz}+\frac{\{R,S\}d^2}{dydz}\end{aligned}$$

Thus, in order to obtain in this way the Klein-Gordon operator, the conditions in the statement must be satisfied. □

As a technical comment here, normally when extracting a square root, we should look for a hermitian operator. In view of this, observe that we have:

$$\left(\frac{i}{c}\cdot\frac{Pd}{dt}+\frac{Qd}{dx}+\frac{Rd}{dy}+\frac{Sd}{dz}\right)^*=-\frac{i}{c}\cdot\frac{P^*d}{dt}+\frac{Q^*d}{dx}+\frac{R^*d}{dy}+\frac{S^*d}{dz}$$

Thus, we should normally add the conditions $P^*=-P$ and $Q^*=Q$, $R^*=R$, $S^*=S$ to those above. But, the thing is that due to some subtle reasons, the natural square root of the Klein-Gordon operator is not hermitian. More on this later.

Looking for matrices P, Q, R, S as above is not exactly trivial, and the simplest solutions appear in $M_4(\mathbb{C})$, in connection with the Pauli matrices, as follows:

PROPOSITION 13.5. *The simplest matrices P, Q, R, S as above appear as*

$$P = \gamma_0 \quad , \quad Q = i\gamma_1 \quad , \quad R = i\gamma_2 \quad , \quad S = i\gamma_3$$

with $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ being the Dirac matrices, given by

$$\gamma_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad , \quad \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}$$

where $\sigma_1, \sigma_2, \sigma_3$ are the Pauli spin matrices, given by:

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad , \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

PROOF. We have $\gamma_0^2 = 1$, and by using $\sigma_i^2 = 1$ for any $i = 1, 2, 3$, we have as well the following formula, which shows that we have $(i\gamma_i)^2 = 1$, as needed:

$$\gamma_i^2 = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

As in what regards the commutators, we first have, for any $i = 1, 2, 3$:

$$\gamma_0\gamma_i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}$$

$$\gamma_i\gamma_0 = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -\sigma_i \\ -\sigma_i & 0 \end{pmatrix}$$

Regarding now the remaining commutators, observe here that we have:

$$\gamma_i\gamma_j = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_i\sigma_j & 0 \\ 0 & -\sigma_i\sigma_j \end{pmatrix}$$

Now since the Pauli matrices anticommute, we obtain $\gamma_i\gamma_j = -\gamma_j\gamma_i$, as desired. □

We can now put everything together, and we obtain:

THEOREM 13.6. *The following operator, called Dirac operator,*

$$D = i\left(\frac{\gamma_0 d}{cdt} + \frac{\gamma_1 d}{dx} + \frac{\gamma_2 d}{dy} + \frac{\gamma_3 d}{dz}\right)$$

has the property that its square is the Klein-Gordon operator.

PROOF. With notations from Proposition 13.4 and Proposition 13.5, and by making the choices in Proposition 13.5, we have:

$$\begin{aligned}\frac{i}{c}\cdot\frac{Pd}{dt}+\frac{Qd}{dx}+\frac{Rd}{dy}+\frac{Sd}{dz} &= \frac{i}{c}\cdot\frac{\gamma_0 d}{dt}+\frac{i\gamma_1 d}{dx}+\frac{i\gamma_2 d}{dy}+\frac{i\gamma_3 d}{dz}\\ &= i\left(\frac{\gamma_0 d}{cdt}+\frac{\gamma_1 d}{dx}+\frac{\gamma_2 d}{dy}+\frac{\gamma_3 d}{dz}\right)\end{aligned}$$

Thus, we have here a square root of the Klein-Gordon operator, as desired. □

We can now extract the square root of the Klein-Gordon equation, as follows:

DEFINITION 13.7. *We have the following equation, called Dirac equation,*

$$ih\left(\frac{\gamma_0 d}{cdt}+\frac{\gamma_1 d}{dx}+\frac{\gamma_2 d}{dy}+\frac{\gamma_3 d}{dz}\right)\psi=mc\psi$$

obtained by extracting the square root of the Klein-Gordon equation.

As a first observation, all this is very related to spin. In fact, as we will see later, the Dirac equation is the correct relativistic equation describing the spin 1/2 particles.

The Dirac equation comes with a price to pay, which is that of opening Pandora's box of particles. To be more precise, once we adopt this equation, we must surely adopt all its free solutions. And bad news here, the solution which is complementary to the electron is not the proton, but rather a weird new particle, called the positron.

In order to explain all this, let us start with the following observation:

PROPOSITION 13.8. *For a particle at rest, meaning under the assumption*

$$\frac{d\psi}{dx}=\frac{d\psi}{dy}=\frac{d\psi}{dz}=0$$

the Dirac equation takes the form

$$\frac{ih}{c}\cdot\gamma_0\cdot\frac{d\psi}{dt}=mc\psi$$

with $\gamma_0=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$ *being as usual the first Dirac matrix.*

PROOF. Consider indeed the Dirac equation, as formulated in Definition 13.7:

$$ih\left(\frac{\gamma_0 d}{cdt}+\frac{\gamma_1 d}{dx}+\frac{\gamma_2 d}{dy}+\frac{\gamma_3 d}{dz}\right)\psi=mc\psi$$

With the above rest assumption, we are led to the equation in the statement. □

The above equation at rest is very easy to solve, the result being as follows:

THEOREM 13.9. *The solutions of the Dirac equation for particles at rest are*

$$\psi = \begin{pmatrix} e^{-imc^2t/h}\xi \\ e^{imc^2t/h}\eta \end{pmatrix}$$

with $\xi, \eta \in \mathbb{R}^2$ being arbitrary vectors.

PROOF. In order to solve the Dirac equation in Proposition 13.8, let us write:

$$\psi = \begin{pmatrix} \varphi \\ \phi \end{pmatrix}$$

With this notation, the Dirac equation at rest takes the following form:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} d\varphi/dt \\ d\phi/dt \end{pmatrix} = -\frac{imc^2}{h} \begin{pmatrix} \varphi \\ \phi \end{pmatrix}$$

Now by looking at the components, the equations are as follows:

$$\frac{d\varphi}{dt} = -\frac{imc^2}{h}\,\varphi \quad , \quad \frac{d\phi}{dt} = \frac{imc^2}{h}\,\phi$$

But the solutions of these latter equations are as follows, with $\xi, \eta \in \mathbb{R}^2$:

$$\varphi = e^{-imc^2t/h}\xi \quad , \quad \phi = e^{imc^2t/h}\eta$$

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

The question is now, is the above result good news or not? Not really, because in view of what we know from quantum mechanics, an $e^{-iEt/h}$ factor should correspond to the time dependence of a quantum state with energy E, which at rest is $E = mc^2$. And from this perspective, while the above φ functions look very good, the other components, the ϕ functions, look bad, seemingly coming from particles having "negative energy".

So, what to do? The solution is that of talking about antiparticles with positive energy, and to formulate, as a continuation of Theorem 13.9:

THEOREM 13.10. *The basic solutions of the Dirac equation for particles at rest are*

$$\psi^1 = e^{-imc^2t/h} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \psi^2 = e^{-imc^2t/h} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

corresponding to the electron with spin up, and spin down, plus

$$\psi^3 = e^{imc^2t/h} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad \psi^4 = e^{imc^2t/h} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

corresponding to a new particle, the positron, with spin up, and spin down.

PROOF. Here the mathematics comes from what we found in Theorem 13.9, and the terminology and philosophy comes from the above discussion. With the remark that the newly introduced positron is rather an antiparticle, but more on this later. □

Moving now forward, let us attempt to solve the following question:

QUESTION 13.11. *What are the plane wave solutions*

$$\psi(s) = ae^{-i<k,s>}u$$

of the Dirac equation?

To be more precise, we are using here, as argument of the function ψ, the standard relativistic space-time position $s \in \mathbb{R}^4$ of our particle, namely:

$$s = \begin{pmatrix} ct \\ r \end{pmatrix} \quad , \quad r = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Next, we have in the above a constant $a \in \mathbb{R}$, which will be quite irrelevant to our computations, the Dirac equation being linear. Regarding now k, it is convenient to write this vector split over components, as we did in the above with s, as follows:

$$k = \begin{pmatrix} f \\ g \end{pmatrix} \quad , \quad g = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix}$$

With these conventions, along with the standard relativistic convention that the space coordinates contribute with $-$ signs, the scalar product in Question 13.11 is given by:

$$< k, s >= cft - < g, r >$$

Now observe that the real part of the exponential in Question 13.11 is given by:

$$Re(e^{-i<k,s>}) = \cos\big(cft - < g, r >\big)$$

Thus, what we have here, justifying the terminology, is a sinusoidal wave propagating in the direction g, with angular frequency and wavelenght as follows:

$$\omega = cf \quad , \quad \lambda = 2\pi/||g||$$

In order to answer Question 13.11, we must first plug into the Dirac equation our special function ψ. We are led in this way to a quite simple equation, as follows:

PROPOSITION 13.12. *The Dirac equation for plane wave functions*

$$\psi(s) = ae^{-i<k,s>}u$$

takes the following special form, no longer involving derivatives,

$$h\big(\gamma_0 f - \gamma_1 g_1 - \gamma_2 g_2 - \gamma_3 g_3\big)u = mcu$$

with the above conventions for indices and vectors.

PROOF. Consider indeed the Dirac equation, as formulated in Definition 13.7:

$$ih\left(\frac{\gamma_0 d}{cdt}+\frac{\gamma_1 d}{dx}+\frac{\gamma_2 d}{dy}+\frac{\gamma_3 d}{dz}\right)\psi=mc\psi$$

For the function ψ in the statement, the derivatives are given by:

$$\frac{d\psi}{ds_i}=-ik_i\psi$$

Thus, with our above conventions for indices and vectors, we have:

$$\frac{d\psi}{cdt}=-if\psi\quad,\quad\frac{d\psi}{dr_i}=ig_i\psi$$

By plugging these quantities in the Dirac equation, this equation becomes:

$$h\big(\gamma_0 f-\gamma_1 g_1-\gamma_2 g_2-\gamma_3 g_3\big)\psi=mc\psi$$

Now by using again $\psi=ae^{-i<k,s>}u$, this equation takes the following form:

$$h\big(\gamma_0 f-\gamma_1 g_1-\gamma_2 g_2-\gamma_3 g_3\big)ae^{-i<k,s>}u=mcae^{-i<k,s>}u$$

Thus, by simplifying, we are led to the equation in the statement. □

Let us study now the equation that we found. As a first observation, we can further fine-tune the equation in Proposition 13.12, via some simple manipulations, as follows:

PROPOSITION 13.13. *In the context of Proposition 13.12, with the notation*

$$u=\begin{pmatrix}v\\w\end{pmatrix}$$

the Dirac equation takes the following form, in terms of the components v,w,

$$v=\frac{<g,\sigma>}{f-mc/h}\,w\qquad,\qquad w=\frac{<g,\sigma>}{f+mc/h}\,v$$

where $\sigma_1,\sigma_2,\sigma_3$ stand as usual for the Pauli spin matrices.

PROOF. According to the definition of the Dirac matrices, in terms of the Pauli ones, we have the following computation, for the operator appearing in Proposition 13.12:

$$\begin{aligned}\gamma_0 f-\gamma_1 g_1-\gamma_2 g_2-\gamma_3 g_3&=\begin{pmatrix}1&0\\0&-1\end{pmatrix}f-\sum_{i=1}^{3}\begin{pmatrix}0&\sigma_i\\-\sigma_i&0\end{pmatrix}g_i\\&=\begin{pmatrix}f&0\\0&-f\end{pmatrix}-\begin{pmatrix}0&<g,\sigma>\\-<g,\sigma>&0\end{pmatrix}\\&=\begin{pmatrix}f&-<g,\sigma>\\<g,\sigma>&-f\end{pmatrix}\end{aligned}$$

Thus, the quantity which must vanish in Proposition 13.12 is given by:

$$\begin{aligned}
&\Big(h\big(\gamma_0 f-\gamma_1 g_1-\gamma_2 g_2-\gamma_3 g_3\big)-mc\Big)u\\
=\ &\begin{pmatrix} hf-mc & -h<g,\sigma>\\ h<g,\sigma> & -hf-mc\end{pmatrix}\begin{pmatrix}v\\w\end{pmatrix}\\
=\ &\begin{pmatrix}(hf-mc)v-h<g,\sigma>w\\ h<g,\sigma>v-(hf+mc)w\end{pmatrix}
\end{aligned}$$

We therefore conclude that, in our case, the Dirac equation reads:

$$(hf-mc)v=h<g,\sigma>w$$

$$h<g,\sigma>v=(hf+mc)w$$

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

In order to solve now our equation, let us make the following observation:

PROPOSITION 13.14. *In the context of Proposition 13.13 we must have*

$$||g||^2=f^2-\left(\frac{mc}{h}\right)^2$$

under the assumption that the solution is nonzero, $u\neq 0$.

PROOF. Consider the equations found in Proposition 13.13, namely:

$$v=\frac{<g,\sigma>}{f-mc/h}\,w\qquad,\qquad w=\frac{<g,\sigma>}{f+mc/h}\,v$$

By substituting, we are led to the following formulae:

$$v=\frac{<g,\sigma>^2}{f^2-(mc/h)^2}\,v\quad,\quad w=\frac{<g,\sigma>^2}{f^2-(mc/h)^2}\,w$$

Thus, assuming that the solution is nonzero, $u\neq 0$, we must have:

$$\frac{<g,\sigma>^2}{f^2-(mc/h)^2}=\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

Now, let us compute the left term. According to our various conventions above, and to the formulae for the Pauli matrices, we have the following formula:

$$\begin{aligned}
<g,\sigma>\ &=\ g_1\begin{pmatrix}0&1\\1&0\end{pmatrix}+g_2\begin{pmatrix}0&-i\\i&0\end{pmatrix}+g_3\begin{pmatrix}1&0\\0&-1\end{pmatrix}\\
&=\ \begin{pmatrix}g_3&g_1-ig_2\\ g_1+ig_2&-g_3\end{pmatrix}
\end{aligned}$$

By raising this quantity to the square, we obtain:

$$\begin{aligned}
< g, \sigma >^2 &= \begin{pmatrix} g_3 & g_1 - ig_2 \\ g_1 + ig_2 & -g_3 \end{pmatrix} \begin{pmatrix} g_3 & g_1 - ig_2 \\ g_1 + ig_2 & -g_3 \end{pmatrix} \\
&= \begin{pmatrix} g_3^2 + (g_1 - ig_2)(g_1 + ig_2) & g_3(g_1 - ig_2) - (g_1 - ig_2)g_3 \\ (g_1 + ig_2)g_3 - g_3(g_1 + ig_2) & (g_1 + ig_2)(g_1 - ig_2) + g_3^2 \end{pmatrix} \\
&= \begin{pmatrix} g_1^2 + g_2^2 + g_3^2 & 0 \\ 0 & g_1^2 + g_2^2 + g_3^2 \end{pmatrix} \\
&= ||g||^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

Thus, the condition that we found above, coming from $u \neq 0$, reads:

$$\frac{||g||^2}{f^2 - (mc/h)^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We conclude that we must have the following equality:

$$||g||^2 = f^2 - \left(\frac{mc}{h}\right)^2$$

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

The point now is that the above result invites us to use the rescaled energy-momentum four-vector as variable, $k = \pm p/h$, and we are led in this way to the following result:

THEOREM 13.15. *The basic plane wave solutions, of type*

$$\psi(s) = ae^{-i<k,s>}u$$

of the Dirac equation, come from the functions

$$u^1 = \frac{1}{E + mc^2} \begin{pmatrix} E + mc^2 \\ 0 \\ cp_z \\ cp_x + icp_y \end{pmatrix} \quad , \quad u^2 = \frac{1}{E + mc^2} \begin{pmatrix} 0 \\ E + mc^2 \\ cp_x - icp_y \\ -cp_z \end{pmatrix}$$

corresponding to particle solutions, plus from the functions

$$u^3 = \frac{1}{E + mc^2} \begin{pmatrix} cp_z \\ cp_x + icp_y \\ E + mc^2 \\ 0 \end{pmatrix} \quad , \quad u^4 = \frac{1}{E + mc^2} \begin{pmatrix} cp_x - icp_y \\ -cp_z \\ 0 \\ E + mc^2 \end{pmatrix}$$

corresponding to antiparticle solutions.

PROOF. This comes by putting together all the above. Indeed, with $k = \pm p/h$, as suggested above, we have four choices, which are as follows:

$$v = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad w = \frac{c}{E + mc^2} \begin{pmatrix} p_z \\ p_x + ip_y \end{pmatrix}$$

$$v = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \quad w = \frac{c}{E + mc^2} \begin{pmatrix} p_x - ip_y \\ -p_z \end{pmatrix}$$

$$w = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad v = \frac{c}{E + mc^2} \begin{pmatrix} p_z \\ p_x + ip_y \end{pmatrix}$$

$$w = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \quad v = \frac{c}{E + mc^2} \begin{pmatrix} p_x - ip_y \\ -p_z \end{pmatrix}$$

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

13b.

13c.

13d.

13e. Exercises

Exercises:

EXERCISE 13.16.

EXERCISE 13.17.

EXERCISE 13.18.

EXERCISE 13.19.

EXERCISE 13.20.

EXERCISE 13.21.

Bonus exercise.

CHAPTER 14

Relativistic aspects

14a.

14b.

14c.

14d.

14e. Exercises

Exercises:

Exercise 14.1.

Exercise 14.2.

Exercise 14.3.

Exercise 14.4.

Exercise 14.5.

Exercise 14.6.

Bonus exercise.

CHAPTER 15

Lattice theory

15a.

15b.

15c.

15d.

15e. Exercises

Exercises:

Exercise 15.1.

Exercise 15.2.

Exercise 15.3.

Exercise 15.4.

Exercise 15.5.

Exercise 15.6.

Bonus exercise.

CHAPTER 16

Into the void

16a.

16b.

16c.

16d.

16e. Exercises

Congratulations for having read this book, and no exercises for this final chapter.

Bibliography

[1] A.A. Abrikosov, Fundamentals of the theory of metals, Dover (1988).

[2] A.A. Abrikosov, L.P. Gorkov and I.E. Dzyaloshinski, Methods of quantum field theory in statistical physics, Dover (1963).

[3] G.W. Anderson, A. Guionnet and O. Zeitouni, An introduction to random matrices, Cambridge Univ. Press (2010).

[4] V.I. Arnold, Mathematical methods of classical mechanics, Springer (1974).

[5] V.I. Arnold, Lectures on partial differential equations, Springer (1997).

[6] V.I. Arnold, Catastrophe theory, Springer (1974).

[7] V.I. Arnold and B.A. Khesin, Topological methods in hydrodynamics, Springer (1998).

[8] N.W. Ashcroft and N.D. Mermin, Solid state physics, Saunders College Publ. (1976).

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

[10] T. Banica, Principles of operator algebras (2024).

[11] T. Banica, Introduction to modern physics (2024).

[12] T. Banica, Principles of thermodynamics (2024).

[13] T. Banica and J. Bichon, Random walk questions for linear quantum groups, *Int. Math. Res. Not.* **24** (2015), 13406–13436.

[14] T. Banica and B. Collins, Integration over the Pauli quantum group, *J. Geom. Phys.* **58** (2008), 942–961.

[15] T. Banica and S. Curran, Decomposition results for Gram matrix determinants, *J. Math. Phys.* **51** (2010), 1–14.

[16] T. Banica and I. Nechita, Flat matrix models for quantum permutation groups, *Adv. Appl. Math.* **83** (2017), 24–46.

[17] G.K. Batchelor, An introduction to fluid dynamics, Cambridge Univ. Press (1967).

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

[19] S.J. Blundell and K.M. Blundell, Concepts in thermal physics, Oxford Univ. Press (2006).

[20] P.M. Chaikin and T.C. Lubensky, Principles of condensed matter physics, Cambridge Univ. Press (1995).

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

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

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

[24] P.A. Davidson, Introduction to magnetohydrodynamics, Cambridge Univ. Press (2001).
[25] P. Di Francesco, Meander determinants, *Comm. Math. Phys.* **191** (1998), 543–583.
[26] P. Di Francesco, P. Mathieu and D. Sénéchal, Conformal field theory, Springer (1996).
[27] V.G. Drinfeld, Quantum groups, Proc. ICM Berkeley (1986), 798–820.
[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] L.D. Faddeev, Instructive history of the quantum inverse scattering method, *Acta Appl. Math.* **39** (1995), 69–84.
[32] L.D. Faddeev, N.Y. Reshetikhin and L. A. Takhtajan, Quantization of Lie groups and Lie algebras, in "Algebraic analysis", Academic Press (1988), 129–139.
[33] L.D. Faddeev and L. A. Takhtajan, Hamiltonian methods in the theory of solitons, Springer (2007).
[34] W. Feller, An introduction to probability theory and its applications, Wiley (1950).
[35] E. Fermi, Thermodynamics, Dover (1937).
[36] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics I: mainly mechanics, radiation and heat, Caltech (1963).
[37] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics II: mainly electromagnetism and matter, Caltech (1964).
[38] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics III: quantum mechanics, Caltech (1966).
[39] R.P. Feynman and A.R. Hibbs, Quantum mechanics and path integrals, Dover (1965).
[40] N. Goldenfeld, Lectures on phase transitions and the renormalization group, CRC Press (1992).
[41] H. Goldstein, C. Safko and J. Poole, Classical mechanics, Addison-Wesley (1980).
[42] D.L. Goodstein, States of matter, Dover (1975).
[43] M.B. Green, J.H. Schwarz and E. Witten, Superstring theory, Cambridge Univ. Press (2012).
[44] D.J. Griffiths, Introduction to electrodynamics, Cambridge Univ. Press (2017).
[45] D.J. Griffiths and D.F. Schroeter, Introduction to quantum mechanics, Cambridge Univ. Press (2018).
[46] D.J. Griffiths, Introduction to elementary particles, Wiley (2020).
[47] D.J. Griffiths, Revolutions in twentieth-century physics, Cambridge Univ. Press (2012).
[48] W.A. Harrison, Solid state theory, Dover (1970).
[49] W.A. Harrison, Electronic structure and the properties of solids, Dover (1980).
[50] K. Huang, Introduction to statistical physics, CRC Press (2001).
[51] K. Huang, Quantum field theory, Wiley (1998).
[52] K. Huang, Quarks, leptons and gauge fields, World Scientific (1982).
[53] K. Huang, A superfluid universe, World Scientific (2017).
[54] M. Jimbo, A q-difference analog of $U(\mathfrak{g})$ and the Yang-Baxter equation, *Lett. Math. Phys.* **10** (1985), 63–69.

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

[56] V.F.R. Jones, A polynomial invariant for knots via von Neumann algebras, *Bull. Amer. Math. Soc.* **12** (1985), 103–111.

[57] V.F.R. Jones, Hecke algebra representations of braid groups and link polynomials, *Ann. of Math.* **126** (1987), 335–388.

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

[59] V.F.R. Jones, Subfactors and knots, AMS (1991).

[60] V.F.R. Jones, The Potts model and the symmetric group, in "Subfactors, Kyuzeso 1993" (1994), 259–267.

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

[62] V.F.R. Jones, The planar algebra of a bipartite graph, in "Knots in Hellas '98", World Sci. Publishing (2000), 94–117.

[63] V.F.R. Jones, The annular structure of subfactors, *Monogr. Enseign. Math.* **38** (2001), 401–463.

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

[65] T. Kibble and F.H. Berkshire, Classical mechanics, Imperial College Press (1966).

[66] C. Kittel, Introduction to solid state physics, Wiley (1953).

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

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

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

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

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

[72] R.K. Pathria and and P.D. Beale, Statistical mechanics, Elsevier (1972).

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

[74] D.V. Schroeder, An introduction to thermal physics, Oxford Univ. Press (1999).

[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] J. Smit, Introduction to quantum fields on a lattice, Cambridge Univ. Press (2002).

[79] J.R. Taylor, Classical mechanics, Univ. Science Books (2003).

[80] 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.

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

[82] J. von Neumann, Mathematical foundations of quantum mechanics, Princeton Univ. Press (1955).
[83] S. Wang, Free products of compact quantum groups, *Comm. Math. Phys.* **167** (1995), 671–692.
[84] S. Wang, Quantum symmetry groups of finite spaces, *Comm. Math. Phys.* **195** (1998), 195–211.
[85] A. Wassermann, Coactions and Yang-Baxter equations for ergodic actions and subfactors, *London Math. Soc. Lect. Notes* **136** (1988), 203–236.
[86] D. Weingarten, Asymptotic behavior of group integrals in the limit of infinite rank, *J. Math. Phys.* **19** (1978), 999–1001.
[87] S. Weinberg, Foundations of modern physics, Cambridge Univ. Press (2011).
[88] S. Weinberg, Lectures on quantum mechanics, Cambridge Univ. Press (2012).
[89] S. Weinberg, Lectures on astrophysics, Cambridge Univ. Press (2019).
[90] S. Weinberg, Cosmology, Oxford Univ. Press (2008).
[91] H. Weyl, The theory of groups and quantum mechanics, Princeton Univ. Press (1931).
[92] H. Weyl, The classical groups: their invariants and representations, Princeton Univ. Press (1939).
[93] H. Weyl, Space, time, matter, Princeton Univ. Press (1918).
[94] E. Witten, Quantum field theory and the Jones polynomial, *Comm. Math. Phys.* **121** (1989), 351–399.
[95] S.L. Woronowicz, Compact matrix pseudogroups, *Comm. Math. Phys.* **111** (1987), 613–665.
[96] S.L. Woronowicz, Tannaka-Krein duality for compact matrix pseudogroups. Twisted SU(N) groups, *Invent. Math.* **93** (1988), 35–76.
[97] J.M. Yeomans, Statistical mechanics of phase transitions, Oxford Univ. Press (1992).
[98] J. Zinn-Justin, Path integrals in quantum mechanics, Oxford Univ. Press (2004).
[99] J. Zinn-Justin, Phase transitions and renormalization group, Oxford Univ. Press (2005).
[100] B. Zwiebach, A first course in string theory, Cambridge Univ. Press (2004).

Index