Fourier analysis and waves

Teo Banica

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

2010 *Mathematics Subject Classification.* 35A09

Key words and phrases. Fourier analysis, Wave equation

Abstract. This is an introduction to the partial differential equations of physics, and their Fourier analysis, by focusing on the wave equation, and its versions. We first discuss the fundamentals of physics, including the wave equation. Then we discuss Fourier analysis, done mathematically, in detail, assuming basic calculus known, and its applications to various PDE, including the wave equation. Finally, we discuss some theoretical generalizations of both Fourier analysis, and the wave equation.

Preface

You surely know well calculus, so what comes next? If you are a student in mathematics, this should not be a concern, just follow the many classes that are offered there, and you will end up learning anything anyway. However, if you are into other sciences, and have to make a choice for your next math class, here is the situation:

(1) A first solid option, coming after calculus, is that of learning some measure theory, or advanced probability, or functional analysis, with all these being more or less the same thing. This is particularly interesting if doing something heavily relying on probability, such as statistical or quantum mechanics in physics, or for instance finance.

(2) Another good option is learning some algebra. Algebra can be of many kinds, and as examples here, if doing chemistry, or solid state physics, often dealing with symmetry, learning group theory would be useful. Also, if interested in computer science, learn some graphs, or other form of discrete mathematics, and some number theory.

(3) Yet another good option is learning some geometry. This is quite tough learning, because geometry is very wide, having been heavily developed by mathematicians, since ever, with many things to be learned. But this would be very rewarding, especially if you are a physicist, or an engineer, always struggling with weird curves and surfaces.

(4) Finally, yet another option is that of learning some advanced calculus, physics style. With this meaning more calculus, not along the line of measure theory, as evoked above, which is something quite mathematical, but rather tough, hardline, applied calculus. This would be a good option for instance if doing physics or chemistry.

The present book deals with this latter direction that calculus can take, namely advanced calculus, physics style. As already mentioned, this would be useful if you are into physics or chemistry. As for mathematicians, definitely stay with me. Calculus comes from physics, this is why Newton himself invented it, for solving questions from physics, and what you most likely learned so far, calculus stripped of physics, is just half of the story. Time to get now interested in the other half, involving physics, and with this being something very useful, no matter what kind of math you want to do.

Let me particularly insist on this latter fact, useful no matter what kind of math you want to do. First, because no one in the academia is excited about hiring math professors having weaknesses in calculus. And second, because of research. Believe me, even if interested in something seemingly away from calculus, such as logic, or abstract algebra, or design theory, knowing well calculus will enormously help your research.

Getting now to the present book, the topics that we want to discuss are basic physics, partial differential equations (PDE) coming from basic physics, and then solving these PDE via various advanced calculus methods, the main of which is Fourier analysis. We have organized all this around Fourier analysis, which is the main mathematical thing that we want to talk about, and around the wave equation, which is the main equation, or at least the one that I prefer, in physics. The book has 4 parts, as follows:

I - General physics, with the story told by focusing on waves. That is, we will first learn classical mechanics, motivated by mechanical waves, then fluid mechanics, motivated by sound waves, and then electrodynamics, motivated by electromagnetic waves. Thermodynamics will be not forgotten either, motivated by, well, heat waves.

II - Fourier analysis, developed very mathematically, as a continuation of the calculus that you already know. Technically this requires a bit of measure theory and functional analysis, disciplines that could easily take one full class, and here we will short-circuit a bit all this, with some simplified arguments. As illustrations for all this, 1D waves.

III - Partial differential equations (PDE), mainly the waves ones found in Part I, solved with techniques from Part II. Again, this part will be rather mathematical, and with various ramifications, depending on the precise equations that we have to study, and notably on the differences between the wave equation, and the heat equation.

IV - Generalizations. This last part will be somewhere between mathematics and physics, mostly mathematical, but with potential applications to many physics problems, which are more specialized. We will discus here the Fourier transform over arbitrary groups, and its various generalizations, and their potential applications to physics.

It is a pleasure to thank my algebra colleagues, at the many universities I have been to. Normally I know some algebra, but with such a fight everywhere on who has the algebra classes, and with me being lazy to fight, I always end up teaching analysis. Which is now rewarding, look at this book that I wrote, about it. Many thanks go as well to my cats, who definitely know well Fourier analysis, and helped me a lot, with all this.

Contents

Part I

Physics, waves

Choo, choo, train
Tracking down the track
Gotta travel on
Ain't never coming back

CHAPTER 1

Mechanical waves

1a. Chains and whips

Welcome to the waves. Or perhaps to general physics, to the main partial differential equations (PDE) coming from general physics, to the functional and Fourier analysis needed for studying these PDE, with this being also known as "advanced calculus", and why not to some algebra and group theory as well, in order to do some further Fourier tricks, for some further waves. So many things this book is advertised for, and I must admit myself that, now that I started to write page one, I am a bit confused.

The plan is to do physics and phenomenology in the present Part I, and leave mathematics for later. That is, in this Part I we will do hardline physics, all delicate modelling work, in order to reach to as many interesting PDE as possible, and with these PDE of course as close to the real-life physics as possible. And we will not bother much with solving these PDE, unless of course we have some quick, trivial things to say, usually in 1D, we have the whole middle of the book, Parts II and III, for doing that.

Before starting, a piece of advice, depending on what kind of student you are:

(1) If genuinely interested in physics, chemistry or engineering, you surely know well all this phenomenology stuff, perhaps even better than myself, so you can safely skip this Part I, and go ahead with Parts II and III. That is indeed useful mathematics, explained by your favorite math professor, and that's what I have interesting to say, to you.

(2) On the opposite, if you are into mathematics, mathematical physics, perhaps even theoretical physics, stay with me. This Part I is what I have interesting to say to you, this is all tricky phenomenology that you must know, and in absolutely any case, not be clueless about. As for Parts II and III, these are just some math, routine things.

Getting started now, our idea for this Part I will be to have a tour through basic physics, guided by the waves. 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 physics.

In practice now, 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 the fearsome electrodynamics, with its Maxwell equations. Exciting enough you would say, and with the remark of course that in each case, the equation $\ddot{\varphi} = v^2\Delta\varphi$ must be fine-tuned a bit, and we will discuss this too.

Finally, let us mention that advanced thermodynamics will be not forgotten, we will talk about this in chapter 4, by adopting the main equation there, called heat diffusion equation, and which is 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 as well, as a close cousin of this heat equation.

Getting started for good now, 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 modelling problem, and we find:

THEOREM 1.2. *We have the wave equation, namely*

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

where Δ is the Laplace operator on the functions $\varphi : \mathbb{R}^N \to \mathbb{C}$, given by

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

and $v > 0$ is the propagation speed, obtained via chains.

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, $N \to \infty$. Alternatively, we can speculate as well directly on a continuous model, of whip type, again via classical mechanics. □

In higher dimensions, pretty much the same thing happens, as stated above.

1b. Springs, Hooke

We present here 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:

THEOREM 1.3. *The wave equation in $\mathbb{R}^N$ is*

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

where Δ is the Laplace operator, given by

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

and $v > 0$ is the propagation speed.

PROOF. As already mentioned, 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 $N >> 0$ balls, having a total mass M, and

spanning a total distance L. Thus, our previous infinitesimal parameters are as follows, with K being the spring constant of the total system, which is of course lower than k:

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

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

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

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

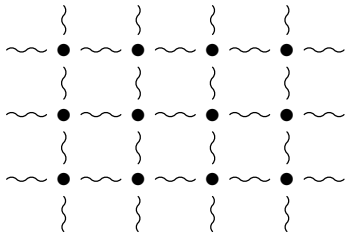
$$\ddot{\varphi}(x) = \frac{KL^2}{M} \cdot \frac{\varphi(x+l) - 2\varphi(x) + \varphi(x-l)}{l^2}$$

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

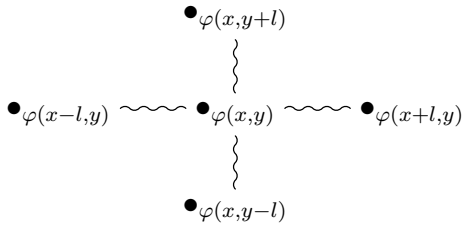
$$\ddot{\varphi}(x) = \frac{KL^2}{M} \cdot \frac{d^2\varphi}{dx^2}(x)$$

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 $N^2 >> 0$ 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:

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

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

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

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:

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

Thus, we are led in this way 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. □

Observe that there are some subtleties in the above, with the convention for the total spring constant K, which varies with the dimension N. We will be back to this.

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, leads to different results.

1c. Bars and elasticity

In order to reach to some further insight into our spring models above, we must get into elesticity. Indeed, the Hooke law that we used has behind it some non-trivial elasticity, of "linear" type, and understanding all this, and further modifying our models, according to what elasticity theory 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 go into elasticity.

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

THEOREM 1.4. *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. □

1d. D'Alembert formula

Time for some math. As explained in the beginning of this chapter, our policy in the present Part I will be that of not solving our PDE, unless we have something really simple to say. And in what regards the wave equation, in 1D we have the following result:

THEOREM 1.5. *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 lattice model discretization, 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:

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

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 out 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. □

In $N\geq 2$ dimensions things are more complicated. We will be back to this.

1e. Exercises

Exercises:

Exercise 1.6.

Exercise 1.7.

Exercise 1.8.

Exercise 1.9.

Exercise 1.10.

Exercise 1.11.

Bonus exercise.

CHAPTER 2

Water and sound

2a. Gases, pressure

The material in chapter 1 was quite theoretical, and we discuss here how mechanical waves exactly propagate through solids, liquids and gases, and how the general wave equation must be corrected, in each case. For this, we first need a crash course in thermodynamics, leading to the theory of states of matter. Let us start with:

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 has, 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".

In order to understand how all this works, and as a first result now, dealing with pressure only, and for the gases without collisions between molecules, we have:

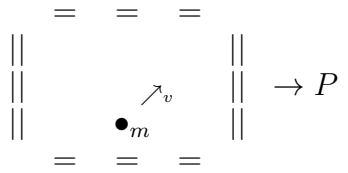
THEOREM 2.2. *The pressure P, volume V and total kinetic energy K of a gas, having point molecules, with no collisions between them, satisfy*

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

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:

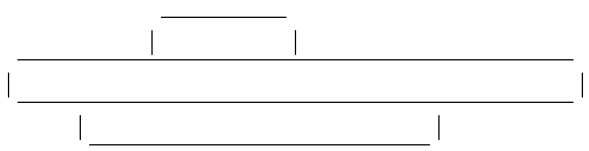
$$P = \frac{m||v||^2}{dV} = \frac{2K}{dV}$$

(4) It remains to extend our result to arbitrary volume shapes. For this purpose, let us first redo the above computations for a parallelepiped, $V = L_1 L_2 L_3$. Here the above 1D gas computation carries on, and gives the same result, as follows:

$$P = \frac{F}{L_2 L_3} = \frac{\Delta p}{L_2 L_3 \Delta t} = \frac{2mv}{L_2 L_3 \cdot 2L_1/v} = \frac{mv^2}{L_1 L_2 L_3} = \frac{2K}{V}$$

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:

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

(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. □

Along the same lines, but taking now into account the collisions between gas molecules, and also dealing with the internal mechanism of the $PV = kT$ formula, we have:

THEOREM 2.3 (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. □

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 $V = vol(B) \in \mathbb{R}$.

Let us first record, regarding the gases, the following key result:

THEOREM 2.4. *Beyond the ideal gas setting, stating that we should have $PV = kT$, the gases are subject to the Van der Waals equation*

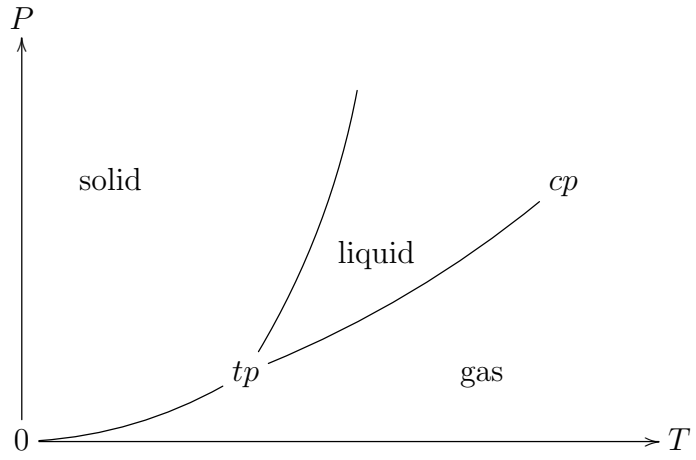
$$\left(P + \frac{\alpha}{V^2}\right)(V - \beta) = kT$$

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

PROOF. This is something quite tricky, with both the parameters $\alpha, \beta > 0$ coming naturally from the kinetic viewpoint, and also of course, from experiments. □

The Van der Waals gases have some interesting points on the isothermals, called triple and critical points of the gas. By further buidling on this, we are led to:

FACT 2.5. *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.

Many other things can be said here, especially regarding the extremes.

2c. Water waves, sound

Generally speaking, fluids are by definition the non-solids, and as such, they fall into liquids, gases and plasma. 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 all 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 2.6. *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 2.7. *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:

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

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:

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

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.

For further fun, we can add magnetism to our incompressible fluids. However, in regards with discrete modelling questions, all this looks quite complicated.

2d. Seismic waves

Soil liquefaction and related topics. Seismic waves.

2e. Exercises

Exercises:

Exercise 2.8.

Exercise 2.9.

Exercise 2.10.

Exercise 2.11.

Exercise 2.12.

Exercise 2.13.

Bonus exercise.

CHAPTER 3

Electromagnetic waves

3a. Radiation, light

Quite remarkably, the wave equation, which was mechanical so far, appears as well in the context of electromagnetism. In order to talk about this, we will need:

THEOREM 3.1. *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 fundamental, appearing as a tricky mixture of physics facts and mathematical results, the idea being as follows:

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

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

(3) Fortunately, there is a full set of equations relating the electric field E and the magnetic field B, those above. Regarding the math, the dots denote derivatives with respect to time, and ∇ is the gradient operator, or space derivative, given by:

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

(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. □

We will need the following consequence of the Maxwell equations:

THEOREM 3.2. *In regions of space where there is no charge or current present the Maxwell equations for electrodynamics read*

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

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

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

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

where $\Delta = \sum_i d^2/dx_i^2$ is the Laplace operator, and c is the speed of light.

PROOF. Under the circumstances in the statement, namely no charge or current present, the Maxwell equations in Theorem 3.1 simply read:

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

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

Here we have used the following famous formula, due to Biot and Savart:

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

Now by applying the curl operator to the last two equations, we obtain:

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

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

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

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

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

So, what is light? Light is the wave predicted by Theorem 3.2, travelling in vacuum at the maximum possible speed, c, and with an important extra property being that it depends on a real positive parameter, that can be called, upon taste, frequency, wavelength, or color. And in what regards the creation of light, the mechanism here is as follows:

FACT 3.3. *An accelerating or decelerating charge produces electromagnetic radiation, called light, whose frequency and wavelength can be explicitely computed.*

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

Getting back now to Fact 3.3, in its general form, as stated above, this is something which can be deduced via some math, based on the Maxwell equations.

3b. Frequency, color

Moving ahead, let us go back to the wave equation $\ddot{\varphi} = v^2\Delta\varphi$ from Theorem 3.2, and try to understand its simplest solutions. In 1D, the situation is as follows:

THEOREM 3.4. *The 1D wave equation, with speed v, namely*

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

has as basic solutions the following functions,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

DEFINITION 3.5. *A monochromatic plane wave is a solution of the 3D wave equation which moves in only 1 direction, making it in practice a solution of the 1D wave equation, and which is of the special from found in Theorem 3.4, with no frequencies mixed.*

In other words, we are making here two assumptions on our wave. First is the 1-dimensionality assumption, which gets us into the framework of Theorem 3.4. And second is the assumption, in connection with the Fourier decomposition result from the end of Theorem 3.4, that our solution is of "pure" type, meaning a wave having a well-defined wavelenght and frequency, instead of being a "packet" of such pure waves.

All this is still mathematics, and making now the connection with physics and electromagnetism, and more specifically with Theorem 3.2 and Fact 3.3, we have:

FACT 3.6. *Physically speaking, a monochromatic plane wave is the electromagnetic radiation appearing as in Theorem 3.2 and Fact 3.3, via equations of type*

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

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

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

To be more precise, what we are doing here is to import the conclusions of our mathematical discussion so far, from Theorem 3.4 and Definition 3.5, into the context of our original physics discussion, from Fact 3.3. And also to add an extra twist coming from physics, and more specifically from the notion of polarization. More on this later.

In any case, we have now a decent intuition about what light is, and more on this later, and let us discuss now the examples. The idea is that we have various types of light, depending on frequency and wavelength. These are normally referred to as

"electromagnetic waves", but for keeping things simple and luminous, we will keep using the familiar term "light". The classification, in a rough form, is as follows:

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

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

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

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

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

Importantly, both ends of the table are a bit loose. At the high frequency end there are some restrictions coming from quantum mechanics, and more on them later. As for the low frequency end, what's wave and what's not is a bit of a philosophical question,

but which is actually not that philosophical, because waves having huge wavelengths can easily turn around mountains, full countries and so on, and so are of military interest. Secret research here, more of engineering type of course, is still ongoing.

3c. Snell law, optics

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

FACT 3.7. *Inside a linear, homogeneous medium, where there is no free charge or current present, the Maxwell equations for electrodynamics read*

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

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

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

Observe that this statement is precisely the first part of Theorem 3.2, with the vacuum constants ε_0, μ_0 being now replaced by their versions ε, μ, concerning the medium in question. In what regards now the second part of Theorem 3.2, we have:

THEOREM 3.8. *Inside a linear, homogeneous medium, where there is no free charge or free current present, both E and B are subject to the wave equation*

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

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

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

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

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

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

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

As a first observation here, while the above is something quite trivial, mathematically speaking, from the physical viewpoint we are here into complicated things. Materials can be transparent or opaque, with the distinction between them being something very subtle, and advanced, and Theorem 3.8 obviously deals with the transparent case.

In short, we are here inside advanced materials theory, that we cannot really understand, with our knowledge so far. In what follows we will be interested in transparent materials only, such as glass. Regarding the other materials, such as rock, let us just mention that light dissapears inside them, converted into heat. Of course glass heats too when light crosses it, with this being related to $v < c$ inside it. More on this later.

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

FACT 3.9. *When travelling through a material, and hitting a new material, some of the light gets reflected, at the same angle, and some of it gets refracted, at a different angle, depending both on the old and the new material, and on the wavelength.*

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

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

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

Here $n_i(\lambda)$ are the refraction indices of the two materials, adjusted for the wavelength, and with this adjustment for wavelength being the whole point, which is something quite complicated. For an introduction to all this, we refer for instance to Griffiths [43].

As a simple consequence of the above, we have:

THEOREM 3.10. *Light can be decomposed, by using a prism.*

PROOF. This follows from Fact 3.9. Indeed, when hitting a piece of glass, provided that the hitting angle is not 90°, the light will decompose over the wavelenghts present, with the corresponding refraction angles depending on these wavelengths. And we can capture these split components at the exit from the piece of glass, again deviated a bit, provided that the exit surface is not parallel to the entry surface. And the simplest device doing the job, that is, having two non-parallel faces, is a prism. □

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

FACT 3.11. *We can study events via spectroscopy, by capturing the light the event has produced, decomposing it with a prism, carefully recording its "spectral signature", consisting of the wavelenghts present, and their density, and then doing some reverse engineering, consisting in reconstructing the event out of its spectral signature.*

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

3d. Polarized light

Polarized light, collisions, other topics.

3e. Exercises

Exercises:

Exercise 3.12.

Exercise 3.13.

Exercise 3.14.

Exercise 3.15.

Exercise 3.16.

Exercise 3.17.

Bonus exercise.

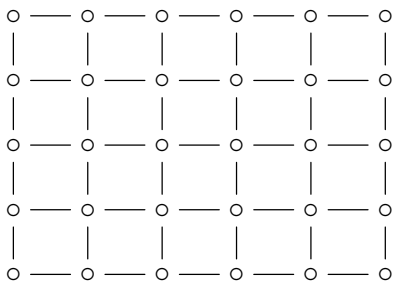
CHAPTER 4

Heat waves

4a. Basic diffusion

For our discussion of general physics to be complete, we must talk as well about thermodynamics, and heat diffusion. Here the main equation is a bit different from the wave one, namely $\dot{\varphi} = \alpha\Delta\varphi$, but we will agree to adopt it, 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 as well, as a close cousin of this heat equation.

The simplest heat diffusion question, 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, in this setting, and obviously, we can model this by focusing on the membrane, with a basic grid model for it, as follows:



There is some sort of "game" played by the two gases, over this grid, and we can model 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 modelling. Also, in this setting, we can actually keep the membrane, but allow it to inflate.

4b. The heat equation

Let us go now into heavier, fully powerful models and equations for the heat diffusion mechanism, involving this time more advanced mathematics and physics.

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:

$$\text{---}\circ_{x-l}\overset{l}{\text{---}}\circ_{x}\overset{l}{\text{---}}\circ_{x+l}\text{---}$$

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:

$$\frac{\varphi(x,t+\delta)-\varphi(x,t)}{\delta}=\frac{\alpha}{l^2}\sum_{x\sim y}\left[\varphi(y,t)-\varphi(x,t)\right]$$

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:

$$\frac{\varphi(x,t+\delta)-\varphi(x,t)}{\delta}=\alpha\cdot\varphi''(x,t)$$

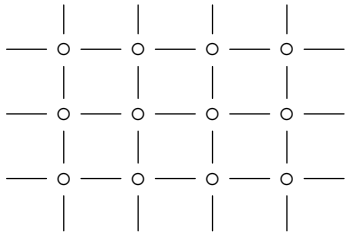
Now with $t \to 0$, we are led in this way to the heat equation, namely:

$$\dot{\varphi}(x,t)=\alpha\cdot\varphi''(x,t)$$

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.

4c. Stefan-Boltzmann

There are some corrections to the heat equation to be made, due to Stefan-Boltzmann, and we will explain this now, with the story being something quite exciting.

Our discussion so far implicitely used the fact that heat is light too, and so time now, before getting into quantum mechanics, to get back to the theory of heat, as developed before, and see what our new viewpoint on it can bring. The main problem here is to compute the radiation of black bodies, and we will discuss this now.

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 4.2. *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:

$$\mathcal{E} = \frac{V}{16\pi} \sum_n (< e_n, e_n > + < \bar{e}_n, \bar{e}_n >)$$

(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 4.3. *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:

$$a=\frac{16\pi^8 b^4}{15h^3c^3}$$

PROOF. This is something quite technical, obtained along the lines of the proof of Theorem 4.2, 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.

In what concerns us, we can now go back to the heat equation, and add a correction term, accounting for heat loss via radiation, due to Stefan-Boltzmann.

4d. The heat kernel

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 4.4. *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 now the resolution of the heat equation, we have here:

THEOREM 4.5. *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. □

Regarding now our discretization questions, things here are quite tricky. In relation with Theorem 4.5, 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:

DEFINITION 4.6. *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 4.7. *With $F_f(x) = \mathbb{E}(e^{ixf})$ 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 4.8 (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, $F_f(x) = \mathbb{E}(e^{ixf})$ is given by the following formula:

$$\begin{aligned} F_f(x) &= \mathbb{E}\left(\sum_{k=0}^{\infty}\frac{(ixf)^k}{k!}\right) \\ &= \sum_{k=0}^{\infty}\frac{(ix)^k\mathbb{E}(f^k)}{k!} \\ &= \sum_{k=0}^{\infty}\frac{i^kM_k(f)}{k!}\,x^k \end{aligned}$$

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.

Finally, as a related question, we can talk as well about the Schrödinger equation, as being some sort of heat equation, in imaginary time. We will be back to this.

4e. Exercises

Exercises:

EXERCISE 4.9.

EXERCISE 4.10.

EXERCISE 4.11.

EXERCISE 4.12.

EXERCISE 4.13.

EXERCISE 4.14.

Bonus exercise.

Part II

Fourier analysis

See the white light
The light within
Be your own disciple
Fan the sparks of will

CHAPTER 5

Calculus, revised

5a. Partial derivatives

Let us first discuss the differentiability in several variables. At order 1, the situation is quite similar to the one in 1 variable, but this time involving matrices. In order to explain this material, let us start with a straightforward definition, as follows:

DEFINITION 5.1. *We say that a map $f : \mathbb{R}^N \to \mathbb{R}^M$ is differentiable at $x \in \mathbb{R}^N$ if*

$$f(x+t) \simeq f(x) + f'(x)t$$

for some linear map $f'(x) : \mathbb{R}^N \to \mathbb{R}^M$, called derivative of f at the point $x \in \mathbb{R}^N$.

But is this the correct definition. I can hear you screaming that we are probably going the wrong way, because for functions $f : \mathbb{R} \to \mathbb{R}$ the derivative is something much simpler, as follows, and that we should try to imitate, in our higher dimensional setting:

$$f'(x) = \lim_{t \to 0} \frac{f(x+t) - f(x)}{t}$$

However, this is not possible. So, moving ahead, we know that the linear maps correspond to rectangular matrices $A \in M_{M \times N}(\mathbb{R})$, and we are led in this way to:

QUESTION 5.2. *Given a differentiable map $f : \mathbb{R}^N \to \mathbb{R}^M$, in the abstract sense of Definition 5.1, what exactly is its derivative*

$$f'(x) : \mathbb{R}^N \to \mathbb{R}^M$$

regarded as a rectangular matrix, $f'(x) \in M_{M \times N}(\mathbb{R})$?

Again, I might hear scream you here, arguing that you come after a long battle, just agreeing that the derivative is a linear map, and not a number, and now what, we are trying to replace this linear map by a matrix, and so by a bunch of numbers.

Good point, and I must admit that I have no good answer to this. In fact, what we are doing here are quite deep things, that took mankind several centuries to develop, and that we are now presenting in a compressed form. So yes, all this is difficult mathematics, when you first see it, I perfectly agree with you. In any case, hope that you're still with me, and in order to further clarify all this, here is the answer to Question 5.2:

THEOREM 5.3. *The derivative of a differentiable function $f : \mathbb{R}^N \to \mathbb{R}^M$, making*

$$f(x+t) \simeq f(x) + f'(x)t$$

work, is the matrix of partial derivatives at x, namely

$$f'(x) = \left(\frac{df_i}{dx_j}(x) \right)_{ij} \in M_{M \times N}(\mathbb{R})$$

acting on the vectors $t \in \mathbb{R}^N$ by usual multiplication.

PROOF. As a first observation, the formula in the statement makes sense indeed, as an equality, or rather approximation, of vectors in $\mathbb{R}^M$, as follows:

$$f \begin{pmatrix} x_1 + t_1 \\ \vdots \\ x_N + t_N \end{pmatrix} \simeq f \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \frac{df_1}{dx_1}(x) & \ldots & \frac{df_1}{dx_N}(x) \\ \vdots & & \vdots \\ \frac{df_M}{dx_1}(x) & \ldots & \frac{df_M}{dx_N}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix}$$

In order to prove now this formula, which does make sense, the idea is as follows:

(1) First of all, at $N = M = 1$ what we have is a usual 1-variable function $f : \mathbb{R} \to \mathbb{R}$, and the formula in the statement is something that we know well, namely:

$$f(x+t) \simeq f(x) + f'(x)t$$

(2) Let us discuss now the case $N = 2, M = 1$. Here what we have is a function $f : \mathbb{R}^2 \to \mathbb{R}$, and by using twice the basic approximation result from (1), we obtain:

$$\begin{aligned} f \begin{pmatrix} x_1 + t_1 \\ x_2 + t_2 \end{pmatrix} &\simeq f \begin{pmatrix} x_1 + t_1 \\ x_2 \end{pmatrix} + \frac{df}{dx_2}(x) t_2 \\ &\simeq f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{df}{dx_1}(x) t_1 + \frac{df}{dx_2}(x) t_2 \\ &= f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \frac{df}{dx_1}(x) & \frac{df}{dx_2}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \end{aligned}$$

(3) More generally, we can deal in this way with the general case $M = 1$, with the formula here, obtained via a straightforward recurrence, being as follows:

$$\begin{aligned} f \begin{pmatrix} x_1 + t_1 \\ \vdots \\ x_N + t_N \end{pmatrix} &\simeq f \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \frac{df}{dx_1}(x) t_1 + \ldots + \frac{df}{dx_N}(x) t_N \\ &= f \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \frac{df}{dx_1}(x) & \ldots & \frac{df}{dx_N}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix} \end{aligned}$$

(4) But this gives the result in the case where both $N, M \in \mathbb{N}$ are arbitrary too. Indeed, consider a function $f : \mathbb{R}^N \to \mathbb{R}^M$, and let us write it as follows:

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_M \end{pmatrix}$$

We can apply (3) to each of the components $f_i : \mathbb{R}^N \to \mathbb{R}$, and we get:

$$f_i \begin{pmatrix} x_1 + t_1 \\ \vdots \\ x_N + t_N \end{pmatrix} \simeq f_i \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \frac{df_i}{dx_1}(x) & \ldots & \frac{df_i}{dx_N}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix}$$

But this collection of M formulae tells us precisely that the following happens, as an equality, or rather approximation, of vectors in $\mathbb{R}^M$:

$$f \begin{pmatrix} x_1 + t_1 \\ \vdots \\ x_N + t_N \end{pmatrix} \simeq f \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \frac{df_1}{dx_1}(x) & \ldots & \frac{df_1}{dx_N}(x) \\ \vdots & & \vdots \\ \frac{df_M}{dx_1}(x) & \ldots & \frac{df_M}{dx_N}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix}$$

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

The above result, while nice, clear and useful, does not close the discussion regarding the derivative. The result here, which is something quite technical, is as follows:

THEOREM 5.4. *For a function $f : X \to \mathbb{R}^M$, with $X \subset \mathbb{R}^N$, the following conditions are equivalent, and in this case we say that f is continuously differentiable:*

(1) *f is differentiable, and the map $x \to f'(x)$ is continuous.*
(2) *f has partial derivatives, which are continuous with respect to $x \in X$.*

If these conditions are satisfied, $f'(x)$ is the matrix fomed by the partial derivatives at x.

PROOF. We already know, from Theorem 5.3, that the last assertion holds. Regarding now the proof of the equivalence, this goes as follows:

(1) $\implies$ (2) Assuming that f is differentiable, we know from Theorem 5.3 that $f'(x)$ is the matrix fomed by the partial derivatives at x. Thus, for any $x, y \in X$:

$$\frac{df_i}{dx_j}(x) - \frac{df_i}{dx_j}(y) = f'(x)_{ij} - f'(y)_{ij}$$

By applying now the absolute value, we obtain from this the following estimate:

$$\begin{aligned} \left| \frac{df_i}{dx_j}(x) - \frac{df_i}{dx_j}(y) \right| &= |f'(x)_{ij} - f'(y)_{ij}| \\ &= |(f'(x) - f'(y))_{ij}| \\ &\leq ||f'(x) - f'(y)|| \end{aligned}$$

But this gives the result, because if the map $x \to f'(x)$ is assumed to be continuous, then the partial derivatives follow to be continuous with respect to $x \in X$.

(2) $\implies$ (1) This is something more technical. For simplicity, let us assume $M = 1$, the proof in general being similar. Given $x \in X$ and $\varepsilon > 0$, let us pick $r > 0$ such that the ball $B = B_x(r)$ belongs to X, and such that the following happens, over B:

$$\left| \frac{df}{dx_j}(x) - \frac{df}{dx_j}(y) \right| < \frac{\varepsilon}{N}$$

Our claim is that, with this choice made, we have the following estimate, for any $t \in \mathbb{R}^N$ satisfying $||t|| < r$, with A being the vector of partial derivatives at x:

$$|f(x+t) - f(x) - At| \leq \varepsilon ||t||$$

In order to prove this claim, the idea will be that of suitably applying the mean value theorem, over the N directions of $\mathbb{R}^N$. Indeed, consider the following vectors:

$$t^{(k)} = \begin{pmatrix} t_1 \\ \vdots \\ t_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

In terms of these vectors, we have the following formula:

$$f(x+t) - f(x) = \sum_{j=1}^{N} f(x + t^{(j)}) - f(x + t^{(j-1)})$$

Also, the mean value theorem gives a formula as follows, with $s_j \in [0, 1]$:

$$f(x + t^{(j)}) - f(x + t^{(j-1)}) = \frac{df}{dx_j}(x + s_j t^{(j)} + (1 - s_j) t^{(j-1)}) \cdot t_j$$

But, according to our assumption on $r > 0$ from the beginning, the derivative on the right differs from $\frac{df}{dx_j}(x)$ by something which is smaller than ε/N:

$$\left| \frac{df}{dx_j}(x + s_j t^{(j)} + (1 - s_j) t^{(j-1)}) - \frac{df}{dx_j}(x) \right| < \frac{\varepsilon}{N}$$

Now by putting everything together, we obtain the following estimate:

$$\begin{aligned}
|f(x+t)-f(x)-At| &= \left|\sum_{j=1}^{N} f(x+t^{(j)})-f(x+t^{(j-1)})-\frac{df}{dx_j}(x)\cdot t_j\right| \\
&\leq \sum_{j=1}^{N}\left|f(x+t^{(j)})-f(x+t^{(j-1)})-\frac{df}{dx_j}(x)\cdot t_j\right| \\
&= \sum_{j=1}^{N}\left|\frac{df}{dx_j}(x+s_jt^{(j)}+(1-s_j)t^{(j-1)})\cdot t_j-\frac{df}{dx_j}(x)\cdot t_j\right| \\
&= \sum_{j=1}^{N}\left|\frac{df}{dx_j}(x+s_jt^{(j)}+(1-s_j)t^{(j-1)})-\frac{df}{dx_j}(x)\right|\cdot|t_j| \\
&\leq \sum_{j=1}^{N}\frac{\varepsilon}{N}\cdot|t_j| \\
&\leq \varepsilon||t||
\end{aligned}$$

Thus we have proved our claim, and this gives the result. □

In practice now, our first task will be that of extending to several variables our basic results from one-variable calculus. As a standard result here, we have:

THEOREM 5.5. *We have the chain derivative formula*

$$(f\circ g)'(x)=f'(g(x))\cdot g'(x)$$

as an equality of matrices.

PROOF. This is something standard in one variable, and in several variables the proof is similar, by using the notion of derivative coming from Theorem 5.3. To be more precise, consider a composition of functions, as follows:

$$f:\mathbb{R}^N\to\mathbb{R}^M\quad,\quad g:\mathbb{R}^K\to\mathbb{R}^N\quad,\quad f\circ g:\mathbb{R}^K\to\mathbb{R}^M$$

According to Theorem 5.3, the derivatives of these functions are certain linear maps, corresponding to certain rectangular matrices, as follows:

$$f'(g(x))\in M_{M\times N}(\mathbb{R})\quad,\quad g'(x)\in M_{N\times K}(\mathbb{R})\qquad(f\circ g)'(x)\in M_{M\times K}(\mathbb{R})$$

Thus, our formula makes sense indeed. As for proof, this comes from:

$$\begin{aligned}
(f\circ g)(x+t) &= f(g(x+t)) \\
&\simeq f(g(x)+g'(x)t) \\
&\simeq f(g(x))+f'(g(x))g'(x)t
\end{aligned}$$

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

The above formula is quite interesting, and working its main particular cases, when some of parameters M, N, K equal 1, is a very instructive exercice. But, the best is to rather ignore all these particular cases, and use Theorem 5.5 as stated.

Moving ahead now, we have several useful consequences of the above result, a standard application, generalizing things that we know in one variable, being as follows:

THEOREM 5.6. *Assuming that $f : X \to \mathbb{R}^M$ is differentiable, with $X \subset \mathbb{R}^N$ being convex, we have the estimate*

$$||f(x) - f(y)|| \leq M||x - y||$$

for any $x, y \in X$, where the quantity on the right is given by:

$$M = \sup_{x \in X} ||f'(x)||$$

Moreover, this estimate can be sharp, for instance for the linear functions.

PROOF. This is something quite tricky, which in several variables cannot be proved with bare hands. However, we can get it by using our chain derivative formula. Consider indeed the path $\gamma : [0, 1] \to \mathbb{R}^M$ given by the following formula:

$$\gamma(t) = tx + (1 - t)y$$

Now let us set $g(t) = f(\gamma(t))$. We have then, according to the chain rule formula:

$$\begin{aligned} g'(t) &= f'(\gamma(t))\gamma'(t) \\ &= f'(\gamma(t))(x - y) \end{aligned}$$

But this gives the following estimate, with $M > 0$ being as in the statement:

$$\begin{aligned} |g'(t)| &\leq ||f'(\gamma(t))|| \cdot ||x - y|| \\ &\leq M||x - y|| \end{aligned}$$

Now by using one-variable results that we know, we obtain from this:

$$||g(1) - g(0)|| \leq ||M|| \cdot ||x - y||$$

But since we have $g(1) = f(x), g(0) = f(y)$, this gives the formula in the statement. Finally, the last assertion is clear. □

In analogy with what we know in 1 variable, we can formulate:

THEOREM 5.7. *The Taylor formula at order 1 for a function $f : \mathbb{R}^N \to \mathbb{R}$ is*

$$f(x + t) \simeq f(x) + f'(x)t$$

and in particular, in order for x to be a local extremum, we must have $f'(x) = 0$.

PROOF. Here the first assertion is something that we know, as explained above, and the second assertion follows from it. Indeed, let us look at the order 1 term, given by:

$$f'(x)t = \sum_{i=1}^{N} \frac{df}{dx_i}\, t_i$$

Now since this linear combination of the entries of $t \in \mathbb{R}^N$ can range among positives and negatives, unless all the coefficients are zero, which means $f'(x) = 0$, we are led to the conclusion that local extremum needs $f'(x) = 0$ to hold, as stated. □

Let us discuss now the Taylor formula at order 2. We have here:

THEOREM 5.8. *Given a twice differentiable function $f : \mathbb{R}^N \to \mathbb{R}$, we have*

$$f(x+t) \simeq f(x) + f'(x)t + \frac{< f''(x)t, t >}{2}$$

where $f''(x) \in M_N(\mathbb{R})$ stands as usual for the Hessian matrix.

PROOF. This is something more tricky, the idea being as follows:

(1) As a first observation, at $N = 1$ the Hessian matrix is the 1×1 matrix having as entry the second derivative $f''(x)$, and the formula in the statement hold indeed, as:

$$f(x+t) \simeq f(x) + f'(x)t + \frac{f''(x)t^2}{2}$$

(2) In general now, this is in fact something which does not need a new proof, because it follows from the one-variable formula above, applied to the restriction of f to the following segment in $\mathbb{R}^N$, which can be regarded as being a one-variable interval:

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

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

$$g(r) = f(x+ry)$$

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

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

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

(3) 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:

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

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

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

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

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

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

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

Thus, we have obtained the formula in the statement.

(5) Finally, for completness, let us record as well a more numeric formulation of what we found. According to our usual rules for matrix calculus, what we found is:

$$f(x+t) \simeq f(x) + \sum_{i=1}^N \frac{df}{dx_i}\,t_i + \frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N \frac{d^2f}{dx_idx_j}\,t_it_j$$

Observe that, since the Hessian matrix $f''(x)$ is symmetric, most of the terms on the right will appear in pairs, making it clear what the $1/2$ is there for, namely avoiding redundancies. However, this is only true for the off-diagonal terms, so instead of further messing up our numeric formula above, we will just leave it like this. □

Many other things can be said, as a continuation of the above.

5b. Multiple integrals

Multiple integrals.

5c. Spaces of functions

Let us study now the spaces that the functions $f : \mathbb{R} \to \mathbb{C}$ can form. These functions can be continuous, differentiable, infinitely differentiable, and so on, but there are many more properties that these functions can have, that we will investigate now. This will lead to various spaces of functions $f : \mathbb{R} \to \mathbb{C}$, that can be used, among others, in order to well-define the Fourier transform operation $f \to \widehat{f}$. We will need:

THEOREM 5.9. *Given a convex function $f : \mathbb{R} \to \mathbb{R}$, we have the following Jensen inequality, for any $x_1, \ldots, x_N \in \mathbb{R}$, and any $\lambda_1, \ldots, \lambda_N > 0$ summing up to 1,*

$$f(\lambda_1 x_1 + \ldots + \lambda_N x_N) \leq \lambda_1 f(x_1) + \ldots + \lambda_N x_N$$

with equality when $x_1 = \ldots = x_N$. In particular, by taking the weights λ_i to be all equal, we obtain the following Jensen inequality, valid for any $x_1, \ldots, x_N \in \mathbb{R}$,

$$f\left(\frac{x_1 + \ldots + x_N}{N}\right) \leq \frac{f(x_1) + \ldots + f(x_N)}{N}$$

and once again with equality when $x_1 = \ldots = x_N$. We have a similar statement holds for the concave functions, with all the inequalities being reversed.

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

(1) First, we can talk about convex functions in a usual, intuitive way, with this meaning by definition that the following inequality must be satisfied:

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}$$

(2) But this means, via a simple argument, by approximating numbers $t \in [0, 1]$ by sums of powers 2^{-k}, that for any $t \in [0, 1]$ we must have:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

Alternatively, via yet another simple argument, this time by doing some geometry with triangles, this means that we must have:

$$f\left(\frac{x_1 + \ldots + x_N}{N}\right) \leq \frac{f(x_1) + \ldots + f(x_N)}{N}$$

But then, again alternatively, by combining the above two simple arguments, the following must happen, for any $\lambda_1, \ldots, \lambda_N > 0$ summing up to 1:

$$f(\lambda_1 x_1 + \ldots + \lambda_N x_N) \leq \lambda_1 f(x_1) + \ldots + \lambda_N x_N$$

(3) Summarizing, all our Jensen inequalities, at $N = 2$ and at $N \in \mathbb{N}$ arbitrary, are equivalent. The point now is that, if we look at what the first Jensen inequality, that we took as definition for the convexity, means, this is simply equivalent to:

$$f''(x) \geq 0$$

(4) Thus, we are led to the conclusions in the statement, regarding the convex functions. As for the concave functions, the proof here is similar. Alternatively, we can say that f is concave precisely when $-f$ is convex, and get the results from what we have. □

As a second result on the subject, which is very classical as well, we have:

THEOREM 5.10. *For $p \in (1,\infty)$ we have the following Hölder inequality,*

$$\left|\frac{x_1+\ldots+x_N}{N}\right|^p \leq \frac{|x_1|^p+\ldots+|x_N|^p}{N}$$

and for $p \in (0,1)$ we have the following reverse Hölder inequality,

$$\left|\frac{x_1+\ldots+x_N}{N}\right|^p \geq \frac{|x_1|^p+\ldots+|x_N|^p}{N}$$

with in both cases equality precisely when $|x_1| = \ldots = |x_N|$.

PROOF. This follows indeed from Theorem 5.9, because we have:

$$(x^p)'' = p(p-1)x^{p-2}$$

Thus x^p is convex for $p > 1$ and concave for $p < 1$, which gives the results. □

Observe that at $p = 2$ we obtain as particular case of the Hölder inequality the Cauchy-Schwarz inequality, or rather something equivalent to it, namely:

$$\left(\frac{x_1+\ldots+x_N}{N}\right)^2 \leq \frac{x_1^2+\ldots+x_N^2}{N}$$

We will be back to this in a moment, when talking scalars products and Hilbert spaces, with some more conceptual proofs for such inequalities.

Getting now to spaces of functions, let us start with something standard, namely:

THEOREM 5.11. *Given two functions $f, g : \mathbb{R} \to \mathbb{C}$ and an exponent $p \geq 1$, we have*

$$\left(\int_{\mathbb{R}} |f+g|^p\right)^{1/p} \leq \left(\int_{\mathbb{R}} |f|^p\right)^{1/p} + \left(\int_{\mathbb{R}} |g|^p\right)^{1/p}$$

called Minkowski inequality. Also, assuming that $p, q \geq 1$ satisfy $1/p + 1/q = 1$, we have

$$\int_{\mathbb{R}} |fg| \leq \left(\int_{\mathbb{R}} |f|^p\right)^{1/p} \left(\int_{\mathbb{R}} |g|^q\right)^{1/q}$$

called Hölder inequality. These inequalities hold as well for ∞ values of the exponents.

PROOF. All this is very standard, the idea being as follows:

(1) As a first observation, at $p = 2$, which is a special exponent, we have $q = 2$ as well, and the Minkowski and Hölder inequalities are as follows:

$$\left(\int_{\mathbb{R}} |f+g|^2\right)^{1/2} \leq \left(\int_{\mathbb{R}} |f|^2\right)^{1/2} + \left(\int_{\mathbb{R}} |g|^2\right)^{1/2}$$

$$\int_{\mathbb{R}} |fg| \leq \left(\int_{\mathbb{R}} |f|^2\right)^{1/2} \left(\int_{\mathbb{R}} |g|^2\right)^{1/2}$$

But the proof of the Hölder inequality, called Cauchy-Schwarz inequality in this case, is something elementary, coming from the fact that $I(t) = \int |f + twg|^2$ with $|w| = 1$ is a positive degree 2 polynomial in $t \in \mathbb{R}$, and so its discriminant must be negative. As for the Minkowski inequality, in the present $p = 2$ case this follows from the Hölder, or Cauchy-Schwarz inequality, by taking squares and simplifying.

(2) In the general case now, where $p \geq 1$ is still finite, but arbitrary, the proofs are a bit more complicated, based on the Jensen inequality that we discussed above. In fact, we already know from there the Minkowski and Hölder inequalities for the sequences of numbers, and the extension to functions is straightforward.

(3) Finally, with the convention that $(\int |f|^p)^{1/p}$ takes as value at $p = \infty$ the essential supremum of f, the Minkowski inequality holds as well at $p = \infty$, trivially:

$$\sup \big|f + g\big| \leq \sup \big|f\big| + \sup \big|g\big|$$

The same goes for the Hölder inequality at $p = \infty, q = 1$, which is simply:

$$\int_{\mathbb{R}} |fg| \leq \sup \big|f\big| \times \int_{\mathbb{R}} |g|$$

And finally, the same goes for the Hölder inequality at $p = 1, q = \infty$. □

As a consequence of the above results, we can formulate:

THEOREM 5.12. *Given an interval $I \subset \mathbb{R}$ and an exponent $p \geq 1$, the following space, with the convention that functions are identified up to equality almost everywhere,*

$$L^p(I) = \left\{ f : I \to \mathbb{C} \Big| \int_I |f(x)|^p dx < \infty \right\}$$

is a vector space, and the following quantity

$$||f||_p = \left(\int_I |f(x)|^p\right)^{1/p}$$

is a norm on it, in the sense that it satisfies the usual conditions for a vector space norm. Moreover, $L^p(I)$ is complete with respect to the distance $d(f, g) = ||f - g||_p$.

PROOF. This basically follows from Theorem 5.11, the idea being as follows:

(1) Again, let us first see what happens at $p = 2$. Here everything is standard from what we have in Theorem 5.11, and with the remark that the space $L^2(I)$ that we obtain is more than just a normed vector space, because we have as well a scalar product, related to the norm by the formula $||f||_2 = \sqrt{< f, f >}$, constructed as follows:

$$< f, g >= \int_I f(x)\,\overline{g(x)}\,dx$$

(2) In the general case now, where $p \geq 1$ is still finite, but arbitrary, the proof is similar, basically coming from the Minkowski inequality from Theorem 5.11.

(3) Finally, the extension at $p = \infty$ is clear too, coming from definitions, and with the various conventions made at the end of the proof of Theorem 5.11. □

5d. Functional analysis

There are many more things that can be said about the above spaces $L^2(I)$. Let us mention, as a key result, that we will not really need in what follows, that the Hölder inequality shows that any continuous linear form $\varphi : L^p(I) \to \mathbb{C}$ must be of the form $\varphi(f) = \int fg$, with $g \in L^q(I)$, where $q \geq 1$ is as usual given by $1/p + 1/q = 1$. Thus, if we denote by $L^p(I)^*$ the space of continuous linear forms $\varphi : L^p(I) \to \mathbb{C}$, we have:

$$L^p(I)^* = L^q(I)$$

Observe in particular that by applying twice this formula, we obtain the following remarkable property, called self-duality, or reflexivity of the Banach space $L^2(I)$:

$$L^p(I)^{**} = L^p(I)$$

For more on all this, we refer to any functional analysis book.

5e. Exercises

Exercises:

EXERCISE 5.13.

EXERCISE 5.14.

EXERCISE 5.15.

EXERCISE 5.16.

EXERCISE 5.17.

EXERCISE 5.18.

Bonus exercise.

CHAPTER 6

Fourier transform

6a. Convolution product

Going ahead now with our study of functions $f : \mathbb{R} \to \mathbb{C}$, let us define an interesting operation on such functions, called convolution, which is useful for many purposes:

DEFINITION 6.1. *The convolution of two functions $f, g : \mathbb{R} \to \mathbb{C}$ is the function*

$$(f * g)(x) = \int_{\mathbb{R}} f(x-y)g(y)dy$$

provided that the function $y \to f(x-y)g(y)$ is indeed integrable, for any x.

There are many reasons for introducing this operation, that we will gradually discover, in what follows. As a basic example, let us take $g = \chi_{[0,1]}$. We have then:

$$(f * g)(x) = \int_0^1 f(x-y)dy$$

Thus, with this choice of g, the operation $f \to f * g$ has some sort of "regularizing effect", that can be useful for many purposes. We will get back to this, later.

Goinh ahead with more theory, let us try to understand when the convolution operation is well-defined. We have here the following basic result:

THEOREM 6.2. *The convolution operation is well-defined on the space*

$$C_c(\mathbb{R}) = \left\{ f \in C(\mathbb{R}) \middle| supp(f) = \text{ compact} \right\}$$

of continuous functions $f : \mathbb{R} \to \mathbb{C}$ having compact support.

PROOF. We have several things to be proved, the idea being as follows:

(1) First we must show that given two functions $f, g \in C_c(\mathbb{R})$, their convolution $f * g$ is well-defined, as a function $f * g : \mathbb{R} \to \mathbb{C}$. But this follows from the following estimate,

where l denotes the length of the compact subsets of $\mathbb{R}$:

$$\begin{aligned}\int_{\mathbb{R}}|f(x-y)g(y)|dy &= \int_{supp(g)}|f(x-y)g(y)|dy\\ &\leq \max(g)\int_{supp(g)}|f(x-y)|dy\\ &\leq \max(g)\cdot l(supp(g))\cdot\max(f)\\ &< \infty\end{aligned}$$

(2) Next, we must show that the function $f*g:\mathbb{R}\to\mathbb{C}$ that we constructed is indeed continuous. But this follows from the following estimate, where K_f is the constant of uniform continuity for the function $f\in C_c(\mathbb{R})$:

$$\begin{aligned}|(f*g)(x+\varepsilon)-(f*g)(x)| &= \left|\int_{\mathbb{R}}f(x+\varepsilon-y)g(y)dy-\int_{\mathbb{R}}f(x-y)g(y)dy\right|\\ &= \left|\int_{\mathbb{R}}\left(f(x+\varepsilon-y)-f(x-y)\right)g(y)dy\right|\\ &\leq \int_{\mathbb{R}}|f(x+\varepsilon-y)-f(x-y)|\cdot|g(y)|dy\\ &\leq K_f\cdot\varepsilon\cdot\int_{\mathbb{R}}|g|\end{aligned}$$

(3) Finally, we must show that the function $f*g\in C(\mathbb{R})$ that we constructed has indeed compact support. For this purpose, our claim is that we have:

$$supp(f+g)\subset supp(f)+supp(g)$$

In order to prove this claim, observe that we have, by definition of $f*g$:

$$\begin{aligned}(f*g)(x) &= \int_{\mathbb{R}}f(x-y)g(y)dy\\ &= \int_{supp(g)}f(x-y)g(y)dy\end{aligned}$$

But this latter quantity being 0 for $x\notin supp(f)+supp(g)$, this gives the result. □

Here are now a few remarkable properties of the convolution operation:

PROPOSITION 6.3. *The following hold, for the functions in $C_c(\mathbb{R})$:*

(1) $f*g=g*f$.
(2) $f*(g*h)=(f*g)*h$.
(3) $f*(\lambda g+\mu h)=\lambda f*g+\mu f*h$.

PROOF. These formulae are all elementary, the idea being as follows:

(1) This follows from the following computation, with $y = x - t$:

$$\begin{aligned}(f * g)(x) &= \int_{\mathbb{R}} f(x-y)g(y)dy \\ &= \int_{\mathbb{R}} f(t)g(x-t)dt \\ &= \int_{\mathbb{R}} g(x-t)f(t)dt \\ &= (g * f)(x)\end{aligned}$$

(2) This is clear from definitions.

(3) Once again, this is clear from definitions. □

In relation with derivatives, we have the following result:

THEOREM 6.4. *Given two functions $f, g \in C_c(\mathbb{R})$, assuming that g is differentiable, then so is $f * g$, with derivative given by the following formula:*

$$(f * g)' = f * g'$$

More generally, given $f, g \in C_c(\mathbb{R})$, and assuming that g is k times differentiable, then so is $f * g$, with k-th derivative given by $(f * g)^{(k)} = f * g^{(k)}$.

PROOF. In what regards the first assertion, with $y = x - t$, then $t = x - y$, we get:

$$\begin{aligned}(f * g)'(x) &= \frac{d}{dx}\int_{\mathbb{R}} f(x-y)g(y)dy \\ &= \frac{d}{dx}\int_{\mathbb{R}} f(t)g(x-t)dt \\ &= \int_{\mathbb{R}} f(t)g'(x-t)dt \\ &= \int_{\mathbb{R}} f(x-y)g'(y)dy \\ &= (f * g')(x)\end{aligned}$$

As for the second assertion, this follows form the first one, by recurrence. □

Finally, getting beyond the compactly supported continous functions, we have the following result, which is of particular theoretical importance:

THEOREM 6.5. *The convolution operation is well-defined on $L^1(\mathbb{R})$, and we have:*

$$||f * g||_1 \leq ||f||_1 ||g||_1$$

Thus, if $f \in L^1(\mathbb{R})$ and $g \in C_c^k(\mathbb{R})$, then $f * g$ is well-defined, and $f * g \in C_c^k(\mathbb{R})$.

PROOF. In what regards the first assertion, this follows from the following computation, involving an intuitive manipulation on the double integrals, called Fubini theorem, that we will use as such here, and that we will fully clarify later on, when talking more in detail about functions of several real variables, and their integrals:

$$\begin{aligned}\int_{\mathbb{R}}|(f*g)(x)|dx &\leq \int_{\mathbb{R}}\int_{\mathbb{R}}|f(x-y)g(y)|dydx\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}}|f(x-y)g(y)|dxdy\\ &= \int_{\mathbb{R}}|f(x)|dx\int_{\mathbb{R}}|g(y)|dy\end{aligned}$$

As for the second assertion, this follows from the first one, and from Theorem 6.4. □

We should mention that it is possible to go beyond the above result, with a suitable extension of the convolution to the space $L^2(\mathbb{R})$ as well. We will be back to this, and to other more specialized questions, at the end of the present chapter.

Summarizing, we have now some good knowledge of the various spaces that the functions $f:\mathbb{R}\to\mathbb{C}$ can form, and we have as well an interesting regularization operation $f\to f*g$ on such functions, that can be used for various purposes.

6b. Fourier transform

We discuss here the construction and main properties of the Fourier transform, which is the main tool in analysis, and even in mathematics in general. We first have:

DEFINITION 6.6. *Given $f\in L^1(\mathbb{R})$, we define a function $\widehat{f}:\mathbb{R}\to\mathbb{C}$ by*

$$\widehat{f}(\xi)=\int_{\mathbb{R}}e^{ix\xi}f(x)dx$$

and call it Fourier transform of f.

As a first observation, even if f is a real function, $\widehat{f}$ is a complex function, which is not necessarily real. Also, $\widehat{f}$ is obviously well-defined, because $f\in L^1(\mathbb{R})$ and $|e^{ix\xi}|=1$. Also, the condition $f\in L^1(\mathbb{R})$ is basically needed for constructing $\widehat{f}$, because:

$$\widehat{f}(0)=\int_{\mathbb{R}}f(x)dx$$

Generally speaking, the Fourier transform is there for helping with various computations, with the above formula $\widehat{f}(0)=\int f$ being something quite illustrating.

Here are some basic properties of the Fourier transform, all providing some good motivations for the study of this transform, adding to the above one:

PROPOSITION 6.7. *The Fourier transform has the following properties:*

(1) *Linearity:* $\widehat{f+g}=\widehat{f}+\widehat{g}$, $\widehat{\lambda f}=\lambda\widehat{f}$.
(2) *Regularity:* $\widehat{f}$ *is continuous and bounded.*
(3) *If* f *is even then* $\widehat{f}$ *is even.*
(4) *If* f *is odd then* $\widehat{f}$ *is odd.*

PROOF. These results are all elementary, as follows:

(1) The additivity formula is clear from definitions, as follows:

$$\begin{aligned}\widehat{f+g}(\xi) &= \int_{\mathbb{R}} e^{ix\xi}(f+g)(x)dx\\ &= \int_{\mathbb{R}} e^{ix\xi}f(x)dx+\int_{\mathbb{R}} e^{ix\xi}f(x)dx\\ &= \widehat{f}(\xi)+\widehat{g}(\xi)\end{aligned}$$

As for the formula $\widehat{\lambda f}=\lambda\widehat{f}$, this is clear as well.

(2) The continuity of $\widehat{f}$ follows indeed from:

$$\begin{aligned}|\widehat{f}(\xi+\varepsilon)-\widehat{f}(\xi)| &\leq \int_{\mathbb{R}}\left|(e^{ix(\xi+\varepsilon)}-e^{ix\xi})f(x)\right|dx\\ &= \int_{\mathbb{R}}\left|e^{ix\xi}(e^{ix\varepsilon}-1)f(x)\right|dx\\ &\leq |e^{ix\varepsilon}-1|\int_{\mathbb{R}}|f|\end{aligned}$$

As for the boundedness of $\widehat{f}$, this is clear as well.

(3) This follows from the following computation, assuming that f is even:

$$\begin{aligned}\widehat{f}(-\xi) &= \int_{\mathbb{R}} e^{-ix\xi}f(x)dx\\ &= \int_{\mathbb{R}} e^{ix\xi}f(-x)dx\\ &= \int_{\mathbb{R}} e^{ix\xi}f(x)dx\\ &= \widehat{f}(\xi)\end{aligned}$$

(4) The proof here is similar to the proof of (3), by changing some signs. □

We will be back to more theory in a moment, but let us explore now the examples. Here are some basic computations of Fourier transforms:

PROPOSITION 6.8. *We have the following Fourier transform formulae,*

$$f = \chi_{[-a,a]} \implies \widehat{f}(\xi) = \frac{2\sin(a\xi)}{\xi}$$

$$f = e^{-ax}\chi_{[0,\infty]}(x) \implies \widehat{f}(\xi) = \frac{1}{a - i\xi}$$

$$f = e^{ax}\chi_{[-\infty,0]}(x) \implies \widehat{f}(\xi) = \frac{1}{a + i\xi}$$

$$f = e^{-a|x|} \implies \widehat{f}(\xi) = \frac{2a}{a^2 + \xi^2}$$

$$f = sgn(x)e^{-a|x|} \implies \widehat{f}(\xi) = \frac{2i\xi}{a^2 + \xi^2}$$

valid for any number $a > 0$.

PROOF. All this follows from some calculus, as follows:

(1) In what regards first formula, assuming $f = \chi_{[-a,a]}$, we have, by using the fact that $\sin(x\xi)$ is an odd function, whose integral vanishes on centered intervals:

$$\begin{aligned}
\widehat{f}(\xi) &= \int_{-a}^{a} e^{ix\xi}dx \\
&= \int_{-a}^{a} \cos(x\xi)dx + i\int_{-a}^{a} \sin(x\xi)dx \\
&= \int_{-a}^{a} \cos(x\xi)dx \\
&= \left[\frac{\sin(x\xi)}{\xi}\right]_{-a}^{a} \\
&= \frac{2\sin(a\xi)}{\xi}
\end{aligned}$$

(2) With $f(x) = e^{-ax}\chi_{[0,\infty]}(x)$, the computation goes as follows:

$$\begin{aligned}
\widehat{f}(\xi) &= \int_{0}^{\infty} e^{ix\xi - ax}dx \\
&= \int_{0}^{\infty} e^{(i\xi - a)x}dx \\
&= \left[\frac{e^{(i\xi - a)x}}{i\xi - a}\right]_{0}^{\infty} \\
&= \frac{1}{a - i\xi}
\end{aligned}$$

(3) Regarding the third formula, this follows from the second one, by using the following fact, generalizing the parity computations from Proposition 6.7:

$$F(x) = f(-x) \implies \widehat{F}(\xi) = \widehat{f}(-\xi)$$

(4) The last 2 formulae follow from what we have, by making sums and differences, and the linearity properties of the Fourier transform, from Proposition 6.7. □

We will see many other examples, in what follows. Getting back now to theory, we have the following result, adding to the various general properties in Proposition 6.7, and providing more motivations for the Fourier transform:

PROPOSITION 6.9. *Given $f, g \in L^1(\mathbb{R})$ we have $\widehat{f}g, f\widehat{g} \in L^1(\mathbb{R})$ and*

$$\int_{\mathbb{R}} f(\xi)\widehat{g}(\xi)d\xi = \int_{\mathbb{R}} \widehat{f}(x)g(x)dx$$

called "exchange of hat" formula.

PROOF. Regarding the fact that we have indeed $\widehat{f}g, f\widehat{g} \in L^1(\mathbb{R})$, this is actually a bit non-trivial, but we will be back to this later. Assuming this, we have:

$$\int_{\mathbb{R}} f(\xi)\widehat{g}(\xi)d\xi = \int_{\mathbb{R}}\int_{\mathbb{R}} f(\xi)e^{ix\xi}g(x)dxd\xi$$

On the other hand, we have as well the following formula:

$$\int_{\mathbb{R}} \widehat{f}(x)g(x)dx = \int_{\mathbb{R}}\int_{\mathbb{R}} e^{ix\xi}f(x)g(\xi)dxd\xi$$

Thus, with $x \leftrightarrow \xi$, we are led to the formula in the statement. □

As an important result now, showing the power of the Fourier transform, this transforms the derivative, which can be a quite complicated operation, into something very simple, namely the multiplication by the variable, the precise result being as follows:

THEOREM 6.10. *Given $f : \mathbb{R} \to \mathbb{C}$ such that $f, f' \in L^1(\mathbb{R})$, we have:*

$$\widehat{f'}(\xi) = -i\xi\widehat{f}(\xi)$$

More generally, assuming $f, f', f'', \ldots, f^{(n)} \in L^1(\mathbb{R})$, we have

$$\widehat{f^{(k)}}(\xi) = (-i\xi)^k\widehat{f}(\xi)$$

for any $k = 1, 2, \ldots, n$.

PROOF. These results follow by doing a partial integration, as follows:

(1) Assuming that $f:\mathbb{R}\to\mathbb{C}$ has compact support, we have indeed:

$$\begin{aligned}\widehat{f'}(\xi) &= \int_{\mathbb{R}} e^{ix\xi}f'(x)dx\\ &= -\int_{\mathbb{R}} i\xi e^{ix\xi}f(x)dx\\ &= -i\xi\int_{\mathbb{R}} e^{ix\xi}f(x)dx\\ &= -i\xi\widehat{f}(\xi)\end{aligned}$$

(2) Regarding the higher derivatives, the formula here follows by recurrence. □

Importantly, we have a converse statement as well, as follows:

THEOREM 6.11. *Assuming that $f\in L^1(\mathbb{R})$ is such that $F(x)=xf(x)$ belongs to $L^1(\mathbb{R})$ too, the function $\widehat{f}$ is differentiable, with derivative given by:*

$$(\widehat{f})'(\xi)=i\widehat{F}(\xi)$$

More generally, if $F_k(x)=x^kf(x)$ belongs to $L^1(\mathbb{R})$, for $k=0,1,\ldots,n$, we have

$$(\widehat{f})^{(k)}(\xi)=i^k\widehat{F_k}(\xi)$$

for any $k=1,2,\ldots,n$.

PROOF. These results are both elementary, as follows:

(1) Regarding the first assertion, the computation here is as follows:

$$\begin{aligned}(\widehat{f})'(\xi) &= \frac{d}{d\xi}\int_{\mathbb{R}} e^{ix\xi}f(x)dx\\ &= \int_{\mathbb{R}} ixe^{ix\xi}f(x)dx\\ &= i\int_{\mathbb{R}} e^{ix\xi}xf(x)dx\\ &= i\widehat{F}(\xi)\end{aligned}$$

(2) As for the second assertion, this follows from the first one, by recurrence. □

As a conclusion to all this, we are on a good way with our theory, and we have:

CONCLUSION 6.12. *Modulo normalization factors, the Fourier transform converts the derivatives into multiplications by the variable, and vice versa.*

And isn't this interesting, because isn't computing derivatives a difficult task. Here is now another useful result, of the same type, this time regarding convolutions:

THEOREM 6.13. *Assuming $f, g \in L^1(\mathbb{R})$, the following happens:*

$$\widehat{f * g} = \widehat{f} \cdot \widehat{g}$$

Moreover, under suitable assumptions, the formula $\widehat{fg} = \widehat{f} * \widehat{g}$ holds too.

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

(1) Regarding the first assertion, this is something elementary, as follows:

$$\begin{aligned}
\widehat{f * g}(\xi) &= \int_{\mathbb{R}} e^{ix\xi}(f * g)(x)dx \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{ix\xi} f(x-y)g(y)dxdy \\
&= \int_{\mathbb{R}} e^{iy\xi}\left(\int e^{i(x-y)\xi} f(x-y)dx\right) g(y)dy \\
&= \int_{\mathbb{R}} e^{iy\xi}\left(\int e^{it\xi} f(t)dt\right) g(y)dy \\
&= \int_{\mathbb{R}} e^{iy\xi}\widehat{f}(\xi)g(y)dy \\
&= \widehat{f}(\xi)\widehat{g}(\xi)
\end{aligned}$$

(2) As for the second assertion, this is something more tricky, and we will be back to this later. In the meantime, here is however some sort of proof, not very honest:

$$\begin{aligned}
(\widehat{f} * \widehat{g})(\xi) &= \int_{\mathbb{R}} \widehat{f}(\xi-\eta)\widehat{g}(\eta)d\eta \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{ix(\xi-\eta)} f(x) e^{iy\eta} g(y)dxdyd\eta \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{ix\eta} e^{i(y-x)\eta} f(x)g(y)dxdyd\eta \\
&= \int_{\mathbb{R}} e^{ix\eta} f(x)g(x)dx \\
&= \widehat{fg}(\eta)
\end{aligned}$$

To be more precise, the point here is that we can pass from the triple to the single integral by arguing that "we must have $x = y$". We will be back to this later. □

As an updated conclusion to all this, we have, modulo a few bugs, to be fixed:

CONCLUSION 6.14. *The Fourier transform converts the derivatives into multiplications by the variable, and convolutions into products, and vice versa.*

We will see applications of this later, after developing some more general theory.

6c. Fourier inversion

Let us develop now more theory for the Fourier transform. We first have:

THEOREM 6.15. *Given $f \in L^1(\mathbb{R})$, its Fourier transform satisfies*

$$\lim_{\xi\to\pm\infty} \widehat{f}(\xi) = 0$$

called Riemann-Lebesgue property of $\widehat{f}$.

PROOF. This is something quite technical, as follows:

(1) Given a function $f : \mathbb{R} \to \mathbb{C}$ and a number $y \in \mathbb{R}$, let us set:

$$f_y(x) = f(x - y)$$

Our claim is then is that if $f \in L^p(\mathbb{R})$, then the following function is uniformly continuous, with respect to the usual p-norm on the right:

$$\mathbb{R} \to L^p(\mathbb{R}) \quad , \quad y \to f_y$$

(2) In order to prove this, fix $\varepsilon > 0$. Since $f \in L^p(\mathbb{R})$, we can find a function of type $g : [-K, K] \to \mathbb{C}$ which is continuous, such that:

$$||f - g||_p < \varepsilon$$

Now since g is uniformly continuous, we can find $\delta \in (0, K)$ such that:

$$|s - t| < \delta \implies |g(s) - g(t)| < (3K)^{-1/p}\varepsilon$$

But this shows that we have the following estimate:

$$\begin{aligned} ||g_s - g_t||_p &= \left(\int_{\mathbb{R}} \big|g(x - s) - g(x - t)\big|^p dx\right)^{1/p} \\ &< \big[(3K)^{-1}\varepsilon^p(2k + \delta)\big]^{1/p} \\ &< \varepsilon \end{aligned}$$

By using now the formula $||f||_p = ||f_s||_p$, which is clear, we obtain:

$$\begin{aligned} ||f_s - f_t||_p &\leq ||f_s - g_s||_p + ||g_s - g_t||_p + ||g_t - f_t||_p \\ &< \varepsilon + \varepsilon + \varepsilon \\ &= 3\varepsilon \end{aligned}$$

But this being true for any $|s - t| < \delta$, we have proved our claim.

(3) Let us prove now the Riemann-Lebesgue property of $\widehat{f}$, as formulated in the statement. By using $e^{\pi i}=-1$, and the change of variables $x\to x-\pi/\xi$, we have:

$$\begin{aligned}\widehat{f}(\xi)&=\int_{\mathbb{R}}e^{ix\xi}f(x)dx\\&=-\int_{\mathbb{R}}e^{ix\xi}e^{\pi i}f(x)dx\\&=-\int_{\mathbb{R}}e^{i\xi(x+\pi/\xi)}f(x)dx\\&=-\int_{\mathbb{R}}e^{ix\xi}f\left(x-\frac{\pi}{\xi}\right)dx\end{aligned}$$

On the other hand, we have as well the following formula:

$$\widehat{f}(\xi)=\int_{\mathbb{R}}e^{ix\xi}f(x)dx$$

Thus by summing, we obtain the following formula:

$$2\widehat{f}(\xi)=\int_{\mathbb{R}}e^{ix\xi}\left(f(x)-f\left(x-\frac{\pi}{\xi}\right)\right)dx$$

But this gives the following estimate, with notations from (1):

$$2|\widehat{f}(\xi)|\leq||f-f_{\pi/\xi}||_1$$

Since by (1) this goes to 0 with $\xi\to\pm\infty$, this gives the result. □

Quite remarkably, and as a main result now regarding Fourier transforms, a function $f:\mathbb{R}\to\mathbb{C}$ can be recovered from its Fourier transform $\widehat{f}:\mathbb{R}\to\mathbb{C}$, as follows:

THEOREM 6.16. *Assuming $f,\widehat{f}\in L^1(\mathbb{R})$, we have*

$$f(x)=\int_{\mathbb{R}}e^{-ix\xi}\widehat{f}(\xi)d\xi$$

almost everywhere, called Fourier inversion formula.

PROOF. This is something quite tricky, due to the fact that a direct attempt by double integration fails. Consider the following function, depending on a parameter $\lambda>0$:

$$\varphi_\lambda(x)=\int_{\mathbb{R}}e^{-ix\xi-\lambda|\xi|}d\xi$$

We have then the following computation:

$$\begin{aligned}(f*\varphi_\lambda)(x) &= \int_{\mathbb{R}} f(x-y)\varphi_\lambda(y)dy\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)e^{-iy\xi-\lambda|\xi|}d\xi dy\\ &= \int_{\mathbb{R}} e^{-\lambda|\xi|}\left(\int_{\mathbb{R}} f(x-y)e^{-iy\xi}dy\right)d\xi\\ &= \int_{\mathbb{R}} e^{-\lambda|\xi|}e^{-ix\xi}\widehat{f}(\xi)d\xi\end{aligned}$$

By letting now $\lambda\to 0$, we obtain from this the following formula:

$$\lim_{\lambda\to 0}(f*\varphi_\lambda)(x)=\int_{\mathbb{R}} e^{-ix\xi}\widehat{f}(\xi)d\xi$$

On the other hand, by using Theorem 6.15 we obtain that, almost everywhere:

$$\lim_{\lambda\to 0}(f*\varphi_\lambda)(x)=f(x)$$

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

As an application, we can now fully prove $\widehat{fg}=\widehat{f}*\widehat{g}$, via Fourier inversion.

6d. Plancherel and more

There are many more things that can be said about Fourier transforms, a key result being the Plancherel formula, allowing us to talk about the Fourier transform over the space $L^2(\mathbb{R})$. Also, we can talk about the Fourier transform over the space $\mathcal{S}$ of functions all whose derivatives are rapidly decreasing, called Schwartz space.

6e. Exercises

Exercises:

Exercise 6.17.

Exercise 6.18.

Exercise 6.19.

Exercise 6.20.

Exercise 6.21.

Exercise 6.22.

Bonus exercise.

CHAPTER 7

Advanced aspects

7a.

7b.

7c.

7d.

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

Wave packets

8a.

8b.

8c.

8d.

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

PDE and waves

Nothing here remains
No future and no past
No one could foresee
The end that came so fast

CHAPTER 9

Basic theory

9a.

9b.

9c.

9d.

9e. Exercises

Exercises:

Exercise 9.1.

Exercise 9.2.

Exercise 9.3.

Exercise 9.4.

Exercise 9.5.

Exercise 9.6.

Bonus exercise.

CHAPTER 10

Back to waves

10a.

10b.

10c.

10d.

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

More mathematics

11a.

11b.

11c.

11d.

11e. Exercises

Exercises:

Exercise 11.1.

Exercise 11.2.

Exercise 11.3.

Exercise 11.4.

Exercise 11.5.

Exercise 11.6.

Bonus exercise.

CHAPTER 12

Waves and heat

12a.

12b.

12c.

12d.

12e. Exercises

Exercises:

Exercise 12.1.

Exercise 12.2.

Exercise 12.3.

Exercise 12.4.

Exercise 12.5.

Exercise 12.6.

Bonus exercise.

Part IV

Generalizations

Farewell the ashtray girl
Forbidden snowflake
Beware this troubled world
Watch out for earthquakes

CHAPTER 13

Group theory

13a.

13b.

13c.

13d.

13e. Exercises

Exercises:

EXERCISE 13.1.

EXERCISE 13.2.

EXERCISE 13.3.

EXERCISE 13.4.

EXERCISE 13.5.

EXERCISE 13.6.

Bonus exercise.

CHAPTER 14

Discrete Fourier

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

Quantum Fourier

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

Exotic waves

16a.

16b.

16c.

16d.

16e. Exercises

Congratulations for having read this book, and no exercises for this final chapter.

Bibliography

[1] V.I. Arnold, Ordinary differential equations, Springer (1973).

[2] V.I. Arnold, Mathematical methods of classical mechanics, Springer (1974).

[3] V.I. Arnold, Lectures on partial differential equations, Springer (1997).

[4] V.I. Arnold, Catastrophe theory, Springer (1974).

[5] V.I. Arnold and B.A. Khesin, Topological methods in hydrodynamics, Springer (1998).

[6] V.I. Arnold, V.V. Kozlov and A.I. Neishtadt, Mathematical aspects of classical and celestial mechanics, Springer (2006).

[7] M.F. Atiyah, K-theory, CRC Press (1964).

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

[9] T. Banica, Calculus and applications (2024).

[10] T. Banica, Principles of operator algebras (2023).

[11] T. Banica, Introduction to modern physics (2024).

[12] G.K. Batchelor, An introduction to fluid dynamics, Cambridge Univ. Press (1967).

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

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

[15] N. Berline, E. Getzler and M. Vergne, Heat kernels and Dirac operators, Springer (2004).

[16] S.J. Blundell and K.M. Blundell, Concepts in thermal physics, Oxford Univ. Press (2006).

[17] S.M. Carroll, Spacetime and geometry, Cambridge Univ. Press (2004).

[18] P.M. Chaikin and T.C. Lubensky, Principles of condensed matter physics, Cambridge Univ. Press (1995).

[19] E.A. Coddington, An introduction to ordinary differential equations, Dover (1961).

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

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

[22] J.B. Conway, A course in functional analysis, Springer (1985).

[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. Deift, Orthogonal polynomials and random matrices: a Riemann-Hilbert approach, AMS (1999).

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

[27] M.P. do Carmo, Differential geometry of curves and surfaces, Dover (1976).

[28] M.P. do Carmo, Riemannian geometry, Birkhäuser (1992).

[29] S. Dodelson, Modern cosmology, Academic Press (2003).

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

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

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

[33] L.D. Faddeev and L. A. Takhtajan, Hamiltonian methods in the theory of solitons, Springer (2007).

[34] S.J. Farlow, Partial differential equations for scientists and engineers, Dover (1982).

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

[36] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics, Caltech (1963).

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

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

[39] A.P. French, Special relativity, Taylor and Francis (1968).

[40] D. Gilbarg and N.S. Trudinger, Elliptic partial differential equations of second order, Springer (2001).

[41] H. Goldstein, C. Safko and J. Poole, Classical mechanics, Addison-Wesley (1980).

[42] M.B. Green, J.H. Schwarz and E. Witten, Superstring theory, Cambridge Univ. Press (2012).

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

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

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

[46] R.W. Hamming, Numerical methods for scientists and engineers, Dover (1962).

[47] J. Harris, Algebraic geometry, Springer (1992).

[48] W.A. Harrison, Solid state theory, Dover (1970).

[49] W.A. Harrison, Electronic structure and the properties of solids, Dover (1980).

[50] H. Hofer and E. Zehnder, Symplectic invariants and Hamiltonian dynamics, Birkhäuser (1994).

[51] L. Hörmander, The analysis of linear partial differential operators, Springer (1983).

[52] R.A. Horn and C.R. Johnson, Matrix analysis, Cambridge Univ. Press (1985).

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

[54] V.F.R. Jones, Subfactors and knots, AMS (1991).

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

[56] T. Kato, Perturbation theory for linear operators, Springer (1995).

[57] T. Kibble and F.H. Berkshire, Classical mechanics, Imperial College Press (1966).

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

[59] L.D. Landau and E.M. Lifshitz, Course of theoretical physics, Pergamon Press (1960).

[60] G. Landi, An introduction to noncommutative spaces and their geometry, Springer (1997).

[61] S. Lang, Algebra, Addison-Wesley (1993).

[62] P. Lax, Linear algebra and its applications, Wiley (2007).

[63] P. Lax, Functional analysis, Wiley (2002).

[64] P. Lax and R.S. Phillips, Scattering theory, Academic Press (1967).

[65] P. Lax and M.S. Terrell, Calculus with applications, Springer (2013).

[66] P. Lax and M.S. Terrell, Multivariable calculus with applications, Springer (2018).

[67] J.M. Lee, Introduction to topological manifolds, Springer (2011).

[68] J.M. Lee, Introduction to smooth manifolds, Springer (2012).

[69] J.M. Lee, Introduction to Riemannian manifolds, Springer (2019).

[70] D. McDuff and D. Salamon, Introduction to symplectic topology, Oxford Univ. Press (2017).

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

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

[73] R.K. Pathria and and P.D. Beale, Statistical mechanics, Elsevier (1972).

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

[75] P. Petersen, Riemannian geometry, Springer (2006).

[76] W. Rudin, Principles of mathematical analysis, McGraw-Hill (1964).

[77] W. Rudin, Real and complex analysis, McGraw-Hill (1966).

[78] W. Rudin, Functional analysis, McGraw-Hill (1973).

[79] W. Rudin, Fourier analysis on groups, Dover (1974).

[80] B. Ryden, Introduction to cosmology, Cambridge Univ. Press (2002).

[81] D.V. Schroeder, An introduction to thermal physics, Oxford Univ. Press (1999).

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

[83] I.R. Shafarevich, Basic algebraic geometry, Springer (1974).

[84] R. Shankar, Quantum field theory and condensed matter: an introduction, Cambridge Univ. Press (2017).

[85] A.M. Steane, Thermodynamics, Oxford Univ. Press (2016).

[86] S. Sternberg, Dynamical systems, Dover (2010).

[87] J.R. Taylor, Classical mechanics, Univ. Science Books (2003).

[88] M. Tenenbaum and H. Pollard, Ordinary differential equations, Dover (1963).

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

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

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

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

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

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

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

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

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

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

[99] D.H. White and H.E. White, Physics and music: the science of musical sound, Dover (1980).

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