The magic of extreme physics

Teo Banica

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

2010 *Mathematics Subject Classification.* 85A40

Key words and phrases. High energy, Low energy

Abstract. This is an introduction to various extreme aspects of physics. We first discuss selected aspects of high energy physics, chosen as extreme as possible, which are quite commonplace in our universe. Then we go on a complementary discussion, of low energy physics, again by insisting on the extremes, which are quite commonplace too, in other parts of our universe. With this done, extremes of our universe explored, we declare ourselves still unsatisfied, and we go back in time, and also forward into the future, with the aim of looking for even more extreme phenomena and physics.

Preface

Do we live in normal conditions, here on planet Earth? Certainly not, from a cosmic perspective, and with this being a real issue for us, humans, constantly hampering our attempts to understand mathematics, physics, and other sciences.

A well-known example for this is the simple fact, that you surely know since childhood, that feathers fall slower than rocks. That looks as rock-solid as a scientific discovery can look, and go develop a theory of gravity based on this, certainly no way. Fortunately, mankind eventually found a solution to this, by some kind of miracle, and with Newton and the others developing, on their way, calculus too. Responsible for this is of course the Earth atmosphere, which is something quite unnatural, physically speaking.

Along the same lines, another annoying feature of our Earth life is that the way matter is organized here, into atoms and molecules as we know them, does not really allow considerable, naturally occurring electric charges. So, go develop electromagnetism, which by the way happens to be the number 1 force here on Earth, just based on some bizarre experiments that can be made, involving friction, or sightings of lightning, with you being completely clueless on what happens there, and so on. Again, as with gravity, we can feel here that our planet Earth is not the right place, for doing physics.

And there is more. Getting now to the point, really exciting action, appearing in almost every corner of our universe, one extremely annoying feature of Earth, for us physicists, is the lack of naturally occurring atomic or thermonuclear explosions, that we would love to observe, in order to get an idea on the various forces at work. And so again, as before with gravity and electromagnetism, looks like we are in the wrong place.

But wait, in fact, thinking well, what has been said above is just the tip of the iceberg. Think temperature, distance, and many other things, that can widely vary in our universe, and are annoyingly fixed or almost, here on Earth. What a boring planet, do we have. And add to this the time factor, with the present time being maybe something quite boring too, compared to what happened in the past, or what will happen in the future, and you get the picture: we humans are here in the wrong place, at the wrong time.

The present book will be an introduction to this, various extreme aspects of physics, presented with love and interest, and somewhat from a cosmic perspective, instead of the usual fear. We will first discuss selected aspects of high energy physics, chosen as extreme as possible, which are quite commonplace in our universe, and with of course due admiration for nuclear fission and fusion. Then we will go on a complementary discussion, of low energy physics, again by insisting on the extremes, which are quite commonplace too, in other parts of our universe. With this done, extremes of our universe explored, we will declare ourselves still unsatisfied, and we go back in time, and also forward into the future, with the aim of looking for even more extreme phenomena and physics.

All in all, this will be some sort of a cosmology book, that I hope you will enjoy. Let me also mention that our planet Earth will be not forgotten, we will eventually talk about it too, at the end, with the idea that integrating over the boundary of physics, all extreme things, should give you what happens in the middle, that is, where nothing happens.

Many thanks to my cats, for their teachings on high and low energy, that no theoretical physics learning can replace. In fact, we humans are some sort of mice, in this world, and going beyond our modest mice status will be a main theme of discussion, in this book.

Cergy, January 2025
Teo Banica

Contents

Part I

High energy

I got gloss on my lips, a man on my hips
Got me tighter in my Dereon jeans
Acting up, drink in my cup
I can care less what you think

CHAPTER 1

Light, relativity

1a. Speed addition

Welcome to extreme physics. We will be talking in this book about all sorts of exciting physics, and its mathematical interpretations, and the various speculations that can be made. Before that, however, we need to talk about simpler things, namely space and time. And here, we have the following philosophical question, to start with:

Question 1.1. *What is our usual space, the one that we live in?*

And here, we have plenty of answers. Normally that is $\mathbb{R}^3$, where our three coordinates (x, y, z) live. Or perhaps $\mathbb{R}^4$, with one dimension for time t added. But then, for being mathematical and relaxed, why not $\mathbb{R}^N$, extra dimensions can only help for lots of tricks. But then, why not saying $\mathbb{C}^N$, first because $\mathbb{R}^4 \simeq \mathbb{C}^2$ is something very useful, and then also because complex numbers in general can only help, again with a lot of tricks.

However, the story is not over here. Algebraic geometry tells us that being projective, and compact, is the way, so the space becomes $P^{N-1}_{\mathbb{R}}$, or $P^{N-1}_{\mathbb{C}}$. On the other hand, quantum mechanics naturally lives over $\mathbb{C}^\infty$. But the same quantum mechanics, when examined at very small scales, quarks and below, tends to become "confined", somehow of projective nature too, while remaining complex, and infinite dimensional.

And there is more. Again quantum mechanics, but thermodynamics too, tells us that everything is "quantized", and so discrete, with our usual $\mathbb{R}^N$ and its various versions above appearing as thermodynamic limits of this, what happens in the discrete setting, at extremely small scales, somewhere between the quarks scale, and the Planck scale.

Not very easy to navigate through all this, hope you agree with me. Getting a bit organized now, let us first try to answer a modest version of Question 1.1, as follows:

Question 1.2. *What is the usual space for classical mechanics?*

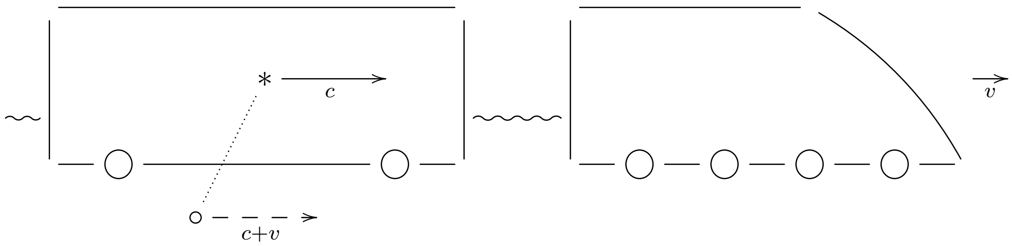
And you would probably say here, of course $\mathbb{R}^3$, as learned from Newton, and many others. And so would I, but Einstein discovered a bug with all this.

Indeed, based on experiments by Fizeau, then Michelson-Morley and others, and some physics by Maxwell and Lorentz too, Einstein came upon the following principles:

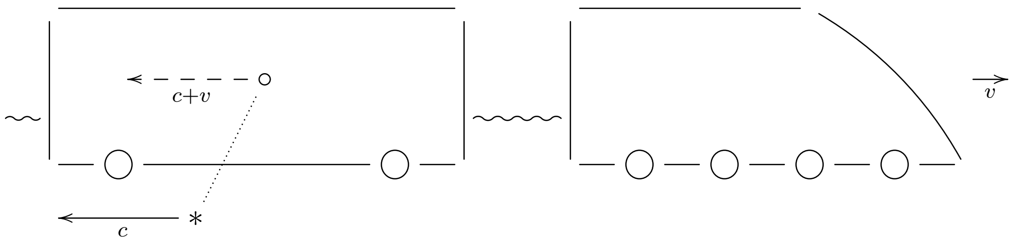
Fact 1.3 (Einstein principles). *The following happen:*

1. *Light travels in vacuum at a finite speed, $c < \infty$.*
2. *This speed c is the same for all inertial observers.*
3. *In non-vacuum, the light speed is lower, $v < c$.*
4. *Nothing can travel faster than light, $v \not> c$.*

The point now is that, obviously, something is wrong here. Indeed, assuming for instance that we have a train, running in vacuum at speed $v > 0$, and someone on board lights a flashlight $*$ towards the locomotive, then an observer $\circ$ on the ground will see the light travelling at speed $c + v > c$, which is a contradiction:



Equivalently, with the same train running, in vacuum at speed $v > 0$, if the observer on the ground lights a flashlight $*$ towards the back of the train, then viewed from the train, that light will travel at speed $c + v > c$, which is a contradiction again:



Summarizing, Fact 1.3 implies $c + v = c$, so contradicts classical mechanics, which therefore needs a fix. By dividing all speeds by c, as to have $c = 1$, and by restricting the attention to the 1D case, to start with, we are led to the following puzzle:

Puzzle 1.4. *How to define speed addition on the space of* 1D *speeds, which is*

$$I = [-1, 1]$$

with our $c = 1$ convention, as to have $1 + c = 1$, as required by physics?

In view of our geometric knowledge so far, a natural idea here would be that of wrapping $[-1, 1]$ into a circle, and then stereographically projecting on $\mathbb{R}$. Indeed, we can then "import" to $[-, 1, 1]$ the usual addition on $\mathbb{R}$, via the inverse of this map.

So, let us see where all this leads us. First, the formula of our map is as follows:

Proposition 1.5. *The map wrapping $[-1,1]$ into the unit circle, and then stereographically projecting on $\mathbb{R}$ is given by the formula*

$$\varphi(u)=\tan\left(\frac{\pi u}{2}\right)$$

with the convention that our wrapping is the most straightforward one, making correspond $\pm 1\to i$, with negatives on the left, and positives on the right.

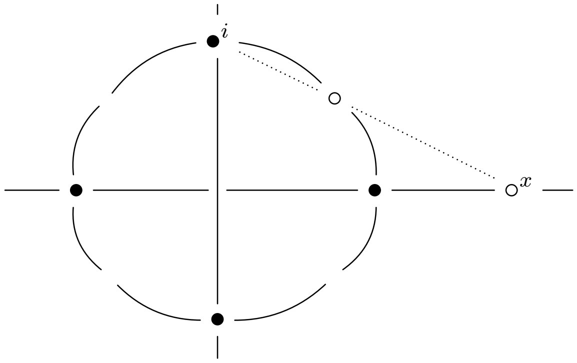
Proof. Regarding the wrapping, as indicated, this is given by:

$$u\to e^{it}\quad,\quad t=\pi u-\frac{\pi}{2}$$

Indeed, this correspondence wraps $[-1,1]$ as above, the basic instances of our correspondence being as follows, and with everything being fine modulo 2π:

$$-1\to\frac{\pi}{2}\quad,\quad -\frac{1}{2}\to-\pi\quad,\quad 0\to-\frac{\pi}{2}\quad,\quad \frac{1}{2}\to 0\quad,\quad 1\to\frac{\pi}{2}$$

Regarding now the stereographic projection, the picture here is as follows:



Thus, by Thales, the formula of the stereographic projection is as follows:

$$\frac{\cos t}{x}=\frac{1-\sin t}{1}\implies x=\frac{\cos t}{1-\sin t}$$

Now if we compose our wrapping operation above with the stereographic projection, what we get is, via the above Thales formula, and some trigonometry:

$$\begin{aligned}
x &= \frac{\cos t}{1-\sin t}\\
&= \frac{\cos\left(\pi u-\frac{\pi}{2}\right)}{1-\sin\left(\pi u-\frac{\pi}{2}\right)}\\
&= \frac{\cos\left(\frac{\pi}{2}-\pi u\right)}{1+\sin\left(\frac{\pi}{2}-\pi u\right)}\\
&= \frac{\sin(\pi u)}{1+\cos(\pi u)}\\
&= \frac{2\sin\left(\frac{\pi u}{2}\right)\cos\left(\frac{\pi u}{2}\right)}{2\cos^2\left(\frac{\pi u}{2}\right)}\\
&= \tan\left(\frac{\pi u}{2}\right)
\end{aligned}$$

Thus, we are led to the conclusion in the statement. $\square$

The above result is very nice, but when it comes to physics, things do not work, for instance because of the wrong slope of the function $\varphi(u)=\tan\left(\frac{\pi u}{2}\right)$ at the origin, which makes our summing on $[-1,1]$ not compatible with the Galileo addition, at low speeds.

So, what to do? Obviously, trash Proposition 1.5, and start all over again. Getting back now to Puzzle 1.4, this has in fact a simpler solution, based this time on algebra, and which in addition is the good, physically correct solution, as follows:

THEOREM 1.6. *If we sum the speeds according to the Einstein formula*

$$u+_e v=\frac{u+v}{1+uv}$$

then the Galileo formula still holds, approximately, for low speeds

$$u+_e v\simeq u+v$$

and if we have $u=1$ or $v=1$, the resulting sum is $u+_e v=1$.

PROOF. All this is self-explanatory, and clear from definitions, and with the Einstein formula of $u+_e v$ itself being just an obvious solution to Puzzle 1.4, provided that, importantly, we know 0 geometry, and rely on very basic algebra only. $\square$

So, very nice, problem solved, at least in 1D. But, shall we give up with geometry, and the stereographic projection? Certainly not, let us try to recycle that material. In order to do this, let us recall that the usual trigonometric functions are given by:

$$\sin x=\frac{e^{ix}-e^{-ix}}{2i}\quad,\quad\cos x=\frac{e^{ix}+e^{-ix}}{2}\quad,\quad\tan x=\frac{e^{ix}-e^{-ix}}{i(e^{ix}+e^{-ix})}$$

The point now is that, and you might know this from calculus, the above functions have some natural "hyperbolic" or "imaginary" analogues, constructed as follows:

$$\sinh x=\frac{e^{x}-e^{-x}}{2}\quad,\quad\cosh x=\frac{e^{x}+e^{-x}}{2}\quad,\quad\tanh x=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$

But the function on the right, tanh, starts reminding the formula of Einstein addition, from Theorem 1.6. So, we have our idea, and we are led to the following result:

Theorem 1.7. *The Einstein speed summation in 1D is given by*

$$\tanh x+_e\tanh y=\tanh(x+y)$$

with $\tanh:[-\infty,\infty]\to[-1,1]$ *being the hyperbolic tangent function.*

Proof. This follows by putting together our various formulae above, but it is perhaps better, for clarity, to prove this directly. Our claim is that we have:

$$\tanh(x+y)=\frac{\tanh x+\tanh y}{1+\tanh x\tanh y}$$

But this can be checked via direct computation, from the definitions, as follows:

$$\begin{aligned}
&\frac{\tanh x+\tanh y}{1+\tanh x\tanh y}\\
=&\left(\frac{e^x-e^{-x}}{e^x+e^{-x}}+\frac{e^y-e^{-y}}{e^y+e^{-y}}\right)\Big/\left(1+\frac{e^x-e^{-x}}{e^x+e^{-x}}\cdot\frac{e^y-e^{-y}}{e^y+e^{-y}}\right)\\
=&\frac{(e^x-e^{-x})(e^y+e^{-y})+(e^x+e^{-x})(e^y-e^{-y})}{(e^x+e^{-x})(e^y+e^{-y})+(e^x-e^{-x})(e^y+e^{-y})}\\
=&\frac{2(e^{x+y}-e^{-x-y})}{2(e^{x+y}+e^{-x-y})}\\
=&\tanh(x+y)
\end{aligned}$$

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

Very nice all this, hope you agree. As a conclusion, passing from the Riemann stereographic projection sum to the Einstein summation basically amounts in replacing:

$$\tan\to\tanh$$

Let us formulate as well this finding more philosophically, as follows:

Conclusion 1.8. *The Einstein speed summation in 1D is the imaginary analogue of the summation on* $[-1,1]$ *obtained via Riemann's stereographic projection.*

Very nice all this, and time to formulate a more precise version of Question 1.2:

QUESTION 1.9. *What is the Einstein speed summation formula in 3D? And, what does this tell us about our usual spacetime $\mathbb{R}^4$, how does this exactly get curved?*

And we will stop with the philosophy here, we have a very good and concrete question now, and time to get to work. Let us attempt to construct $u +_e v$ in arbitrary dimensions, just by using our common sense and intuition. When the vectors $u, v \in \mathbb{R}^N$ are proportional, we are basically in 1D, and so our addition formula must satisfy:

$$u \sim v \implies u +_e v = \frac{u+v}{1+ < u, v >}$$

However, the formula on the right will not work as such in general, for arbitrary speeds $u, v \in \mathbb{R}^N$, and this because we have, as main requirement for our operation, in analogy with the $1 +_e v = 1$ formula from 1D, the following condition:

$$||u|| = 1 \implies u +_e v = u$$

Equivalently, in analogy with $u +_e 1 = 1$ from 1D, we would like to have:

$$||v|| = 1 \implies u +_e v = v$$

Summarizing, our $u \sim v$ formula above is not bad, as a start, but we must add a correction term to it, for the above requirements to be satisfied, and of course with the correction term vanishing when $u \sim v$. So, we are led to a math puzzle:

PUZZLE 1.10. *What vanishes when $u \sim v$, and then how to correctly define*

$$u +_e v = \frac{u + v + \gamma_{uv}}{1+ < u, v >}$$

as for the correction term γ_{uv} to vanish when $u \sim v$?

But the solution to the first question is well-known in 3D. Indeed, here we can use the vector product $u \times v$, that we met before, which notoriously satisfies:

$$u \sim v \implies u \times v = 0$$

Thus, our correction term γ_{uv} must be something containing $w = u \times v$, which vanishes when this vector w vanishes, and in addition arranged such that $||u|| = 1$ produces a simplification, with $u +_e v = u$ as end result, and with $||v|| = 1$ producing a simplification too, with $u +_e v = v$ as end result. Thus, our vector calculus puzzle becomes:

PUZZLE 1.11. *How to correctly define the Einstein summation in 3 dimensions,*

$$u +_e v = \frac{u + v + \gamma_{uvw}}{1+ < u, v >}$$

with $w = u \times v$, in such a way as for the correction term γ_{uvw} to satisfy

$$w = 0 \implies \gamma_{uvw} = 0$$

and also such that $||u|| = 1 \implies u +_e v = u$, and $||v|| \implies u +_e v = v$?

In order to solve this latter puzzle, the first observation is that $\gamma_{uvw} = w$ will not do, and this for several reasons. First, this vector points in the wrong direction, orthogonal to the plane spanned by u, v, and we certainly don't want to leave this plane, with our correction. Also, as a technical remark to be put on top of this, the choice $\gamma_{uvw} = w$ will not bring any simplifications, as required above, in the cases $||u|| = 1$ or $||v|| = 1$. Thus, certainly wrong choice, and we must invent something more complicated.

Moving ahead now, as obvious task, we must "transport" the vector w to the plane spanned by u, v. But this is simplest done by taking the vector product with any vector in this plane, and so as a reasonable candidate for our correction term, we have:

$$\gamma_{uvw} = (\alpha u + \beta v) \times w$$

Here $\alpha, \beta \in \mathbb{R}$ are some scalars to be determined, but let us take a break, and leave the computations for later. We did some good work, time to update our puzzle:

PUZZLE 1.12. *How to define the Einstein summation in 3 dimensions,*

$$u +_e v = \frac{u + v + \gamma_{uvw}}{1 + < u, v >}$$

with the correction term being of the following form, with $w = u \times v$, and $\alpha, \beta \in \mathbb{R}$,

$$\gamma_{uvw} = (\alpha u + \beta v) \times w$$

in such a way as to have $||u|| = 1 \implies u +_e v = u$, and $||v|| \implies u +_e v = v$?

In order to investigate what happens when $||u|| = 1$ or $||v|| = 1$, we must compute the vector products $u \times w$ and $v \times w$. So, pausing now our study for consulting the vector calculus database, and then coming back, here is the formula that we need:

$$u \times (u \times v) = < u, v > u - < u, u > v$$

As for the formula of $v \times w$, that I forgot to record, we can recover it from the one above of $u \times w$, by using the basic properties of the vector products, as follows:

$$\begin{aligned}
v \times (u \times v) &= -v \times (v \times u) \\
&= -(< v, u > v - < v, v > u) \\
&= < v, v > u - < u, v > v
\end{aligned}$$

With these formulae in hand, we can now compute the correction term, with the result here, that we will need several times in what comes next, being as follows:

PROPOSITION 1.13. *The correction term $\gamma_{uvw} = (\alpha u + \beta v) \times w$ is given by*

$$\gamma_{uvw} = (\alpha < u, v > + \beta < v, v >)u - (\alpha < u, u > + \beta < u, v >)v$$

for any values of the scalars $\alpha, \beta \in \mathbb{R}$.

PROOF. According to our vector product formulae above, we have:

$$\begin{aligned}\gamma_{uvw} &= (\alpha u+\beta v)\times w\\ &= \alpha(<u,v>u-<u,u>v)+\beta(<v,v>u-<u,v>v)\\ &= (\alpha<u,v>+\beta<v,v>)u-(\alpha<u,u>+\beta<u,v>)v\end{aligned}$$

Thus, we are led to the conclusion in the statement. $\square$

Time now to get into the real thing, see what happens when $||u||=1$ and $||v||=1$, if we can get indeed $u+_e v=u$ and $u+_e v=v$. It is convenient here to do some reverse engineering. Regarding the first desired formula, namely $u+_e v=u$, we have:

$$\begin{aligned}u+_e v=u &\iff u+v+\gamma_{uvw}=(1+<u,v>)u\\ &\iff \gamma_{uvw}=<u,v>u-v\\ &\iff \alpha=1,\ \beta=0,\ ||u||=1\end{aligned}$$

Thus, with the parameter choice $\alpha=1,\beta=0$, we will have, as desired:

$$||u||=1\implies u+_e v=u$$

In what regards now the second desired formula, namely $u+_e v=v$, here the computation is almost identical, save for a sign switch, which after some thinking comes from our choice $w=u\times v$ instead of $w=v\times u$, clearly favoring u, as follows:

$$\begin{aligned}u+_e v=v &\iff u+v+\gamma_{uvw}=(1+<u,v>)v\\ &\iff \gamma_{uvw}=-u+<u,v>v\\ &\iff \alpha=0,\ \beta=-1,\ ||v||=1\end{aligned}$$

Thus, with the parameter choice $\alpha=0,\beta=-1$, we will have, as desired:

$$||v||=1\implies u+_e v=v$$

All this is mixed news, because we managed to solve both our problems, at $||u||=1$ and at $||v||=1$, but our solutions are different. So, time to breathe, decide that we did enough interesting work for the day, and formulate our conclusion as follows:

PROPOSITION 1.14. *When defining the Einstein speed summation in* 3D *as*

$$u+_e v=\frac{u+v+u\times(u\times v)}{1+<u,v>}$$

in $c=1$ units, the following happen:

1. *When $u\sim v$, we recover the previous* 1D *formula.*
2. *When $||u||=1$, speed of light, we have $u+_e v=u$.*
3. *However, $||v||=1$ does not imply $u+_e v=v$.*
4. *Also, the formula $u+_e v=v+_e u$ fails.*

PROOF. Here (1) and (2) follow from the above discussion, with the following choice for the correction term, by favoring the $||u||=1$ problem over the $||v||=1$ one:

$$\gamma_{uvw}=u\times w$$

In fact, with this choice made, the computation is very simple, as follows:

$$\begin{aligned}||u||=1\quad&\implies\quad\gamma_{uvw}=<u,v>u-v\\&\implies\quad u+v+\gamma_{uvw}=u+<u,v>u\\&\implies\quad\frac{u+v+\gamma_{uvw}}{1+<u,v>}=u\end{aligned}$$

As for (3) and (4), these are also clear from the above discussion, coming from the obvious lack of symmetry of our summation formula. $\square$

Looking now at Proposition 1.14 from an abstract, mathematical perspective, there are still many things missing from there, which can be summarized as follows:

QUESTION 1.15. *Can we fine-tune the Einstein speed summation in* 3D *into*

$$u+_e v=\frac{u+v+\lambda\cdot u\times(u\times v)}{1+<u,v>}$$

with $\lambda\in\mathbb{R}$, chosen such that $||u||=1\implies\lambda=1$, as to have:

1. $||u||,||v||<1\implies||u+_e v||<1$.
2. $||v||=1\implies||u+_e v||=1$.

All this is quite tricky, and deserves some explanations. First, if we add a scalar $\lambda\in\mathbb{R}$ into our formula, as above, we will still have, exactly as before:

$$u\sim v\implies u+_e v=\frac{1+uv}{1+<u,v>}$$

On the other hand, we already know from our previous computations, those preceding Proposition 1.14, that if we ask for $\lambda\in\mathbb{R}$ to be a plain constant, not depending on u,v, then $\lambda=1$ is the only good choice, making the following formula happen:

$$||u||=1\implies u+_e v=u$$

But, and here comes our point, $\lambda=1$ is not an ideal choice either, because it would be nice to have the properties (1,2) in the statement, and these properties have no reason to be valid for $\lambda=1$, as you can check for instance by yourself by doing some computations. Thus, the solution to our problem most likely involves a scalar $\lambda\in\mathbb{R}$ depending on u,v, and satisfying the following condition, as to still have $||u||=1\implies u+_e v=u$:

$$||u||=1\implies\lambda=1$$

Obviously, as simplest answer, λ must be some well-chosen function of $||u||$, or rather of $||u||^2$, because it is always better to use square norms, when possible. But then, with this idea in mind, after a few computations we are led to the following solution:

$$\lambda=\frac{1}{1+\sqrt{1-||u||^2}}$$

Summarizing, final correction done, and with this being the end of mathematics, we did a nice job, and we can now formulate our findings as a theorem, as follows:

Theorem 1.16. *When defining the Einstein speed summation in* 3D *as*

$$u+_e v=\frac{1}{1+<u,v>}\left(u+v+\frac{u\times(u\times v)}{1+\sqrt{1-||u||^2}}\right)$$

in $c=1$ units, the following happen:

1. *When $u\sim v$, we recover the previous* 1D *formula.*
2. *We have $||u||,||v||<1\implies ||u+_e v||<1$.*
3. *When $||u||=1$, we have $u+_e v=u$.*
4. *When $||v||=1$, we have $||u+_e v||=1$.*
5. *However, $||v||=1$ does not imply $u+_e v=v$.*
6. *Also, the formula $u+_e v=v+_e u$ fails.*

Proof. This follows from the above discussion, as follows:

(1) This is something that we know from Proposition 1.14.

(2) In order to simplify notation, let us set $\delta=\sqrt{1-||u||^2}$, which is the inverse of the quantity $\gamma=1/\sqrt{1-||u||^2}$. With this convention, we have:

$$\begin{aligned}u+_e v&=\frac{1}{1+<u,v>}\left(u+v+\frac{<u,v>u-||u||^2v}{1+\delta}\right)\\&=\frac{(1+\delta+<u,v>)u+(1+\delta-||u||^2)v}{(1+<u,v>)(1+\delta)}\end{aligned}$$

Taking now the squared norm and computing gives the following formula:

$$||u+_e v||^2=\frac{(1+\delta)^2||u+v||^2+(||u||^2-2(1+\delta))(||u||^2||v||^2-<u,v>^2)}{(1+<u,v>)^2(1+\delta)^2}$$

But this formula can be further processed by using $\delta=\sqrt{1-||u||^2}$, and by navigating through the various quantities which appear, we obtain, as a final product:

$$||u+_e v||^2=\frac{||u+v||^2-||u||^2||v||^2+<u,v>^2}{(1+<u,v>)^2}$$

But this type of formula is exactly what we need, for what we want to do. Indeed, by assuming $||u||,||v||<1$, we have the following estimate:

$$\begin{aligned} ||u+_e v||^2<1 &\iff ||u+v||^2-||u||^2||v||^2+<u,v>^2<(1+<u,v>)^2\\ &\iff ||u+v||^2-||u||^2||v||^2<1+2<u,v>\\ &\iff ||u||^2+||v||^2-||u||^2||v||^2<1\\ &\iff (1-||u||^2)(1-||v||^2)>0 \end{aligned}$$

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

(3) This is something that we know from Proposition 1.14.

(4) This comes from the squared norm formula established in the proof of (2) above, because when assuming $||v||=1$, we obtain:

$$\begin{aligned} ||u+_e v||^2 &= \frac{||u+v||^2-||u||^2+<u,v>^2}{(1+<u,v>)^2}\\ &= \frac{||u||^2+1+2<u,v>-||u||^2+<u,v>^2}{(1+<u,v>)^2}\\ &= \frac{1+2<u,v>+<u,v>^2}{(1+<u,v>)^2}\\ &= 1 \end{aligned}$$

(5) This is clear, from the obvious lack of symmetry of our formula.

(6) This is again clear, from the obvious lack of symmetry of our formula. □

That was nice, all this mathematics, and hope you're still with me. And good news, the formula in Theorem 1.16 is the good one, confirmed by experimental physics.

1b. Relativity theory

Time now to draw some concrete conclusions, from the above speed computations. Since speed $v=d/t$ is distance over time, we must fine-tune distance d, or time t, or both. Let us first discuss, following as usual Einstein, what happens to time t. Here the result, which might seem quite surprising, at a first glance, is as follows:

THEOREM 1.17. *Relativistic time is subject to Lorentz dilation*

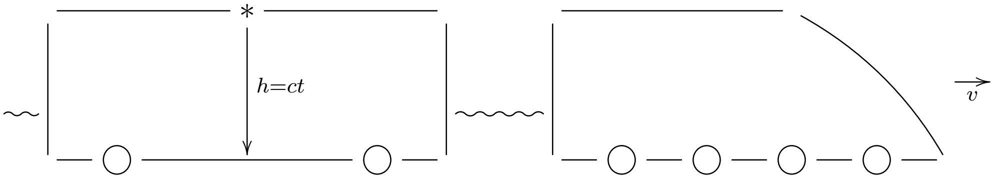
$$t\to\gamma t$$

where the number $\gamma\geq 1$, called Lorentz factor, is given by the formula

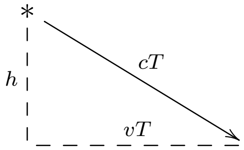
$$\gamma=\frac{1}{\sqrt{1-v^2/c^2}}$$

with v being the moving speed, at which time is measured.

PROOF. Assume indeed that we have a train, moving to the right with speed v, through vacuum. In order to compute the height h of the train, the passenger onboard switches on the ceiling light bulb, measures the time t that the light needs to hit the floor, by travelling at speed c, and concludes that the train height is $h = ct$:



On the other hand, an observer on the ground will see here something different, namely a right triangle, with on the vertical the height of the train h, on the horizontal the distance vT that the train has travelled, and on the hypotenuse the distance cT that light has travelled, with T being the duration of the event, according to his watch:



Now by Pythagoras applied to this triangle, we have:

$$h^2 + (vT)^2 = (cT)^2$$

Thus, the observer on the ground will reach to the following formula for h:

$$h = \sqrt{c^2 - v^2} \cdot T$$

But h must be the same for both observers, so we have the following formula:

$$\sqrt{c^2 - v^2} \cdot T = ct$$

It follows that the two times t and T are indeed not equal, and are related by:

$$T = \frac{ct}{\sqrt{c^2 - v^2}} = \frac{t}{\sqrt{1 - v^2/c^2}} = \gamma t$$

Thus, we are led to the formula in the statement. $\square$

Let us discuss now what happens to length. Intuitively, since speed is distance/time, and since time gets dilated, we can somehow expect distance to get dilated too.

However, and a bit surprisingly, this is wrong, and after due thinking and computations, what we have is in fact the following result:

Theorem 1.18. *Relativistic length is subject to Lorentz contraction*

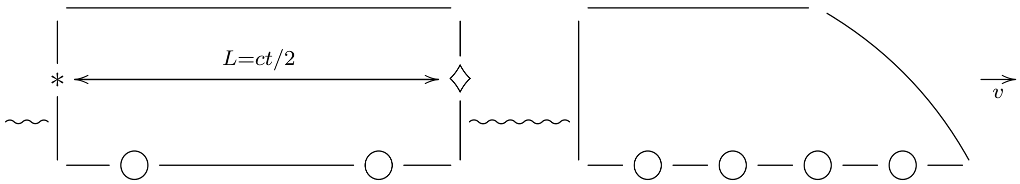
$$L \to L/\gamma$$

where the number $\gamma \geq 1$, called Lorentz factor, is given by the usual formula

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

with v being the moving speed, at which length is measured.

Proof. As before in the proof of Theorem 1.17, meaning in the same train travelling at speed v, in vacuum, imagine now that the passenger wants to measure the length L of the car. For this purpose he switches on the light bulb, now at the rear of the car, and measures the time t needed for the light to reach the front of the car, and get reflected back by a mirror installed there, according to the following scheme:



He concludes that, as marked above, the length L of the car is given by:

$$L = \frac{ct}{2}$$

Now viewed from the ground, the duration of the event is $T = T_1 + T_2$, where $T_1 > T_2$ are respectively the time needed for the light to travel forward, among others for beating v, and the time for the light to travel back, helped this time by v. More precisely, if l denotes the length of the train car viewed from the ground, the formula of T is:

$$T = T_1 + T_2 = \frac{l}{c - v} + \frac{l}{c + v} = \frac{2lc}{c^2 - v^2}$$

With this data, the formula $T = \gamma t$ of time dilation established before reads:

$$\frac{2lc}{c^2 - v^2} = \gamma t = \frac{2\gamma L}{c}$$

Thus, the two lengths L and l are indeed not equal, and related by:

$$l = \frac{\gamma L(c^2 - v^2)}{c^2} = \gamma L\left(1 - \frac{v^2}{c^2}\right) = \frac{\gamma L}{\gamma^2} = \frac{L}{\gamma}$$

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

With this discussed, time now to get to the real thing, see what happens to our usual $\mathbb{R}^4$. Let us start our discussion with a look at the non-relativistic case. Assuming that the object moves with speed v in the x direction, the frame change is given by:

$$x' = x - vt$$

$$y' = y$$

$$z' = z$$

$$t' = t$$

To be more precise, here the first 3 equations come from the law of motion, and $t' = t$ is the old $t' = t$. In the relativistic setting now, the result is more tricky, as follows:

Theorem 1.19. *In the context of a relativistic object moving with speed v along the x axis, the frame change is given by the Lorentz transformation*

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

$$y' = y$$

$$z' = z$$

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

with $\gamma = 1/\sqrt{1 - v^2/c^2}$ being as usual the Lorentz factor.

Proof. We know that, with respect to the non-relativistic formulae, x is subject to the Lorentz dilation by γ, and we obtain as desired:

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

Regarding y, z, these are obviously unchanged, so done with these too. Finally, regarding time t, a naive thought would suggest that this is subject to a Lorentz contraction by $1/\gamma$, but this is not true, and more thinking leads to the conclusion that we must use the reverse Lorentz transformation, given by the following formulae:

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

$$y = y'$$

$$z = z'$$

By using the formula of x' we can compute t', and we obtain the following formula:

$$\begin{aligned} t' &= \frac{x - \gamma x'}{\gamma v} \\ &= \frac{x - \gamma^2(x - vt)}{\gamma v} \\ &= \frac{\gamma^2 vt + (1 - \gamma^2)x}{\gamma v} \end{aligned}$$

On the other hand, we have the following computation:

$$\gamma^2=\frac{c^2}{c^2-v^2}\implies\gamma^2(c^2-v^2)=c^2\implies(\gamma^2-1)c^2=\gamma^2v^2$$

Thus we can finish the computation of t' as follows:

$$\begin{aligned}t'&=\frac{\gamma^2vt+(1-\gamma^2)x}{\gamma v}\\&=\frac{\gamma^2vt-\gamma^2v^2x/c^2}{\gamma v}\\&=\gamma\left(t-\frac{vx}{c^2}\right)\end{aligned}$$

We are therefore led to the conclusion in the statement. □

Now since y,z are irrelevant, we can put them at the end, and put the time t first, as to be close to x. By multiplying as well the time equation by c, our system becomes:

$$ct'=\gamma(ct-vx/c)$$

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

$$y'=y$$

$$z'=z$$

In linear algebra terms, the result is as follows:

THEOREM 1.20. *The Lorentz transformation is given by*

$$\begin{pmatrix}\gamma&-\beta\gamma&0&0\\-\beta\gamma&\gamma&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}ct\\x\\y\\z\end{pmatrix}=\begin{pmatrix}ct'\\x'\\y'\\z'\end{pmatrix}$$

where $\gamma=1/\sqrt{1-v^2/c^2}$ as usual, and where $\beta=v/c$.

PROOF. In terms of $\beta=v/c$, replacing v, the system looks as follows:

$$ct'=\gamma(ct-\beta x)$$

$$x'=\gamma(x-\beta ct)$$

$$y'=y$$

$$z'=z$$

But this gives the formula in the statement. □

As an illustration, let us verify that the inverse Lorentz transformation is indeed given by reversing the speed, $v\to-v$. With notations as in Theorem 1.19, the result is:

Theorem 1.21. *The inverse of the Lorentz transformation is given by $v \to -v$,*

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

$$y = y'$$

$$z = z'$$

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

where $\gamma = 1/\sqrt{1 - v^2/c^2}$ is as usual the Lorentz factor, identical for v and $-v$.

Proof. In terms of the formalism in Theorem 1.20, reversing the speed $v \to -v$ amounts in reversing the $\beta = v/c$ parameter there:

$$\beta \to -\beta$$

What we have to prove, in order to establish the result, is that by doing so, we obtain the inverse of the matrix appearing there, namely:

$$L = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

That is, we want to prove that the inverse of this matrix is as follows:

$$L^{-1} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

But here, for the verification of the inversion formula $LL^{-1} = 1$, we can restrict the attention to the upper left corner, where the result is clear. $\square$

Let us discuss now what happens to momentum, mass and energy. We would like to fix the momentum conservation equations for the plastic collisions, namely:

$$m = m_1 + m_2$$

$$mv = m_1 v_1 + m_2 v_2$$

However, this cannot really be done with bare hands, and by this meaning with mathematics only. But with some help from experiments, the conclusion is as follows:

Fact 1.22. *When defining the relativistic mass of an object of rest mass $m > 0$, moving at speed v, by the formula*

$$M = \gamma m \quad : \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

this relativistic mass M, and the corresponding relativistic momentum $P = Mv$, are both conserved during collisions.

In other words, the situation here is a bit similar to that of the Galileo addition vs Einstein addition for speeds. The collision equations given above are in fact low-speed approximations of the correct, relativistic equations, which are as follows:

$$M = M_1 + M_2$$

$$Mv = M_1 v_1 + M_2 v_2$$

It remains now to discuss kinetic energy. You have certainly heard of the formula $E = mc^2$, which might actually well be on your T-shirt, now as you read this book, and in this case here is the explanation for it, in relation with the above:

Theorem 1.23. *The relativistic energy of an object of rest mass $m > 0$,*

$$\mathcal{E} = Mc^2 \quad : \quad M = \gamma m$$

which is conserved, as being a multiple of M, can be written as $\mathcal{E} = E + T$, with

$$E = mc^2$$

being its $v = 0$ component, called rest energy of m, and with

$$T = (1 - \gamma)mc^2 \simeq \frac{mv^2}{2}$$

being called relativistic kinetic energy of m.

Proof. All this is a bit abstract, coming from Fact 1.22, as follows:

(1) Given an object of rest mass $m > 0$, consider its relativistic mass $M = \gamma m$, as appearing in Fact 1.22, and then consider the following quantity:

$$\mathcal{E} = Mc^2$$

We know from Fact 1.22 that the relativistic mass M is conserved, so $\mathcal{E} = Mc^2$ is conserved too. In view of this, is makes somehow sense to call $\mathcal{E}$ energy. There is of course no clear reason for doing that, but let's just do it, and we'll understand later.

(2) Let us compute $\mathcal{E}$. This quantity is by definition given by:

$$\mathcal{E} = Mc^2 = \gamma mc^2 = \frac{mc^2}{\sqrt{1 - v^2/c^2}}$$

Since $1/\sqrt{1-x} \simeq 1 + x/2$ for x small, by calculus, we obtain, for v small:

$$\mathcal{E} \simeq mc^2\left(1 + \frac{v^2}{2c^2}\right) = mc^2 + \frac{mv^2}{2}$$

And, good news here, we recognize at right the kinetic energy of m.

(3) But this leads to the conclusions in the statement. Indeed, we are certainly dealing with some sort of energies here, and so calling the above quantity $\mathcal{E}$ relativistic energy

is legitimate, and calling $E = mc^2$ rest energy is legitimate too. Finally, the difference between these two energies $T = \mathcal{E} - E$ follows to be given by:

$$T = (1 - \gamma)mc^2 \simeq \frac{mv^2}{2}$$

Thus, calling T relativistic kinetic energy is legitimate too, and we are done. $\square$

1c. Curved spacetime

Getting now to spacetime, in non-relativistic physics two events are separated by space Δx and by time Δt, with these two separation variables being independent. In relativistic physics this is no longer true, and the correct analogue of this comes from:

THEOREM 1.24. *The following quantity, called relativistic spacetime separation*

$$\Delta s^2 = c^2\Delta t^2 - (\Delta x^2 + \Delta y^2 + \Delta z^2)$$

is invariant under relativistic frame changes.

PROOF. We must prove that the quantity $K = c^2t^2 - x^2 - y^2 - z^2$ is invariant under Lorentz transformations. For this purpose, observe that we have:

$$K = \left\langle \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \right\rangle$$

Now recall that the Lorentz transformation is given by the following formula, where $\gamma = 1/\sqrt{1 - v^2/c^2}$ as usual, and where $\beta = v/c$:

$$\begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix}$$

Thus, if we denote by L the matrix of the Lorentz transformation, and by E the matrix found before, we must prove that for any vector ξ we have:

$$< E\xi, \xi >=< EL\xi, L\xi >$$

Since L is symmetric we have $< EL\xi, L\xi >=< LEL\xi, \xi >$, so we must prove:

$$E = LEL$$

But this is the same as proving $L^{-1}E = EL$, and by using the fact that $L \to L^{-1}$ is given by $\beta \to -\beta$, what we eventually want to prove is that:

$$L_{-\beta}E = EL_\beta$$

So, let us prove this. As usual we can restrict the attention to the upper left corner, call that NW corner, and here we have the following computations:

$$(L_{-\beta}E)_{NW}=\begin{pmatrix}\gamma&\beta\gamma\\ \beta\gamma&\gamma\end{pmatrix}\begin{pmatrix}1&0\\ 0&-1\end{pmatrix}=\begin{pmatrix}\gamma&-\beta\gamma\\ \beta\gamma&-\gamma\end{pmatrix}$$

$$(EL_\beta)_{NW}=\begin{pmatrix}1&0\\ 0&-1\end{pmatrix}\begin{pmatrix}\gamma&-\beta\gamma\\ -\beta\gamma&\gamma\end{pmatrix}=\begin{pmatrix}\gamma&-\beta\gamma\\ \beta\gamma&-\gamma\end{pmatrix}$$

The matrices on the right being equal, this gives the result. □

1d. General relativity

Finally, let us discuss gravity. This can be incorporated too, as follows:

THEOREM 1.25 (Einstein). *The theory of gravity can be suitably modified, and merged with relativity, into a theory called general relativity.*

PROOF. All this is a bit complicated, involving some geometry, as follows:

(1) Before anything, we have seen that in the relativistic context, mass m must be replaced by relativistic mass $M=\gamma m$, and momentum $p=mv$ must be replaced by relativistic momentum $P=Mv$. Thus, as with Galileo and many other things, such as the conservation of mass and of momentum, seen above, there is a bug with the Newton formula $F=\dot{p}$, which must be replaced by something of type $F=\dot{P}$.

(2) In practice now, as a starting point, let us go back to the formula $F=-\Delta V$, that we know well. Geometrically, this suggests looking at the gravitational field of k bodies $M_1,\ldots,M_k$ as being represented by $\mathbb{R}^3$ having k holes in it, and with the heavier the M_i, the bigger the hole, and with poor $m\simeq 0$ having to roll on all this.

(3) Of course we are here in 4D, for the full picture, that of the potential V, or rather of its graph, and in order to better understand this, it is of help to first consider the question where our bodies $M_1,\ldots,M_k$ lie in a plane $\mathbb{R}^2$.

(4) Still staying inside classical mechanics, it is possible to further build on the above picture in (2), which was something rather intuitive, now with some precise math formulae, relating the geometry of V to the motion of m under its influence.

(5) The point now is that, with (4) done, the passage to relativity can be understood as well, by modifying a bit the geometry there, as to fit with relativistic spacetime, and by having the $F=\dot{P}$ idea from (1) in mind too. That is the main idea behind general relativity, and in practice, all this needs a bit of technical geometry and formulae. □

This was for the basics of Einstein's relativity theory. For more, we refer to his book [31], which is a must-read, for any mathematician, physicist, scientist, or non-scientist.

1e. Exercises

Exercises:

Exercise 1.26.

Exercise 1.27.

Exercise 1.28.

Exercise 1.29.

Exercise 1.30.

Exercise 1.31.

Bonus exercise.

CHAPTER 2

Particle physics

2a. Atomic theory

Quantum mechanics was born from the study of the hydrogen atom, which consists of a negative charge, the electron, spinning around a positive charge, a proton. Basic electrodynamics, you would say, that I can solve with the Maxwell equations, that I learned the hard way in school. Very good, so let us start indeed with these equations:

Theorem 2.1. *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 fields E and B, namely those above. To be more precise, 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. □

In relation now with the hydrogen atom, we have good and bad news. The bad news is that the Maxwell equations, as formulated above, are something rather statistical and macroscopic, and simply do not apply to the hydrogen atom. However, we have good news as well, the point here being that the same Maxwell equations can lead us, in a rather twisted way, via light, spectroscopy, and many more, to the hydrogen atom. So, let us explain this. We first have the following consequence of the Maxwell equations:

Theorem 2.2. *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 2.1 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 a key formula due to Biot-Savart, as follows:

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

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 2.2, traveling 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 2.3. *An accelerating or decelerating charge produces electromagnetic radiation, called light, whose frequency and wavelength can be explicitly computed.*

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

In practice now, the classification of light, 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 the 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

Many things can be said about light and optics, and going now straight to the point, the idea is that, with this in hand, we can talk about spectroscopy:

FACT 2.4. *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.*

Going now towards atoms, there is a long story here, involving many discoveries, around 1890-1900, focusing on hydrogen H. We will present here things a bit retrospectively. First on our list is the following discovery, by Lyman in 1906:

Fact 2.5 (Lyman). *The hydrogen atom has spectral lines given by the formula*

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

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

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

called Lyman series of the hydrogen atom.

Observe that all the Lyman series lies in UV, which is invisible to the naked eye. Due to this fact, this series, while theoretically being the most important, was discovered only second. The first discovery, which was the big one, and the breakthrough, was by Balmer, the founding father of all this, back in 1885, in the visible range, as follows:

Fact 2.6 (Balmer). *The hydrogen atom has spectral lines given by the formula*

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

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

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

called Balmer series of the hydrogen atom.

So, this was Balmer's original result, which started everything. As a third main result now, this time in IR, due to Paschen in 1908, we have:

Fact 2.7 (Paschen). *The hydrogen atom has spectral lines given by the formula*

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

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

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

called Paschen series of the hydrogen atom.

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

Conclusion 2.8 (Rydberg, Ritz). *The spectral lines of the hydrogen atom are given by the Rydberg formula, depending on integer parameters* $n_1<n_2$,

$$\frac{1}{\lambda_{n_1n_2}}=R\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right)$$

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

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

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

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

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

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

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

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

From a theoretical physics viewpoint, the main result remains the middle assertion, called Ritz-Rydberg combination principle, which is something quite puzzling. But this combination principle reminds the formula $e_{n_1n_2}e_{n_2n_3} = e_{n_1n_3}$ for the usual matrix units $e_{ij} : e_j \to e_i$. Thus, we are in familiar territory here, and we can start dreaming of:

Thought 2.9. *Observables in quantum mechanics should be some sort of infinite matrices, generalizing the Lyman, Balmer, Paschen lines of the hydrogen atom, and multiplying between them as the matrices do, as to produce further observables.*

Now back to more concrete things, as a main problem that we would like to solve, we have the understanding the intimate structure of matter, at the atomic level. There is of course a long story here, regarding the intimate structure of matter, going back centuries and even millennia ago, and our presentation here will be quite simplified. As a starting point, since we need a starting point for all this, let us agree on:

Claim 2.10. *Ordinary matter is made of small particles called atoms, with each atom appearing as a mix of even smaller particles, namely protons* $+$, *neutrons* 0 *and electrons* $-$, *with the same number of protons* $+$ *and electrons* $-$.

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

But, as a natural question, why should be these electrons $-$ and protons $+$ that small? And also, what about the neutron 0? These are not easy questions, and the fact that it is so came from several clever experiments. Let us first recall that careful experiments with tiny particles are practically impossible. However, all sorts of brutal experiments, such as bombarding matter with other pieces of matter, accelerated to the extremes, or submitting it to huge electric and magnetic fields, do work. And it is such kind of experiments, due to Thomson, Rutherford and others, "peeling off" protons $+$, neutrons 0 and electrons $-$ from matter, and observing them, that led to the conclusion that these small beasts $+,0,-$ exist indeed, in agreement with Claim 2.10.

So, taking now Claim 2.10 for granted, how are then the atoms organized, as mixtures of protons $+$, neutrons 0 and electrons $-$? The answer here lies again in the above-mentioned "brutal" experiments of Thomson, Rutherford and others, which not only proved Claim 2.10, but led to an improved version of it, as follows:

Claim 2.11. *The atoms are formed by a core of protons $+$ and neutrons 0, surrounded by a cloud of electrons $-$, gravitating around the core.*

This is a considerable advance, because we are now into familiar territory, namely some kind of mechanics. And with this in mind, all the pieces of our puzzle start fitting together, and we are led to the following grand conclusion:

Claim 2.12 (Bohr and others). *The atoms are formed by a core of protons and neutrons, surrounded by a cloud of electrons, basically obeying to a modified version of electromagnetism. And with a fine mechanism involved, as follows:*

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

And isn't this beautiful. Moreover, following now Heisenberg, the next claim is that the underlying mathematics in all the above can lead to a beautiful axiomatization of quantum mechanics, as a "matrix mechanics", along the lines of Thought 2.9.

2b. Quantum mechanics

Before explaining what Heisenberg was saying, based on Lyman, Balmer, Paschen, namely developing some sort of "matrix mechanics", let us hear as well the point of view of Schrödinger, which came a few years later. His idea was to forget about exact things, and try to investigate the hydrogen atom statistically. Let us start with:

Question 2.13. *In the context of the hydrogen atom, assuming that the proton is fixed, what is the probability density* $\varphi_t(x)$ *of the position of the electron* e*, at time* t*,*

$$P_t(e\in V)=\int_V\varphi_t(x)dx$$

as function of an intial probability density $\varphi_0(x)$*? Moreover, can the corresponding equation be solved, and will this prove the Bohr claims for hydrogen, statistically?*

In order to get familiar with this question, let us first look at examples coming from classical mechanics. In the context of a particle whose position at time t is given by $x_0+\gamma(t)$, the evolution of the probability density will be given by:

$$\varphi_t(x)=\varphi_0(x)+\gamma(t)$$

However, such examples are somewhat trivial, of course not in relation with the computation of γ, usually a difficult question, but in relation with our questions, and do not apply to the electron. The point indeed is that, in what regards the electron, we have:

Fact 2.14. *In respect with various simple interference experiments:*

1. *The electron is definitely not a particle in the usual sense.*
2. *But in most situations it behaves exactly like a wave.*
3. *But in other situations it behaves like a particle.*

Getting back now to the Schrödinger question, all this suggests to use, as for the waves, an amplitude function $\psi_t(x)\in\mathbb{C}$, related to the density $\varphi_t(x)>0$ by the formula $\varphi_t(x)=|\psi_t(x)|^2$. Not that a big deal, you would say, because the two are related by simple formulae as follows, with $\theta_t(x)$ being an arbitrary phase function:

$$\varphi_t(x)=|\psi_t(x)|^2\quad,\quad\psi_t(x)=e^{i\theta_t(x)}\sqrt{\varphi_t(x)}$$

However, such manipulations can be crucial, raising for instance the possibility that the amplitude function satisfies some simple equation, while the density itself, maybe not. And this is what happens indeed. Schrödinger was led in this way to:

Claim 2.15 (Schrödinger). *In the context of the hydrogen atom, the amplitude function of the electron* $\psi=\psi_t(x)$ *is subject to the Schrödinger equation*

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

m *being the mass,* $h=h_0/2\pi$ *the reduced Planck constant, and* V *the Coulomb potential of the proton. The same holds for movements of the electron under any potential* V*.*

Observe the similarity with the wave equation $\ddot{\varphi}=v^2\Delta\varphi$, and with the heat equation $\dot{\varphi}=\alpha\Delta\varphi$ too. Many things can be said here. Following now Heisenberg and Schrödinger, and then especially Dirac, who did the axiomatization work, we have:

Definition 2.16. *In quantum mechanics the states of the system are vectors of a Hilbert space H, and the observables of the system are linear operators*

$$T:H\to H$$

which can be densely defined, and are taken self-adjoint, $T=T^*$. The average value of such an observable T, evaluated on a state $\xi\in H$, is given by:

$$<T>=<T\xi,\xi>$$

In the context of the Schrödinger mechanics of the hydrogen atom, the Hilbert space is the space $H=L^2(\mathbb R^3)$ where the wave function ψ lives, and we have

$$<T>=\int_{\mathbb R^3}T(\psi)\cdot\bar\psi\,dx$$

which is called "sandwiching" formula, with the operators

$$x\quad,\quad -\frac{ih}{m}\nabla\quad,\quad -ih\nabla\quad,\quad -\frac{h^2\Delta}{2m}\quad,\quad -\frac{h^2\Delta}{2m}+V$$

representing the position, speed, momentum, kinetic energy, and total energy.

In other words, we are doing here two things. First, we are declaring by axiom that various "sandwiching" formulae found before by Heisenberg, involving the operators at the end, that we will not get into in detail here, hold true. And second, we are raising the possibility for other quantum mechanical systems, more complicated, to be described as well by the mathematics of the operators on a certain Hilbert space H, as above.

Now, let us go back to the Schrödinger equation from Claim 2.15. We have:

Proposition 2.17. *We have the following formula,*

$$\dot\varphi=\frac{ih}{2m}\left(\Delta\psi\cdot\bar\psi-\Delta\bar\psi\cdot\psi\right)$$

for the time derivative of the probability density function $\varphi=|\psi|^2$.

Proof. According to the Leibnitz product rule, we have the following formula:

$$\dot\varphi=\frac{d}{dt}|\psi|^2=\frac{d}{dt}(\psi\bar\psi)=\dot\psi\bar\psi+\psi\dot{\bar\psi}$$

On the other hand, the Schrödinger equation and its conjugate read:

$$\dot\psi=\frac{ih}{2m}\left(\Delta\psi-\frac{2m}{h^2}V\psi\right)\quad,\quad\dot{\bar\psi}=-\frac{ih}{2m}\left(\Delta\bar\psi-\frac{2m}{h^2}V\bar\psi\right)$$

By plugging this data, we obtain the following formula:

$$\dot\varphi=\frac{ih}{2m}\left[\left(\Delta\psi-\frac{2m}{h^2}V\psi\right)\bar\psi-\left(\Delta\bar\psi-\frac{2m}{h^2}V\bar\psi\right)\psi\right]$$

But this gives, after simplifying, the formula in the statement. $\square$

As an important application of Proposition 2.17, we have:

THEOREM 2.18. *The Schrödinger equation conserves probability amplitudes,*

$$\int_{\mathbb{R}^3}|\psi_0|^2=1\implies\int_{\mathbb{R}^3}|\psi_t|^2=1$$

in agreement with the basic probabilistic requirement, $P=1$ overall.

PROOF. According to the formula in Proposition 2.17, we have:

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^3}|\psi|^2\,dx &= \int_{\mathbb{R}^3}\frac{d}{dt}|\psi|^2\,dx\\ &= \int_{\mathbb{R}^3}\dot{\varphi}\,dx\\ &= \frac{ih}{2m}\int_{\mathbb{R}^3}\left(\Delta\psi\cdot\bar{\psi}-\Delta\bar{\psi}\cdot\psi\right)dx\end{aligned}$$

Now by remembering the definition of the Laplace operator, we have:

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^3}|\psi|^2\,dx &= \frac{ih}{2m}\int_{\mathbb{R}^3}\sum_i\left(\frac{d^2\psi}{dx_i^2}\cdot\bar{\psi}-\frac{d^2\bar{\psi}}{dx_i^2}\cdot\psi\right)dx\\ &= \frac{ih}{2m}\sum_i\int_{\mathbb{R}^3}\frac{d}{dx_i}\left(\frac{d\psi}{dx_i}\cdot\bar{\psi}-\frac{d\bar{\psi}}{dx_i}\cdot\psi\right)dx\\ &= \frac{ih}{2m}\sum_i\int_{\mathbb{R}^2}\left[\frac{d\psi}{dx}\cdot\bar{\psi}-\frac{d\bar{\psi}}{dx}\cdot\psi\right]_{-\infty}^{\infty}\frac{dx}{dx_i}\\ &= 0\end{aligned}$$

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

Moving now towards hydrogen, we have here the following result:

THEOREM 2.19. *In the case of time-independent potentials V, including the Coulomb potential of the proton, the solutions of the Schrödinger equation*

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

which are of the following special form, with the time and space variables separated,

$$\psi_t(x)=w_t\phi(x)$$

are given by the following formulae, with E being a certain constant,

$$w=e^{-iEt/h}w_0\quad,\quad E\phi=-\frac{h^2}{2m}\Delta\phi+V\phi$$

with the equation for ϕ being called time-independent Schrödinger equation.

Proof. By dividing by ψ, the Schrödinger equation becomes:

$$ih\cdot\frac{\dot{w}}{w}=-\frac{h^2}{2m}\cdot\frac{\Delta\phi}{\phi}+V$$

Now since the left-hand side depends only on time, and the right-hand side depends only on space, both quantities must equal a constant E, and this gives the result. $\square$

Moving ahead with theory, we can further build on Theorem 2.19, with a number of key observations on the time-independent Schrödinger equation, as follows:

Theorem 2.20. *In the case of time-independent potentials V, the Schrödinger equation and its time-independent version have the following properties:*

1. *For solutions of type $\psi = w_t\phi(x)$, the density $\varphi = |\psi|$ is time-independent, and more generally, all quantities of type $< T >$ are time-independent.*
2. *The time-independent Schrödinger equation can be written as $\widehat{H}\phi = E\phi$, with $H = T + V$ being the total energy, of Hamiltonian.*
3. *For solutions of type $\psi = w_t\phi(x)$ we have $< H^k >= E^k$ for any k. In particular we have $< H >= E$, and the variance is $< H^2 > - < H >^2 = 0$.*

Proof. All the formulae are clear indeed from the fact that, when using the sandwiching formula for computing averages, the phases will cancel:

$$< T >=\int_{\mathbb{R}^3}\bar{\psi}\cdot T\cdot\psi\,dx=\int_{\mathbb{R}^3}\bar{\phi}\cdot T\cdot\phi\,dx$$

Thus, we are led to the various conclusions in the statement. $\square$

We have as well the following key result, mathematical this time:

Theorem 2.21. *The solutions of the Schrödinger equation with time-independent potential V appear as linear combinations of separated solutions*

$$\psi=\sum_n c_n e^{-iE_nt/h}\phi_h$$

with the absolute values of the coefficients being given by

$$< H >=\sum_n |c_n|^2E_n$$

$|c_n|$ being the probability for a measurement to return the energy value E_n.

Proof. This is something standard, which follows from Fourier analysis, which allows the decomposition of ψ as in the statement, and that we will not really need, in what follows next. As before, for a physical discussion here, we refer to Griffiths [46]. $\square$

2c. Hydrogen atom

In order to solve now the hydrogen atom, by using the Schrödinger equation, the idea will be that of reformulating this equation in spherical coordinates. We have:

THEOREM 2.22. *The time-independent Schrödinger equation in spherical coordinates separates, for solutions of type* $\phi = \rho(r)\alpha(s,t)$*, into two equations, as follows,*

$$\frac{d}{dr}\left(r^2 \cdot \frac{d\rho}{dr}\right) - \frac{2mr^2}{h^2}(V-E)\rho = K\rho$$

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

with K *being a constant, called radial equation, and angular equation.*

PROOF. We use the following well-known formula for the Laplace operator in spherical coordinates, whose proof can be found in any advanced calculus book:

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

By using this formula, the time-independent Schrödinger equation reformulates as:

$$(V-E)\phi = \frac{h^2}{2m}\left[\frac{1}{r^2} \cdot \frac{d}{dr}\left(r^2 \cdot \frac{d\phi}{dr}\right) + \frac{1}{r^2 \sin s} \cdot \frac{d}{ds}\left(\sin s \cdot \frac{d\phi}{ds}\right) + \frac{1}{r^2\sin^2 s} \cdot \frac{d^2\phi}{dt^2}\right]$$

Let us look now for separable solutions for this latter equation, consisting of a radial part and an angular part, as in the statement, namely:

$$\phi(r,s,t) = \rho(r)\alpha(s,t)$$

By plugging this function into our equation, we obtain:

$$(V-E)\rho\alpha = \frac{h^2}{2m}\left[\frac{\alpha}{r^2} \cdot \frac{d}{dr}\left(r^2 \cdot \frac{d\rho}{dr}\right) + \frac{\rho}{r^2 \sin s} \cdot \frac{d}{ds}\left(\sin s \cdot \frac{d\alpha}{ds}\right) + \frac{\rho}{r^2\sin^2 s} \cdot \frac{d^2\alpha}{dt^2}\right]$$

By multiplying everything by $2mr^2/(h^2\rho\alpha)$, and then moving the radial terms to the left, and the angular terms to the right, this latter equation can be written as follows:

$$\frac{2mr^2}{h^2}(V-E) - \frac{1}{\rho} \cdot \frac{d}{dr}\left(r^2 \cdot \frac{d\rho}{dr}\right) = \frac{1}{\alpha\sin^2 s}\left[\sin s \cdot \frac{d}{ds}\left(\sin s \cdot \frac{d\alpha}{ds}\right) + \frac{d^2\alpha}{dt^2}\right]$$

Since this latter equation is now separated between radial and angular variables, both sides must be equal to a certain constant $-K$, and this gives the result. □

Let us first study the angular equation. The result here is as follows:

THEOREM 2.23. *The separated solutions $\alpha=\sigma(s)\theta(t)$ of the angular equation,*

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

are given by the following formulae, where $l\in\mathbb N$ is such that $K=l(l+1)$,

$$\sigma(s)=P_l^m(\cos s)\quad,\quad\theta(t)=e^{imt}$$

and where $m\in\mathbb Z$ is a constant, and with P_l^m being the Legendre function,

$$P_l^m(x)=(-1)^m(1-x^2)^{m/2}\left(\frac{d}{dx}\right)^m P_l(x)$$

where P_l are the Legendre polynomials, given by the following formula:

$$P_l(x)=\frac{1}{2^l l!}\left(\frac{d}{dx}\right)^l(x^2-1)^l$$

These solutions $\alpha=\sigma(s)\theta(t)$ are called spherical harmonics.

PROOF. This follows indeed from all the above, and with the comment that everything is taken up to linear combinations. We will normalize the wave function later. □

In order to finish our study, it remains to solve the radial equation, for the Coulomb potential V of the proton. As a first manipulation on the radial equation, we have:

PROPOSITION 2.24. *The radial equation, written with $K=l(l+1)$,*

$$(r^2\rho')'-\frac{2mr^2}{h^2}(V-E)\rho=l(l+1)\rho$$

takes with $\rho=u/r$ the following form, called modified radial equation,

$$Eu=-\frac{h^2}{2m}\cdot u''+\left(V+\frac{h^2l(l+1)}{2mr^2}\right)u$$

which is a time-independent 1D Schrödinger equation.

PROOF. With $\rho=u/r$ as in the statement, we have:

$$\rho=\frac{u}{r}\quad,\quad\rho'=\frac{u'r-u}{r^2}\quad,\quad(r^2\rho')'=u''r$$

By plugging this data into the radial equation, this becomes:

$$u''r-\frac{2mr}{h^2}(V-E)u=\frac{l(l+1)}{r}\cdot u$$

By multiplying everything by $h^2/(2mr)$, this latter equation becomes:

$$\frac{h^2}{2m}\cdot u''-(V-E)u=\frac{h^2l(l+1)}{2mr^2}\cdot u$$

But this gives the formula in the statement. As for the interpretation, as time-independent 1D Schrödinger equation, this is clear as well, and with the comment here that the term added to the potential V is some sort of centrifugal term. $\square$

It remains to solve the above equation, for the Coulomb potential of the proton. And we have here the following result, which proves the original claims by Bohr:

THEOREM 2.25 (Schrödinger). *In the case of the hydrogen atom, where V is the Coulomb potential of the proton, the modified radial equation, which reads*

$$Eu=-\frac{h^2}{2m}\cdot u''+\left(-\frac{Ke^2}{r}+\frac{h^2l(l+1)}{2mr^2}\right)u$$

leads to the Bohr formula for allowed energies,

$$E_n=-\frac{m}{2}\left(\frac{Ke^2}{h}\right)^2\cdot\frac{1}{n^2}$$

with $n\in\mathbb{N}$, the binding energy being

$$E_1\simeq-2.177\times10^{-18}$$

with means $E_1\simeq-13.591$ eV.

PROOF. This is again something non-trivial, the idea being as follows:

(1) By dividing our modified radial equation by E, this becomes:

$$-\frac{h^2}{2mE}\cdot u''=\left(1+\frac{Ke^2}{Er}-\frac{h^2l(l+1)}{2mEr^2}\right)u$$

In terms of $\alpha=\sqrt{-2mE}/h$, this equation takes the following form:

$$\frac{u''}{\alpha^2}=\left(1+\frac{Ke^2}{Er}+\frac{l(l+1)}{(\alpha r)^2}\right)u$$

In terms of the new variable $p=\alpha r$, this latter equation reads:

$$u''=\left(1+\frac{\alpha Ke^2}{Ep}+\frac{l(l+1)}{p^2}\right)u$$

Now let us introduce a new constant S for our problem, as follows:

$$S=-\frac{\alpha Ke^2}{E}$$

In terms of this new constant, our equation reads:

$$u''=\left(1-\frac{S}{p}+\frac{l(l+1)}{p^2}\right)u$$

(2) The idea will be that of looking for a solution written as a power series, but before that, we must "peel off" the asymptotic behavior. Which is something that can be done,

of course, heuristically. With $p \to \infty$ we are led to $u'' = u$, and ignoring the solution $u = e^p$ which blows up, our approximate asymptotic solution is:

$$u \sim e^{-p}$$

Similarly, with $p \to 0$ we are led to $u'' = l(l+1)u/p^2$, and ignoring the solution $u = p^{-l}$ which blows up, our approximate asymptotic solution is:

$$u \sim p^{l+1}$$

(3) The above heuristic considerations suggest writing our function u as follows:

$$u = p^{l+1} e^{-p} v$$

So, let us do this. In terms of v, we have the following formula:

$$u' = p^l e^{-p} \left[(l+1-p)v + pv' \right]$$

Differentiating a second time gives the following formula:

$$u'' = p^l e^{-p} \left[\left(\frac{l(l+1)}{p} - 2l - 2 + p \right) v + 2(l+1-p)v' + pv'' \right]$$

Thus the radial equation, as modified in (1) above, reads:

$$pv'' + 2(l+1-p)v' + (S - 2(l+1))v = 0$$

(4) We will be looking for a solution v appearing as a power series:

$$v = \sum_{j=0}^{\infty} c_j p^j$$

But our equation leads to the following recurrence formula for the coefficients:

$$c_{j+1} = \frac{2(j+l+1) - S}{(j+1)(j+2l+2)} \cdot c_j$$

(5) We are in principle done, but we still must check that, with this choice for the coefficients c_j, our solution v, or rather our solution u, does not blow up. And the whole point is here. Indeed, at $j >> 0$ our recurrence formula reads, approximately:

$$c_{j+1} \simeq \frac{2c_j}{j}$$

But, surprisingly, this leads to $v \simeq c_0 e^{2p}$, and so to $u \simeq c_0 p^{l+1} e^p$, which blows up.

(6) As a conclusion, the only possibility for u not to blow up is that where the series defining v terminates at some point. Thus, we must have for a certain index j:

$$2(j+l+1) = S$$

In other words, we must have, for a certain integer $n > l$:

$$S = 2n$$

(7) We are almost there. Recall from (1) above that S was defined as follows:

$$S=-\frac{\alpha Ke^2}{E}\quad:\quad\alpha=\frac{\sqrt{-2mE}}{h}$$

Thus, we have the following formula for the square of S:

$$S^2=\frac{\alpha^2K^2e^4}{E^2}=-\frac{2mE}{h^2}\cdot\frac{K^2e^4}{E^2}=-\frac{2mK^2e^4}{h^2E}$$

Now by using the formula $S=2n$ from (6), the energy E must be of the form:

$$E=-\frac{2mK^2e^4}{h^2S^2}=-\frac{mK^2e^4}{2h^2n^2}$$

Calling this energy E_n, depending on $n\in\mathbb{N}$, we have, as claimed:

$$E_n=-\frac{m}{2}\left(\frac{Ke^2}{h}\right)^2\cdot\frac{1}{n^2}$$

(8) Thus, we proved the Bohr formula. Regarding numerics, the data is as follows:

$$K=8.988\times10^9\quad,\quad e=1.602\times10^{-19}$$

$$h=1.055\times10^{-34}\quad,\quad m=9.109\times10^{-31}$$

But this gives the formula of E_1 in the statement. □

As a first remark, all this agrees with the Rydberg formula, due to:

THEOREM 2.26. *The Rydberg constant for hydrogen is given by*

$$R=-\frac{E_1}{h_0c}$$

where E_1 is the Bohr binding energy, and the Rydberg formula itself, namely

$$\frac{1}{\lambda_{n_1n_2}}=R\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right)$$

simply reads, via the energy formula in Theorem 2.25,

$$\frac{1}{\lambda_{n_1n_2}}=\frac{E_{n_2}-E_{n_1}}{h_0c}$$

which is in agreement with the Planck formula $E=h_0c/\lambda$.

PROOF. Here the first assertion is something numeric, coming from the fact that the formula in the statement gives, when evaluated, the Rydberg constant:

$$R=\frac{-E_1}{h_0c}=\frac{2.177\times10^{-18}}{6.626\times10^{-34}\times2.998\times10^8}=1.096\times10^7$$

Regarding now the second assertion, by dividing $R = -E_1/(h_0c)$ by any number of type n^2 we obtain, according to the energy convention in Theorem 2.25:

$$\frac{R}{n^2} = -\frac{E_n}{h_0c}$$

But these are exactly the numbers which are subject to substraction in the Rydberg formula, and so we are led to the conclusion in the statement. □

In order to investigate heavier atoms, we need to know more about hydrogen. So, let us go back to our study of the Schrödinger equation for it. Our conclusions so far are:

Theorem 2.27. *The wave functions of the hydrogen atom are the following functions, labelled by three quantum numbers, n, l, m,*

$$\phi_{nlm}(r,s,t) = \rho_{nl}(r)\alpha_l^m(s,t)$$

where $\rho_{nl}(r) = p^{l+1}e^{-p}v(p)/r$ with $p = \alpha r$ as before, with the coefficients of v subject to

$$c_{j+1} = \frac{2(j+l+1-n)}{(j+1)(j+2l+2)} \cdot c_j$$

and $\alpha_l^m(s,t)$ being the spherical harmonics found before.

Proof. This follows indeed by putting together all the results obtained so far, and with the remark that everything is up to the normalization of the wave function. □

In what regards the main wave function, that of the ground state, we have:

Theorem 2.28. *With the hydrogen atom in its ground state, the wave function is*

$$\phi_{100}(r,s,t) = \frac{1}{\sqrt{\pi a^3}}\,e^{-r/a}$$

where $a = 1/\alpha$ is the inverse of the parameter appearing in our computations above,

$$\alpha = \frac{\sqrt{-2mE}}{h}$$

called Bohr radius of the hydrogen atom. This Bohr radius is given by

$$a = \frac{h^2}{mKe^2}$$

which numerically means $a \simeq 5.291 \times 10^{-11}$.

Proof. According to our formulae above, the parameter α there is given by:

$$\alpha = \frac{\sqrt{-2mE}}{h} = \frac{1}{h} \cdot m \cdot \frac{Ke^2}{h} = \frac{mKe^2}{h^2}$$

Thus, the inverse $\alpha = 1/a$ is indeed given by the formula in the statement. Regarding now the wave function, we know from the above that this consists of:

$$\rho_{10}(r) = \frac{2e^{-r/a}}{\sqrt{a^3}} \quad , \quad \alpha_0^0(s,t) = \frac{1}{2\sqrt{\pi}}$$

By making the product, we obtain the formula of ϕ_{100} in the statement. Finally, in what regards the numerics, these are as follows:

$$a = \frac{1.055^2 \times 10^{-68}}{9.109 \times 10^{-31} \times 8.988 \times 10^9 \times 1.602^2 \times 10^{-38}} = 5.297 \times 10^{-11}$$

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

In order to improve our results, we will need the following standard fact:

PROPOSITION 2.29. *The polynomials $v(p)$ are given by the formula*

$$v(p) = L_{n-l-1}^{2l+1}(p)$$

where the polynomials on the right, called associated Laguerre polynomials, are given by

$$L_q^p(x) = (-1)^p \left(\frac{d}{dx}\right)^p L_{p+q}(x)$$

with L_{p+q} being the Laguerre polynomials, given by the following formula,

$$L_q(x) = \frac{e^x}{q!}\left(\frac{d}{dx}\right)^q (e^{-x}x^q)$$

called Rodrigues formula for the Laguerre polynomials.

PROOF. The story here is very similar to that of the Legendre polynomials. Consider the Hilbert space $H = L^2[0,\infty)$, with the following scalar product on it:

$$< f, g >= \int_0^\infty f(x)g(x)e^{-x}\,dx$$

The orthogonal basis obtained by applying Gram-Schmidt to the Weierstrass basis $\{x^q\}$ is then formed by the Laguerre polynomials $\{L_q\}$, and this gives the results. □

With the above result in hand, we can now improve our main results, as follows:

THEOREM 2.30. *The wave functions of the hydrogen atom are given by*

$$\phi_{nlm}(r,s,t) = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}}\, e^{-r/na}\left(\frac{2r}{na}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right)\alpha_l^m(s,t)$$

with $\alpha_l^m(s,t)$ being the spherical harmonics found before.

PROOF. This follows indeed by putting together what we have, and then doing some remaining work, concerning the normalization of the wave function. □

2d. Fine structure

What is next? All sorts of corrections to the solution that we found, due to various phenomena that we neglected in our computations, or rather in our modeling of the problem, which can be both of electric and relativistic nature.

So, let us explain now the standard corrections to the Schrödinger solution to the hydrogen atom. We will focus on energy only, so let us start by recalling:

Theorem 2.31 (Schrödinger). *The energy of the ϕ_{nlm} state of the hydrogen atom is independent on the quantum numbers l,m, given by the Bohr formula*

$$E_n=-\frac{\alpha^2}{n^2}\cdot\frac{mc^2}{2}$$

where α is a dimensionless constant, called fine structure constant, given by

$$\alpha=\frac{Ke^2}{hc}$$

which in practice means $\alpha\simeq 1/137$.

Proof. This is the Bohr energy formula that we know, proved by Schrödinger, and reformulated in terms of Sommerfeld's fine structure constant:

(1) We know from Theorem 2.25 that we have the following formula, which can be written as in the statement, by using the fine structure constant α:

$$E_n=-\frac{m}{2}\left(\frac{Ke^2}{h}\right)^2\cdot\frac{1}{n^2}$$

(2) Observe now that our modified Bohr formula can be further reformulated as follows, with T_c being the kinetic energy of the electron traveling at speed c:

$$E_n=-\frac{\alpha^2}{n^2}\cdot T_c$$

Thus α^2, and so α too, is dimensionless, as being a quotient of energies.

(3) Let us doublecheck however this latter fact, the check being instructive. With respect to the SI system that we use, the units for K,e,h,c are:

$$U_K=\frac{m^3\cdot kg}{s^2\cdot C^2}\quad,\quad U_e=C\quad,\quad U_h=\frac{m^2\cdot kg}{s}\quad,\quad U_c=\frac{m}{s}$$

Thus the units for the fine structure constant α are, as claimed:

$$U_\alpha=U_C\cdot U_e^2\cdot U_h^{-1}\cdot U_c^{-1}=\frac{m^3\cdot kg}{s^2\cdot C^2}\cdot C^2\cdot\frac{s}{m^2\cdot kg}\cdot\frac{s}{m}=1$$

(4) In what regards now the numerics, these are as follows:

$$\alpha = \frac{Ke^2/h}{c} \simeq \frac{2.186 \times 10^6}{2.998 \times 10^8} = 7.291 \times 10^{-3} \simeq \frac{1}{137}$$

Here we used a standard estimate for Ke^2/h, from the proof of Theorem 2.25. $\square$

The fine structure constant α is a remarkable quantity, as obvious from the above, and more on it in a moment. Among its other magic features, it manages well 2π factors. Indeed, by using $K = 1/(4\pi\varepsilon_0)$ and $h = h_0/2\pi$, we can write this constant as:

$$\alpha = \frac{e^2}{2\varepsilon_0 h_0 c}$$

Finally, let us record the complete official data for α and its inverse α^{-1}:

$$\alpha = 0.007\ 297\ 352\ 5693(11)$$

$$\alpha^{-1} = 137.035\ 999\ 084(21)$$

As a final comment here, all this lengthy discussion about α might sound a bit like mania, or mysticism. But wait for it. Sometimes soon α will be part of your life.

Moving ahead now with corrections to Theorem 2.31, we will be quite brief, and for further details, we refer as usual to our favorite books, Feynman [37], Griffiths [46] and Weinberg [92]. We first have the following result, which is something non-trivial:

THEOREM 2.32. *There is a relativistic correction to be made to the Bohr energy E_n of the state ϕ_{nlm}, depending on the quantum number l, given by*

$$\mathcal{E}_{nl} = \frac{\alpha^2 E_n}{n^2}\left(\frac{n}{l+1/2} - \frac{3}{4}\right)$$

coming by replacing the kinetic energy by the relativistic kinetic energy.

PROOF. According to Einstein, the relativistic kinetic energy is given by:

$$T = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \ldots$$

The Schrödinger equation, based on $T = p^2/2m$, must be therefore corrected with a $\mathcal{T} = -p^4/(8m^3c^2)$ term, and this leads to the above correction term $\mathcal{E}_{nl}$. $\square$

Equally non-trivial is the following correction, independent from the above one:

THEOREM 2.33. *There is a spin-related correction to be made to the Bohr energy E_n of the state ϕ_{nlm}, depending on the number $j = l \pm 1/2$, given by*

$$\mathcal{E}_{nj} = -\frac{\alpha^2 E_n}{n^2} \cdot \frac{n(j-l)}{(l+1/2)(j+1/2)}$$

coming from the torque of the proton on the magnetic moment of the electron.

PROOF. As we will explain later, the electron has a spin ±1/2, which is naturally associated to the quantum number l, leading to the parameter $j = l \pm 1/2$. But, knowing now that the electron has a spin, the proton which moves around it certainly acts on its magnetic moment, and this leads to the above correction term $\mathcal{E}_{nj}$. $\square$

So, these are the first two corrections to be made, and again, we refer to Feynman [37], Griffiths [46], Weinberg [92] for details. Obviously we don't quite know what we're doing here, but let us add now the above corrections to E_n, and see what we get. We obtain in this way one of the most famous formulae in quantum mechanics, namely:

THEOREM 2.34. *The energy levels of the hydrogen atom, taking into account the fine structure coming from the relativistic and spin-related correction, are given by*

$$E_{nj} = E_n \left[1 + \frac{\alpha^2}{n^2}\left(\frac{n}{j+1/2} - \frac{3}{4}\right)\right]$$

with $j = l \pm 1/2$ being as above, and with α being the fine structure constant.

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

$$\begin{aligned}
\mathcal{E}_{nl} + \mathcal{E}_{nj} &= \frac{\alpha^2 E_n}{n^2}\left(\frac{n}{l+1/2} - \frac{3}{4} - \frac{n(j-l)}{(l+1/2)(j+1/2)}\right)\\
&= \frac{\alpha^2 E_n}{n^2}\left(\frac{n}{l+1/2}\left(1 - \frac{j-l}{j+1/2}\right) - \frac{3}{4}\right)\\
&= \frac{\alpha^2 E_n}{n^2}\left(\frac{n}{j+1/2} - \frac{3}{4}\right)
\end{aligned}$$

Thus the corrected formula of the energy is as follows:

$$\begin{aligned}
E_{nj} &= E_n + \mathcal{E}_{nl} + \mathcal{E}_{nj}\\
&= E_n + \frac{\alpha^2 E_n}{n^2}\left(\frac{n}{j+1/2} - \frac{3}{4}\right)
\end{aligned}$$

We are therefore led to the conclusion in the statement. $\square$

Summarizing, quantum mechanics is more complicated than what originally appears from Schrödinger's solution of the hydrogen atom. Which was something quite complicated too, we must admit that. And the story is not over here, because on top of the above fine structure correction, which is of order α^2, we have afterwards the Lamb shift, which is an order α^3 correction, then the hyperfine splitting, and more.

Quite remarkably, Feynman and others managed to find a global way of viewing all the phenomena that can appear, corresponding to an infinite series in α. To be more precise, their theory, called quantum electrodynamics (QED), is an advanced version of quantum mechanics, still used nowadays for any delicate computation.

2e. Exercises

Exercises:

Exercise 2.35.

Exercise 2.36.

Exercise 2.37.

Exercise 2.38.

Exercise 2.39.

Exercise 2.40.

Bonus exercise.

CHAPTER 3

Fusion, fission

3a. Fusion, fission

Let us investigate now the case of arbitrary atoms. We will need some general theory for the many-particle systems in quantum mechanics. Let us start with:

Definition 3.1. *The wave function of a system of electrons $e_1,\ldots,e_Z$, given by*

$$P_t\big(e_1\in V_1,\ldots,e_Z\in V_Z\big)=\int_{V_1\times\ldots\times V_Z}|\psi_t(x_1,\ldots,x_Z)|^2dx$$

is governed by the Schrödinger equation $ih\dot{\psi}=\widehat{H}\psi$, with Hamiltonian as follows,

$$\widehat{H}=-\frac{h}{2m}\sum_i\Delta_i+Ke^2\sum_{i<j}\frac{1}{||x_i-x_j||}+V(x_1,\ldots,x_Z)$$

with the middle sum standing for the Coulomb repulsions between them.

As before with the one-particle Schrödinger equation, there is a long story with all this, and for cutting short with the discussion here, this is what experiments lead to.

In general, and in fact at any $Z>1$, and so even at $Z=2$, the above Schrödinger equation is pretty much impossible to solve, due to the Coulomb repulsion term, which makes the mathematics extremely complicated. In fact, as an illustrating analogy here, managing that Coulomb repulsion term is more or less the same thing as solving the N-body problem in classical mechanics, for bodies with equal mass.

We will be interested here in the case of atoms, where V is the Coulomb attraction potential coming from a Ze charge. Here the problem to be solved is as follows:

Problem 3.2. *Consider an atom of atomic number Z, meaning a fixed Ze charge, surrounded by electrons $e_1,\ldots,e_Z$. The problem is to solve the Schrödinger equation*

$$ih\dot{\psi}=\widehat{H}\psi$$

with Hamiltonian as follows,

$$\widehat{H}=\sum_i\left(-\frac{h}{2m}\Delta_i-\frac{KZe^2}{||x_i||}\right)+Ke^2\sum_{i<j}\frac{1}{||x_i-x_j||}$$

or at least to understand how $e_1,\ldots,e_Z$ manage to live together, in a stable way.

A first idea would be of course that of ignoring the right term, Coulomb repulsion. In the simplest case, that of the helium atom, the situation is as follows:

FACT 3.3. *For the helium atom, $Z = 2$, ignoring the Coulomb repulsion between electrons leads, via separation of variables, to product wave functions*

$$\phi(x_1, x_2) = \phi'_{n_1 l_1 m_1}(x_1)\phi'_{n_2 l_2 m_2}(x_2)$$

with the prime signs standing for the doubling $e \to 2e$ of the central charge, with energies:

$$E_{n_1 n_2} = 4(E_{n_1} + E_{n_2})$$

This model predicts a ground state energy for helium given by

$$E_0 = 8 \times (-13.6) = -109 \text{ eV}$$

which is considerably smaller than the observed $E_0 = -79$ eV.

As a partial conclusion to what we have so far, things not going on every well, and in order to advance, we will probably need to invest a lot of time in learning how to solve complicated Schrödinger equations, and why not buying a super-computer too.

Moving ahead, let us focus on a more modest question, that at the end of Problem 3.2, namely understanding how the electrons $e_1, \ldots, e_Z$ manage to live together. Here our method of ignoring the Coulomb repulsion between electrons is not that bad, and for helium for instance, we are led in this way to some interesting conclusions. For instance the excited states of helium must appear as products as follows:

$$\phi_{100}(x_1)\phi_{nlm}(x_2) \quad , \quad \phi_{nlm}(x_1)\phi_{100}(x_2)$$

And this is actually quite close to reality, if we add the particle spin to our discussion. So, speaking now spin, this is indeed something that we ignored so far in the above. And spin is in fact a key component to our problem, because we have:

FACT 3.4 (Pauli exclusion principle). *Two electrons cannot occupy the same quantum numbers n, l, m, with same spin $s = \pm 1/2$.*

So, this is the famous Pauli exclusion principle, giving the golden key to the understanding of $Z \geq 2$ atoms. There are of course many things that can be said about it. A sample quantum mechanics book will probably tell you first something about bosons and fermions, coming with exactly 0 evidence, then some more things about electrons, of type "they are the same, but not really, and everything is entangled, but is it really entangled", and finally formulate the Pauli exclusion principle, as a theorem.

We will not get into this here, and take the Pauli exclusion principle as it is, a physics fact. However, talking philosophy, personally I always think at it as coming from the "human nature of electrons". To be more precise, when I'm at my office typing the present book, I feel like occupying some precise quantum numbers, with precise spin.

And if 30 colleagues, all typing physics books too, manage to come by surprise to my office, and squeeze there like sardines, I will surely find a way of getting rid of them, and disposing of their bodies. So now that a poor human like me can do this, why shouldn't a mighty electron be able to do the same. This is the Pauli exclusion principle.

We have now all the ingredients for discussing the known atoms, or chemical elements, $Z = 1, \ldots, 118$. These can be arranged in a table, called periodic table, as follows:

	1	2		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	H 1																		He 2
2	Li 3	Be 4												B 5	C 6	N 7	O 8	F 9	Ne 10
3	Na 11	Mg 12												Al 13	Si 14	P 15	S 16	Cl 17	Ar 18
4	K 19	Ca 20		Sc 21	Ti 22	V 23	Cr 24	Mn 25	Fe 26	Co 27	Ni 28	Cu 29	Zn 30	Ga 31	Ge 32	As 33	Se 34	Br 35	Kr 36
5	Rb 37	Sr 38		Y 39	Zr 40	Nb 41	Mo 42	Tc 43	Ru 44	Rh 45	Pd 46	Ag 47	Cd 48	In 49	Sn 50	Sb 51	Te 52	I 53	Xe 54
6	Cs 55	Ba 56	l	Lu 71	Hf 72	Ta 73	W 74	Re 75	Os 76	Ir 77	Pt 78	Au 79	Hg 80	Tl 81	Pb 82	Bi 83	Po 84	At 85	Rn 86
7	Fr 87	Ra 88	a	Lr 103	Rf 104	Db 105	Sg 106	Bh 107	Hs 108	Mt 109	Ds 110	Rg 111	Cn 112	Nh 113	Fl 114	Mc 115	Lv 116	Ts 117	Og 118
			l:	La 57	Ce 58	Pr 59	Nd 60	Pm 61	Sm 62	Eu 63	Gd 64	Tb 65	Dy 66	Ho 67	Er 68	Tm 69	Yb 70		
			a:	Ac 89	Th 90	Pa 91	U 92	Np 93	Pu 94	Am 95	Cm 96	Bk 97	Cf 98	Es 99	Fm 100	Md 101	No 102		

Here the horizontal parameter $1, \ldots, 18$ is called the group, and the vertical parameter $1, \ldots, 7$ is called the period. The two bottom rows consist of lanthanum ${}_{57}$La and its followers, called lanthanides, and of actinium ${}_{89}$Ac and its followers, called actinides.

Thus, the periodic table, when correctly drawn, but no one does that because of obvious typographical reasons, is in fact a 7×32 table. Note here that, according to our 7×18 convention, which is the standard one, lanthanides and actinides don't have a group number $1, \ldots, 18$. Their group is by definition "lanthanides" and "actinides".

We will comment in a moment on all this, but before anything:

ADVICE 3.5. *Learn their names.*

This is a serious advice, the periodic table being the main theorem of mathematics, physics, chemistry, biology and engineering combined. So if there's one theorem to be learned, full statement, that is the one. In case you're out of memory, just erase from your brain everything that you learned so far from this book of mine, and learn instead that 118 elements. Please do it for me, this being my final wish, from the death bed.

Actually, in order to get started here, here are the names up to krypton ${}_{36}$Kr, which are absolutely needed for everything, and must be all learned, to start with:

Definition 3.6. *The elements up to krypton* ${}_{36}$Kr *are as follows:*

1. *Hydrogen* ${}_{1}$H, *helium* ${}_{2}$He.
2. *Lithium* ${}_{3}$Li, *beryllium* ${}_{4}$Be, *boron* ${}_{5}$B, *carbon* ${}_{6}$C, *nitrogen* ${}_{7}$N, *oxygen* ${}_{8}$O, *fluorine* ${}_{9}$F, *neon* ${}_{10}$Ne.
3. *Sodium* ${}_{11}$Na, *magnesium* ${}_{12}$Mg, *aluminium* ${}_{13}$Al, *silicon* ${}_{14}$Si, *phosphorus* ${}_{15}$P, *sulfur* ${}_{16}$S, *chlorine* ${}_{17}$Cl, *argon* ${}_{18}$Ar.
4. *Potassium* ${}_{19}$K, *calcium* ${}_{20}$Ca, *scandium* ${}_{21}$Sc, *titanium* ${}_{22}$Ti, *vanadium* ${}_{23}$V, *chromium* ${}_{24}$Cr, *manganese* ${}_{25}$Mn, *iron* ${}_{26}$Fe, *cobalt* ${}_{27}$Co.
5. *Nickel* ${}_{28}$Ni, *copper* ${}_{29}$Cu, *zinc* ${}_{30}$Zn, *gallium* ${}_{31}$Ga, *germanium* ${}_{32}$Ge, *arsenic* ${}_{33}$As, *selenium* ${}_{34}$Se, *bromine* ${}_{35}$Br, *krypton* ${}_{36}$Kr.

Observe that all names fit with the abbreviations, expect for sodium ${}_{11}$Na, coming from the Latin natrium, potassium ${}_{19}$K, coming from the Latin kalium, iron ${}_{26}$Fe coming from the Latin ferrum, and also copper ${}_{29}$Cu, coming from the Latin cuprum.

In what regards the elements heavier than krypton ${}_{36}$Kr, it is heartbreaking to sort them out, but as a useful complement to the above list, we have:

Fact 3.7. *Remarkable elements heavier than krypton* ${}_{36}$Kr *include:*

1. *Noble gases: xenon* ${}_{54}$Xe, *radon* ${}_{86}$Rn.
2. *Noble metals: silver* ${}_{47}$Ag, *iridium* ${}_{77}$Ir, *platinum* ${}_{78}$Pt, *gold* ${}_{47}$Au.
3. *Heavy metals: mercury* ${}_{80}$Hg, *lead* ${}_{82}$Pb.
4. *Radioactive: polonium* ${}_{84}$Po, *radium* ${}_{88}$Ra, *uranium* ${}_{92}$U, *plutonium* ${}_{94}$Pu.
5. *Miscellaneous: rubidium* ${}_{37}$Rb, *strontium* ${}_{38}$Sr, *molybdenum* ${}_{42}$Mo, *technetium* ${}_{43}$Tc, *cadmium* ${}_{48}$Cd, *tin* ${}_{50}$Sn, *iodine* ${}_{53}$I, *caesium* ${}_{55}$Cs, *tungsten* ${}_{74}$Tu, *bismuth* ${}_{83}$Bi, *francium* ${}_{87}$Fr, *americium* ${}_{95}$Am.

Here the abbreviations not fitting with English names come from the Latin or sometimes Greek argentum ${}_{47}$Ag, aurum ${}_{47}$Au, hydrargyrum ${}_{80}$Hg, plumbum ${}_{82}$Pb and stannum ${}_{50}$Sn. The noble gases in (1) normally include oganesson ${}_{118}$Og as well. The noble metals in (2) are something subjective. There are of course plenty of other heavy metals (3), or radioactive elements (4). As for the list in (5), this is something subjective, basically a mixture of well-known metals used in engineering, and some well-known bad guys in the context of nuclear fallout. Technetium ${}_{43}$Tc is a bizarre element, human-made.

But let us not forget about quantum mechanics, and what we wanted to do, namely discuss electron structure. As a first observation, we have:

FACT 3.8. *Any* $Z = 1, \ldots, 118$ *corresponds to a unique element, having* Z *protons in the core, and* Z *electrons around it. This element might come with isotopes, depending on the number of neutrons in the core, can be in ground state or excited states, can get ionized, and so on, but all there versions are "family", and the element is unique.*

This fact is something which might look very natural, with no need for explanation for it, but after some thinking, is this really that natural. And the answer here is that no, if you don't know quantum mechanics, and yes, if you know some, as we do:

(1) For the purposes of our question, we can assume that we are in the context of Problem 3.2, and with the Coulomb repulsions between electrons ignored.

(2) But then, we are a bit in the same situation as in Fact 3.3, and the analysis there, based on hydrogen theory modified via $e \to Ze$, carries over.

(3) And so, the Z electrons will arrange on various energy levels, subject to Pauli exclusion, as to occupy a state of lowest possible energy, so the solution is unique.

In fact, we can now understand the electron structure of the various elements, and also how the periodic table is exactly made, the conclusions here being as follows:

FACT 3.9. *For the element having atomic number* Z*, the electrons will occupy successively the various positions with quantum numbers* $n, l, m \in \mathbb{N}$ *and spin* $s = \pm 1/2$*, such as the total binding energy to be minimal. In practice, the period* $1, \ldots, 7$ *corresponds to the highest* n *occupied, and the group* $1, \ldots, 18$ *comes from* l, m, s*.*

This is of course something very basic, and there is a detailed analysis to be done afterwards, for $Z = 1, \ldots, 118$. For the elements up to krypton $_{36}$Kr, the list of electron configurations can be found for instance in Feynman [**37**] or Griffiths [**46**].

Getting now to radioactivity, the story here is long and fascinating:

(1) The fact that uranium $_{92}$U, which is not that uncommon in the earth, at certain places, is naturally radioactive has been known since ages, with various items such as pottery being decorated with uranium based paint, as to glow in the dark.

(2) Official science started to investigate this phenomenon quite late, at the very end of the 19th century, with the work of Henri Becquerel and Marie Curie on uranium salts. A bit later, many others joined, and there was particular excitement in regards with polonium $_{84}$Po and radium $_{88}$Ra. In fact, looking a bit retrospectively, there are 8

culprits of this type, relatively light chemical elements having no stable isotopes, namely technetium, promethium, bismuth, polonium, astatine, radon, francium and radium:

$${}_{43}\mathrm{Tc}\quad {}_{61}\mathrm{Pm}\quad {}_{83}\mathrm{Bi}\quad {}_{84}\mathrm{Po}\quad {}_{85}\mathrm{At}\quad {}_{86}\mathrm{Rn}\quad {}_{87}\mathrm{Fr}\quad {}_{88}\mathrm{Ra}$$

(3) As for uranium itself, this is part of a series of 15 heavy elements, called actinides, coming right after radium in the periodic table, which are actinium, thorium, protactinium, uranium, neptunium, plutonium, americium, curium, berkelium, californium, einsteinium, fermium, mendelevium, nobelium and lawrencium:

$${}_{89}\mathrm{Ac}\quad {}_{90}\mathrm{Th}\quad {}_{91}\mathrm{Pa}\quad {}_{92}\mathrm{U}\quad {}_{93}\mathrm{Np}\quad {}_{94}\mathrm{Pu}\quad {}_{95}\mathrm{Am}\quad {}_{96}\mathrm{Cm}$$

$${}_{97}\mathrm{Bk}\quad {}_{98}\mathrm{Cf}\quad {}_{99}\mathrm{Es}\quad {}_{100}\mathrm{Fm}\quad {}_{101}\mathrm{Md}\quad {}_{102}\mathrm{No}\quad {}_{103}\mathrm{Lr}$$

(4) More recently, the open questions remain numerous, especially in what concerns the continuation of the periodic table, beyond actinides. Here there are presently known 15 more elements, called super-heavy, which are rutherfordium, dubnium, seaborgium, bohrium, hassium, meitnerium, darmstadtium, roentgenium, copernicium, nihonium, flerovium, moscovium, livermorium, tennessine and oganesson:

$${}_{104}\mathrm{Rf}\quad {}_{105}\mathrm{Db}\quad {}_{106}\mathrm{Sg}\quad {}_{107}\mathrm{Bh}\quad {}_{108}\mathrm{Hs}\quad {}_{109}\mathrm{Mt}\quad {}_{110}\mathrm{Ds}\quad {}_{111}\mathrm{Rg}$$

$${}_{112}\mathrm{Cn}\quad {}_{113}\mathrm{Nh}\quad {}_{114}\mathrm{Fl}\quad {}_{115}\mathrm{Mc}\quad {}_{116}\mathrm{Lv}\quad {}_{117}\mathrm{Ts}\quad {}_{118}\mathrm{Og}$$

To be more precise, a first challenging problem is that of understanding the chemistry of these elements, which are too short-lived for chemistry experiments, and with the known theoretical models being too complex. As for the other challenging problem, this is of course going beyond oganesson, either with experiments, or with theory.

3b.

3c.

3d.

3e. Exercises

Exercises:

Exercise 3.10.

Exercise 3.11.

Exercise 3.12.

Exercise 3.13.

Exercise 3.14.

Exercise 3.15.

Bonus exercise.

CHAPTER 4

Exotic matter

4a. Exotic matter

We must first talk about gases. Let us start with the following basic fact, which was the beginning of everything, going back to work of Boyle, Charles, Avogardo, Gay-Lussac, Clapeyron and others from the 17th, 18th and 19th centuries, and with final touches from Maxwell, Boltzmann, Gibbs and others, in the late 19th and early 20th centuries:

Fact 4.1. *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".

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

Error. The point indeed comes from the following:

Explanation 4.2. *In the equation of state $PV=kT$, as formulated above, the pressure P and the temperature T appear more precisely as follows,*

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

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

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

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

THEOREM 4.3. *The pressure P, volume V and total kinetic energy K of a gas, having point molecules, with no collisions between them, satisfy*

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

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

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

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

$$\begin{matrix} & = & = & = & & \\ || & & & & || & \\ || & & \nearrow_v & & || & \to P \\ || & & \bullet_m & & || & \\ & = & = & = & & \end{matrix}$$

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

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

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

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

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

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

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

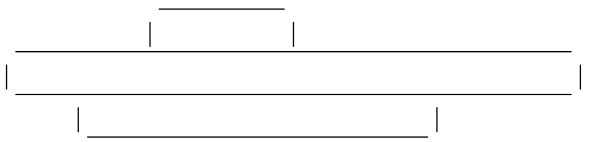
$$P = \frac{m||v||^2}{dV} = \frac{2K}{dV}$$

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

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

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

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



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

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

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

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

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

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

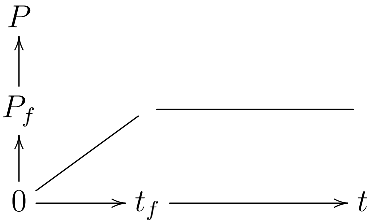
Thus, we are led to the conclusion in the statement. $\square$

Observe that, when carefully looking at the above proof, we can see that pressure P is not something instantaneous, but rather something statistical. In fact, we have:

THEOREM 4.4. *In the context of a gas consisting of point molecules, with no collisions between them, the correct time for reading the correct pressure is*

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

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



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

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

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

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

(2) In the general case now, that of a d-dimensional gas, with $d=1,2,3$, the same argument carries on, with the only change being that each molecular speed $v_i\in\mathbb{R}^d$ is now replaced by its horizontal component $v_{i1}\in\mathbb{R}$, which by statistical reasons has squared magnitude as follows, as explained in the proof of Theorem 4.3:

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

Thus, with respect to (1), the correct final pressure must be adjusted by a d factor, and becomes $P_f=2K/dV$, as in Theorem 4.3. As for the correct reading time, this must be adjusted by a $\sqrt{d}$ factor, and becomes $t_f=2\sqrt{d}\,V^{1/3}/||v||$, as claimed. $\square$

All this is quite interesting, and many other things can be said about pressure P alone, which is a quite fascinating quantity. We will be back to this, later in this book.

At a more advanced level now, dealing with the internal mechanism of the $PV=kT$ formula, we have the following famous result, due to Maxwell and Boltzmann:

Theorem 4.5 (Maxwell, Boltzmann). *The molecular speeds $v \in \mathbb{R}^3$ of a gas in thermal equilibrium are subject to the Maxwell-Boltzmann distribution formula*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

(7) Observe that the above proof has in fact little physical content, with the whole thing being basically obtained by using $PV=2K/3$, which is a mathematical theorem, then $PV=kT$, which is a crucial physics fact, and finally by invoking several times a number of clever symmetry arguments, and doing some calculus.

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

We will be back to the Maxwell-Boltzmann formula later, with some further details, when discussing the internal structure of gases in a more systematic way.

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

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

Questions 4.6. *What is the work done by a gas pushing a piston? What about the work done by a gas evolving on a path γ, in the state space $f(P,V,T)=0$?*

In order to answer this question, we will need some mathematics. Assume that we have a path in $\mathbb R^3$, which can be thought of as corresponding to a function as follows:

$$\gamma:[0,1]\to\mathbb R^3$$

Observe that this function γ is not exactly the path itself, for instance because the functions $r\to\gamma(r^2)$, or $r\to\gamma(1-r)$, and so on, produce the same path, parametrized differently. Our claim now is that we can talk about integrals as follows, with $f,g,h:\mathbb R^3\to\mathbb R$, which are independent on the chosen parametrization of our path:

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

Indeed, by using the obvious symmetry between the coordinates x_1, x_2, x_3, in order to construct such integrals, we just need to construct integrals of the following type:

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

But here, by thinking infinitesimally, we can assume that our path $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ has the property that its first component $\gamma_1 : [0,1] \to \mathbb{R}$ is increasing, or decreasing, and so is bijective on its image $\gamma_1[0,1] = [a,b]$. Thus, we can reparametrize our path in terms of the first coordinate x_1, with the help of some functions φ, ψ, as follows:

$$\gamma(x_1) = (x_1, \varphi(x_1), \psi(x_1))$$

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

$$J = \int_a^b f(x_1, \varphi(x_1), \psi(x_1))dx_1$$

Thus, we have our definition for the contour integrals, and some easy mathematics does rest, such as filling the small gaps in the above construction, and also proving that these contour integrals do not depend on the choice of the parametrization γ.

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

$$\alpha = f(x)dx_1 + g(x)dx_2 + h(x)dx_3$$

Obviously, these are something rather mathematical, and many things can be said here. However, we can have some physical intuition on them. Assume indeed that we are given a function as follows, that you can think for instance as corresponding to an external force, with $F(x) \in \mathbb{R}^3$ being the force vector applied at a given point $x \in \mathbb{R}^3$:

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

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

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

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

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

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

Definition 4.7. *The work done by a force $F = F(x)$ for moving a particle from point $p \in \mathbb{R}^3$ to point $q \in \mathbb{R}^3$ via a given path $\gamma : p \to q$ is the following quantity:*

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

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

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

Proposition 4.8. *The mechanical work done by a gas pushing a piston is*

$$W = \int_a^b PdV$$

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

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

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

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

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

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

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

Theorem 4.9. *The mechanical work done by a gas evolving on a path γ, in the state space $f(P, V, T) = 0$, is given by:*

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

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

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

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

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

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

Definition 4.10. *A transformation of a gas is called:*

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

Here all the terminology comes from Greek, with isos meaning equal, baros meaning weight, chora meaning space, and therme meaning heat.

As an illustration for Theorem 4.9, for an ideal gas, we have the following result:

Theorem 4.11. *For an ideal gas, $PV = kT$, isothermally expanding, we have*

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

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

Proof. This follows indeed from Theorem 4.9, in either formulation:

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

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

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

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

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

Finally, still talking basics, the first law of thermodynamics invites us to introduce, as a complement to the various elementary notions from Definition 4.10:

Definition 4.12. *A transformation is called adiabatic if:*

$$Q=0$$

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

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

Theorem 4.13. *The adiabatic transformations of an ideal gas,* $PV=kT$, *satisfy*

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

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

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

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

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

$$C_VdT+PdV=0$$

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

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

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

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

Theorem 4.14. *Beyond the ideal gas setting, stating that we should have* $PV=kT$, *the gases are subject to the Van der Waals equation*

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

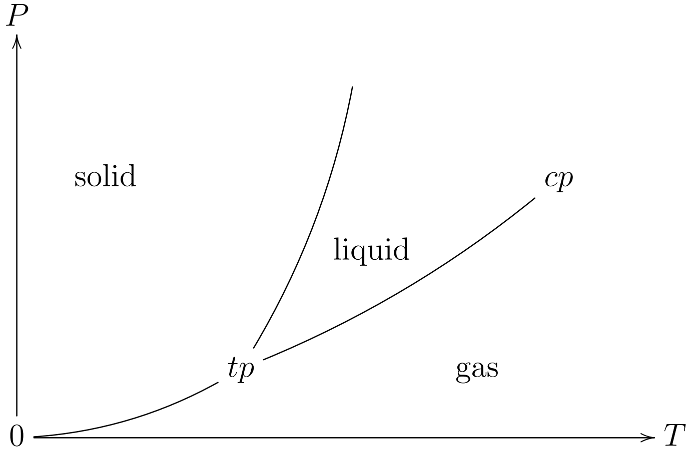
depending on two parameters $\alpha,\beta>0$.

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

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

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

Fact 4.15. *Ordinary matter appears in* 3 *forms, namely solid, liquid and gaseous, roughly appearing according to the following generic diagram*



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

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

Generally speaking, the fluids are by definition the non-solids, and as such, they fall into liquids, gases and plasma. Thus, a fluid is something quite complicated, whose understanding requires advanced thermodynamics, and states of matter theory.

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

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

(1) Compressibility.

(2) Viscosity.

Regarding compressibility, this is certainly a property of the gases, but not of most of the liquids, at least when idealized. We will assume here that our fluid is incompressible, which in practice means more or less that we are dealing with liquids, of rather "regular" type. However, this will be not the general rule, for instance because materials like sand, or snow, that we are particularly interested in, in view of their obvious strong link with classical mechanics, and with our modeling 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 4.16. *We are interested in the basic modeling 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 4.17. *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 modeling questions for the incompressible fluids, we will be mostly interested in the Euler equation, and its versions.

4b.

4c.

4d.

4e. Exercises

Exercises:

Exercise 4.18.

Exercise 4.19.

Exercise 4.20.

Exercise 4.21.

Exercise 4.22.

Exercise 4.23.

Bonus exercise.

Part II

Low energy

I will hold back tears
So I can move in the right direction
I have faced my fears
Now, I can move in the right direction

CHAPTER 5

Low temperature

5a. Low temperature

Let us have a look now at extreme cold. Things are quite fascinating here, and for getting started, we must understand what 0 degrees means. Here is the claim:

Claim 5.1. *The temperature as measured before, with an arbitrarily calibrated thermometer, can be linearly rescaled into an absolute temperature, which is positive,*

$$T > 0$$

and is given by the following equation, valid for any $T_1 < T_2$,

$$\frac{T_1}{T_2} = 1 - \eta$$

where η is the efficiency of a reversible engine operating in the regime $[T_1, T_2]$.

Getting started, let us first recall the second law of thermodynamics, in the Clausius and the Lord Kelvin formulation. In the Clausius formulation, this is as follows:

Law 5.2 (Second law, Clausius' formulation). *Heat flows from warm to cold.*

At the first glance, this might seem like a rather purely scientific, non-profit statement. However, the point is that this law is in fact equivalent to the following statement, due to Lord Kelvin, which is suddenly something very instructive:

Law 5.3 (Second law, Kelvin's formulation). *It is impossible to extract heat from a source staying at constant temperature T, and convert it into work.*

In other words, this tells us that we cannot make money by cooling the land, and converting that heat into work, then electricity, that we can sell afterwards.

Before getting forward, let us mention that Law 5.2 and Law 5.3 are equivalent:

Fact 5.4. *The second law in Clausius' formulation, Law 5.2, is equivalent to the second law in Lord Kelvin's formulation, Law 5.3.*

Of course, while quite intiutive, when thinking a bit, this remains something which is not obvious. However, this fact is true, and is even a theorem. We will prove this theorem in a moment, after knowing a bit more about engines, which are needed in the proof.

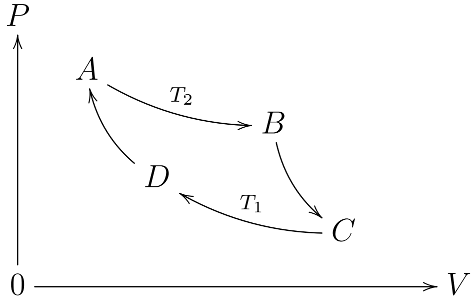
Getting started now, our workhorses for exploiting the second law will be the Carnot engines. These are somewhat mathematical objects, constructed as follows:

Definition 5.5. *A Carnot engine is a 4-cycle engine, functioning between temperatures $T_1 < T_2$, consisting of:*

(1) *An isothermal expansion at T_2, absorbing heat.*
(2) *An adiabatic expansion, down to temperature T_1.*
(3) *An isothermal compression at T_1, expelling heat.*
(4) *An adiabatic compression, back to temperature T_2.*

So, what kind of engine is this. In a nutshell, these are the best possible engines, and quite fascinating objects, no one really knowing to which area of science they belong. We will see that, from the perspective of physics, such things rather belong to mathematics. From the perspective of mathematics, such things look more as engineering. And from the perspective of engineering, such things look more like theoretical physics.

In order to get familiar with such engines, let us discuss their functioning. We can represent a Carnot engine on a (P,V) diagram, as follows, with the horizontals AB, CD being the isothermals, and the verticals BC, DA being the adiabatics:



Regarding now the engineering, the Carnot engines can be implemented as 1-cylinder, 4-stroke engines, with the functioning of one full cycle being as follows:

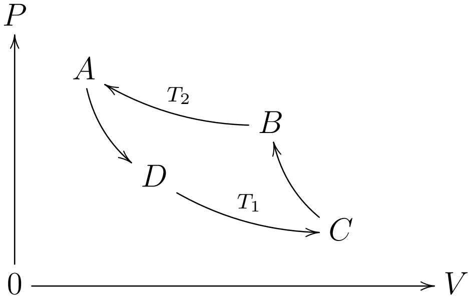
$$\begin{array}{cccccc}
 & A & B & C & D & A' \\
V_C & & & \perp & & \\
V_B & & \perp & & & \\
V_D & & & & \perp & \\
V_A & \perp & & & & \perp \\
 & & & & & \\
 & T_2 & T_2 & \equiv & T_1 & \equiv
\end{array}$$

To be more precise, pictured here is the height of the piston $\perp$ at every step, with the lower data corresponding to the transformations made, T_1, T_2 standing for the temperatures of the isothermals, and $\equiv$ standing for an insulator, needed for the adiabatics.

Obviously, a Carnot engine is an engine as we know them, converting heat into work. To be more precise, an amount of heat Q_2 is absorbed during the upper isotherm AB, and is transformed into an amount of work W during the lower isotherm CD, with an amount of heat Q_1 expelled as well during this lower isotherm CD. And with the adiabatics being there for adjusting the machine, as to make it work as a cycle, in the $[T_1,T_2]$ regime.

As an interesting theoretical remark about the Carnot engines, that we will need in what follows, for theoretical purposes, but which is of practical use as well, we have:

Remark 5.6. *When running a Carnot engine in the other sense, namely*



what we have is a refrigerator, absorbing heat Q_1 on the isotherm DC, helped by some input work W, and expelling it as heat Q_2 during the isotherm BA.

Observe here in passing that due to $Q_2 > Q_1$, using a refrigerator for cooling the room during the Summer, with the door wide open, is not a good idea.

We have seen that a Carnot engine converts heat into work, $Q_2 \to W$, with a loss Q_1. Based on this, let us formulate the following definition:

Definition 5.7. *The efficiency of a cyclic engine functioning between temperatures $T_1 < T_2$ is the quantity $\eta \in (0,1)$ given by*

$$\eta = \frac{W}{Q_2} = \frac{Q_2 - Q_1}{Q_2} = 1 - \frac{Q_1}{Q_2}$$

where Q_2 is the heat absorbed at T_2, from the heat source of the engine, W is the work done, and Q_1 is the loss in the process $Q_2 \to W$, appearing as heat expelled at T_1.

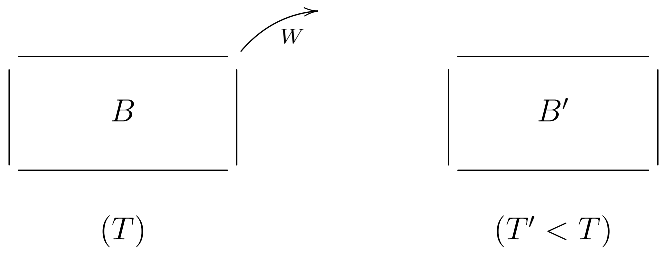
Since we are a bit abstract here, we must comment on the fact that we have indeed $W > 0$ and $Q_2 > Q_1 > 0$. First, we will take $W > 0$ by definition, with the above word “engine” standing for that, that is, “useful engine”. So we are left with $W = Q_2 - Q_1 > 0$, and the problem is whether this really must come from $Q_2 > Q_1 > 0$.

This question might seem a bit futile, but it is not. The second law has certainly something to do with this, because that’s the only principle that we have, preventing heat from flowing from cold to warm, and producing weird things like $Q_1 < 0$, or $Q_2 < 0$. So, let us try now to clarify all this, as our first application of the second law. We have:

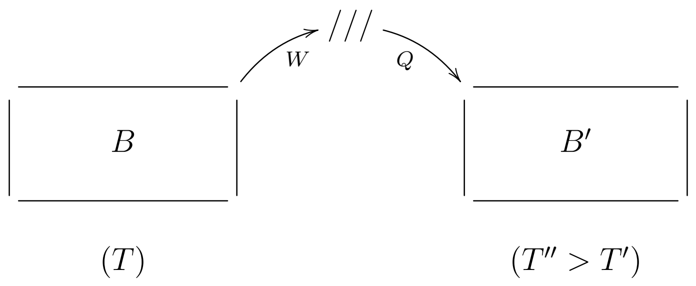
Theorem 5.8. *The second law in Clausius' formulation, Law 5.2, is equivalent to the second law in Lord Kelvin's formulation, Law 5.3.*

Proof. Following Fermi [34], we will do this as mathematicians do, via double implication, and cherry on the cake, by reasoning via negation too:

" $\Longrightarrow$ " Assume that Kelvin is wrong, so that we can extract heat from a body B at temperature T, and convert it into work W. In order to reach to a contradiction with what Clausius says, consider as well a body B', lying nearby, at temperature $T' < T$:

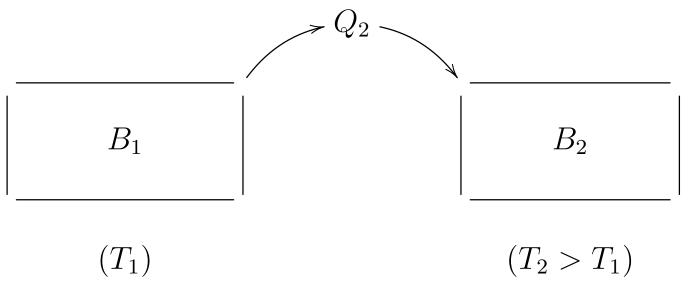


By friction we can convert W into heat, and then heat B' with this heat Q:

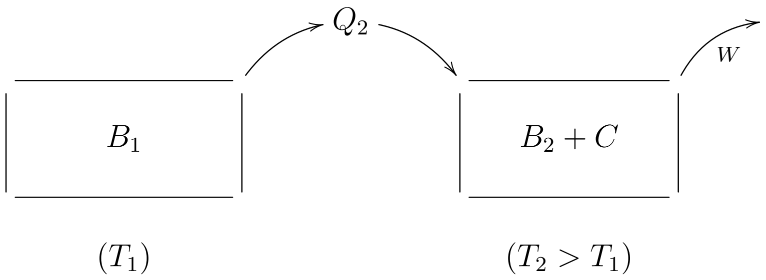


Thus heat has flown from B to B', so Clausius was wrong too. As a conclusion, our assumption that Kelvin is wrong makes no sense, so Kelvin is right, as desired.

"$\Longleftarrow$" Assume now that Clausius is wrong, so that heat Q_2 can flow from a body B_1 at temperature T_1 to a body B_2 of temperature $T_2 > T_1$:



With the help of a Carnot engine C installed on B_2 and functioning in the $[T_1, T_2]$ regime we can convert this heat Q_2 into work W, and with the expelled heat Q_1, at temperature $T_1 < T_2$, not affecting the body B_2:



Thus B_2 has in fact nothing to do with the whole process, which ultimately consists in extracting heat from B_1 and converting it into work W. So Kelvin was wrong too. ☐

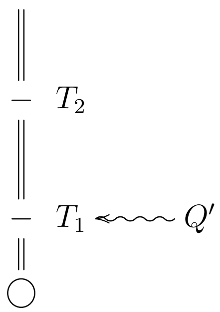
Still following [34], we can now fix the bug in Definition 5.7, as follows:

Proposition 5.9. *In the context of Definition 5.7, assuming that the engine there is a true, useful engine, $W > 0$, we have indeed $Q_2 > Q_1 > 0$.*

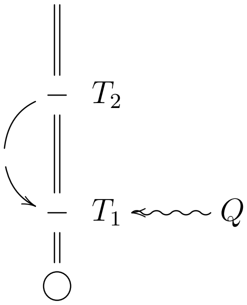
Proof. Since we know that we have $W = Q_2 - Q_1 > 0$, in order to prove the result, we just need to show that we have the following inequality:

$$Q_1 > 0$$

So, assume by contradiction that we have $Q_1 < 0$, say $Q_1 = -Q$ with $Q > 0$. In the context of Definition 5.7, this means that our heat engine functioning in the $[T_1, T_2]$ regime manages to absorb positive heat, $Q' < Q$, at the low temperature T_1:



But then we can install an extra device, letting heat flow from status T_2 to status T_1, during each cycle, as for our machine at status T_1 to absorb exactly $Q > 0$ heat:



And what we have here, in the end, is a machine leaving things at T_1 status unchanged, and making work $W > 0$ out of the source at constant temperature T_2, contradicting Kelvin. Thus, our assumption $Q_1 < 0$ is wrong, and we are done. ☐

Summarizing, everything fine with Definition 5.7, we know now what the efficiency of an engine is. So, let us compare now these engines. We have here:

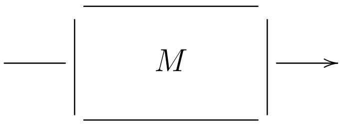
Theorem 5.10. *Among the engines functioning in a given* $[T_1, T_2]$ *regime, the maximum efficiency*

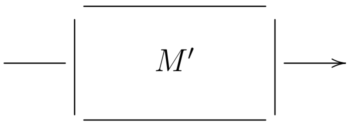
$$\eta = 1 - \frac{Q_1}{Q_2}$$

is acheived by the reversible engines, including the Carnot ones, and with all these reversible engines having the same efficiency.

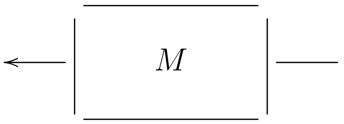
Proof. This is something quite tricky, requiring the use of some clever mathematics. Following as before Fermi [34], the proof goes as follows:

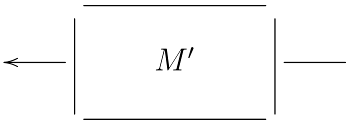
(1) Consider indeed two engines M, M', with the first engine M assumed to be reversible. We will denote these two engines by boxes, as follows:





We will use these engines as well in the reverse sense, as refrigerators:





(2) Our claim is that is enough to prove the following inequality, for our two engines M, M', with the first engine M being assumed, as above, to be reversible:

$$\frac{Q_1}{Q_2} \leq \frac{Q_1'}{Q_2'}$$

Indeed, this certainly proves the first assertion, regarding the maximum efficiency of M. It also proves the last assertion, because assuming that M' was reversible too, we can interchange $M \leftrightarrow M'$, and by double inequality we obtain equality of the above quotients, and so of efficiencies. Finally, the fact that the Carnot engines are reversible is something that we know well, their inverses being refrigerators, as explained in Remark 5.6.

(3) Some math first. We want to prove that we have the following inequality:

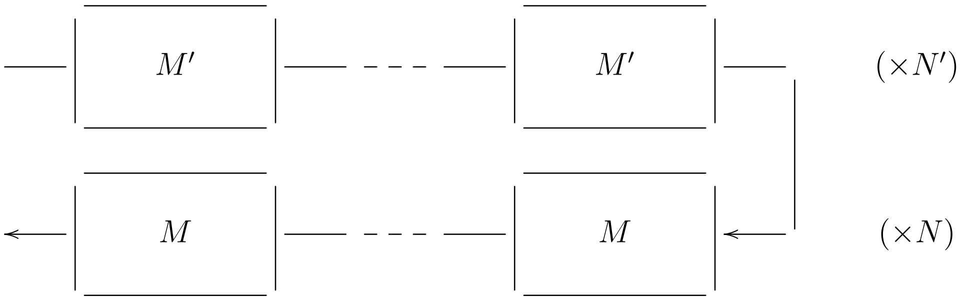
$$\frac{Q_1}{Q_1'} \leq \frac{Q_2}{Q_2'}$$

So, assume that this is wrong, and pick a rational number in between:

$$\frac{Q_1}{Q_1'} > \frac{N'}{N} > \frac{Q_2}{Q_2'}$$

We must come up with a contradiction, out of these inequalities.

(4) In order to do so, we use a trick. Consider the engine $\mathcal{M}$ consisting of N' cycles of M', and N reverse cycles of M, with the picture being as follows:



(5) The data for this new, complex engine $\mathcal{M}$ is then as follows:

$$\mathcal{Q}_1 = N'Q_1' - NQ_1$$

$$\mathcal{Q}_2 = N'Q_2' - NQ_2$$

$$\mathcal{W} = N'W' - NW$$

(6) We have then $\mathcal{W} = \mathcal{Q}_2 - \mathcal{Q}_1$. On the other hand, the inequalities in (3) read:

$$\mathcal{Q}_1 < 0 \quad , \quad \mathcal{Q}_2 > 0$$

Thus, the total work is positive, $\mathcal{W} > 0$, so we have here a true engine, in the sense of Proposition 5.9. But the point now is that, as explained in Proposition 5.9, this should imply $\mathcal{Q}_2 > \mathcal{Q}_1 > 0$. Thus, we have our contradiction, and we are done. $\square$

We have now all the ingredients for stating our main result, as follows:

THEOREM 5.11. *The temperature as measured before, with an arbitrarily calibrated thermometer, can be linearly rescaled into an absolute temperature, which is positive,*

$$T > 0$$

and is given by the following equation, valid for any $T_1 < T_2$*,*

$$\frac{T_1}{T_2} = 1 - \eta$$

where η *is the efficiency of a reversible engine operating in the regime* $[T_1, T_2]$*.*

PROOF. This follows from the theory developed above, and from some supplementary tricks, of the same type, combining engines, the idea being as follows:

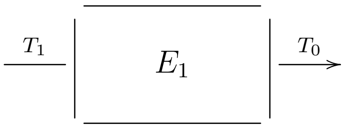
(1) Consider the following function, defined for any temperatures $T_1 < T_2$:

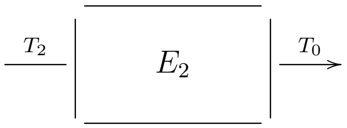
$$f(T_1, T_2) = \frac{1}{1 - \eta(T_1, T_2)} = \frac{Q_2}{Q_1}$$

Our claim, which will basically lead to the result, via some elementary mathematics, is that we have the following formula, valid for any temperatures $T_0 < T_1 < T_2$:

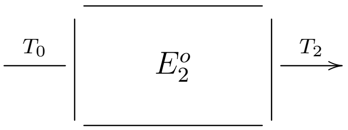
$$f(T_0, T_2) = f(T_0, T_1) f(T_1, T_2)$$

(2) In order to prove this claim, consider a reversible engine E_1 running in the regime $[T_0, T_1]$, and a reversible engine E_2 running in the regime $[T_0, T_2]$:





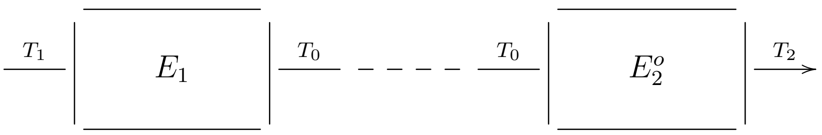
Since E_2 is reversible, we can run it in reverse sense as well, as follows:



(3) We denote by Q_0, Q_1 the amount of heat expelled/absorbed by E_1 at T_0/T_1, and, after suitably normalizing E_2, we denote by Q_0, Q_2 the amount of heat expelled/absorbed by E_2 at T_0/T_2. According to our definition of efficiency, we have the following formulae:

$$f(T_0, T_1) = \frac{Q_1}{Q_0} \quad , \quad f(T_0, T_2) = \frac{Q_2}{Q_0}$$

(4) Consider now an engine E, obtained by running a cycle of E_1, and then a reverse cycle of E_2, denoted as above by E_2^o, with the picture being as follows:



(5) Let us analyze how this complex engine E works at the temperatures T_0, T_1, T_2. According to our assumptions, the situation here is as follows:

– At T_0, no heat is exchanged, Q_0 being absorbed by the reverse E_2.

– At T_1, what happens is that an amount of Q_1 is expelled by E_1.

– At T_2, what happens is that an amount of Q_2 is absorbed by E_2.

(6) We conclude from this that E is a reversible engine running in the regime $[T_1, T_2]$, and so, by our definition of efficiency, we have the following formula:

$$f(T_1, T_2) = \frac{Q_2}{Q_1}$$

Now by comparing with the two formulae obtained at (3), we get, as claimed:

$$f(T_0, T_2) = f(T_0, T_1)f(T_1, T_2)$$

(7) But this tells us that we can find a certain function $\theta = \theta(T)$ such that:

$$f(T_1, T_2) = \frac{\theta(T_2)}{\theta(T_1)}$$

Moreover, this function $\theta = \theta(T)$ must be strictly positive, and increasing.

(8) Getting back now to efficiencies, we have the following formula:

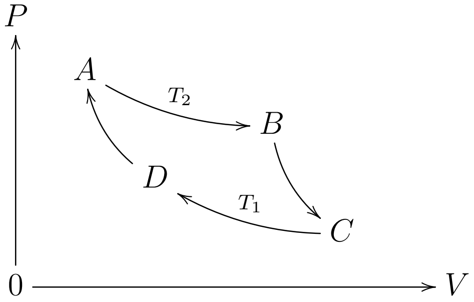
$$\eta(T_1, T_2) = 1 - \frac{1}{f(T_1, T_2)} = 1 - \frac{\theta(T_1)}{\theta(T_2)}$$

The point now is that, for computing θ, we can use this formula, and any reversible engine, of our choice. And by choosing here the Carnot engine operating on $[T_1, T_2]$, and doing the computation, based on the adiabatics for the ideal gases, we obtain:

$$\eta(T_1, T_2) = 1 - \frac{T_1}{T_2}$$

Thus the absolute temperature $\theta = \theta(T) \in (0, \infty)$ constructed in (7) is given by $\theta(T) = \lambda T$, for some $\lambda \in \mathbb{R}$, and we are led to the conclusion in the statement.

(9) To be more precise, and getting now to the details, let us first recall the general functioning scheme of the Carnot engines. This is as follows, with the horizontals AB, CD being the isothermals, and the verticals BC, DA being the adiabatics:



(10) Assuming now that our Carnot engine runs with an ideal gas, and by using the fact, that we know from chapter 4, that the energy of such an ideal gas depends only on the temperature, we have, by looking what happens on the above AB isothermal:

$$Q_2 = L_{AB} = RT_2 \log \frac{V_B}{V_A}$$

Similarly, by looking what happens on the CD isothermal, we have:

$$Q_1 = L_{CD} = RT_1 \log \frac{V_C}{V_D}$$

(11) As a second ingredient now, also as a consequence of the results from chapter 4, since the points A, D lie on an adiabatic, we have the following formula:

$$T_1 V_D^{k-1} = T_2 V_A^{k-1}$$

Similarly, since the points B, C lie on an adiabatic, we have:

$$T_1 V_C^{k-1} = T_2 V_B^{k-1}$$

Now by taking the quotients of these two formulae, we obtain:

$$\frac{V_B}{V_A} = \frac{V_C}{V_D}$$

(12) But with this in hand, we can go back to the two formulae obtained in (10). Since the logarithms on the right are the same, by dividing these formulae, we obtain:

$$\frac{Q_2}{Q_1} = \frac{T_2}{T_1}$$

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

So, that was it, and truly amazing, isn't it. Obviously, having $T > 0$ is something of similar impact to Einstein's $v < c$. And, not only this discovery came before Einstein's, but its proof, explained above, was via a sort of enormous Gedankenexperiment, far more technically involved that the trains and clocks needed for special relativity.

Hommage to the thermodynamics greats, Boyle, Gay-Lussac, Celsius, Joule, Carnot, Otto, Diesel, Clapeyron, Clausius, Lord Kelvin and the others, for this remarkable discovery. More hommage to follow later, to Maxwell, Boltzmann and others, for further discoveries. And even more hommage later, to Planck, and then Bose, Fermi and others, for even further discoveries, such as the difference between bosons and fermions.

In practice now, with Theorem 5.11 in hand, we can do several useful things. First, we can eventually calibrate our thermometer, now after many pages of thermodynamics, a practical definition here, based on the same philosophy as Celsius, being as follows:

Definition 5.12. *We can calibrate our thermometer, as follows:*

(1) *Celsius:* $T = 0$ *is where water freezes, and* $T = 100$ *is where water boils.*
(2) *Kelvin:* $T = 0$ *is the absolute minimum, and the degrees are as in Celsius.*

In practice, this means 0° K $= -273.15$° C. We will use in what follows the Kelvin scale for abstract physics, and Celsius for various engineering matters. We will be also back to all this later on, with various refinements and finer numerics.

Getting back now to the Kelvin degrees, as constructed in Definition 5.12, these are for the moment quite abstract things, and some numerics would certainly help. To start with, here is a table with the melting points of various common materials:

material	melting point
water	273
mercury	234
radon	202
fluorine	200
chlorine	171
xenon	161
krypton	116
argon	83
oxygen	54
nitrogen	63
neon	25
hydrogen	14
helium	1

In relation with this, many interesting things happen around $0 - 1$K, and we will be back to this exciting topic, on regular occasions, in this book.

Along the same lines, let us also mention that engineers nowadays are quite advanced in producing extreme cold, the world record, established in 2021 through matter-wave lensing of rubidium Bose-Einstein condensates, being 38 picokelvin:

$$T = 0.000\ 000\ 000\ 038\ K$$

Finally, as an important remark, all this extreme cold physics is not as fringe as it seems. Indeed, while the universe is mostly known for all sorts of bright and very hot stars, its average temperature is 2.73 K. So, extreme cold is in fact the norm.

There are many interesting things happening at low temperatures, including:

(1) Liquid helium.

(2) Bose-Einstein condensates.

(3) Superconductivity.

(4) And many more.

Regarding the interstellar space, whose temperature is close to 0° K too, many interesting things happen here, often in conjunction with gravity, and with electromagnetism too, and with, thermodynamically, the Maxwell demon roaming around.

5b.

5c.

5d.

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

Interstellar space

6a. Interstellar space

6b.

6c.

6d.

6e. Exercises

Exercises:

Exercise 6.1.

Exercise 6.2.

Exercise 6.3.

Exercise 6.4.

Exercise 6.5.

Exercise 6.6.

Bonus exercise.

CHAPTER 7

The cold jungle

7a. The cold jungle

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

Monsters at rest

8a. Monsters at rest

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

Into the past

Happy people, singing people
Party people, happy people
Jamming on the party session, oh Lord
Sing Hallelujah

CHAPTER 9

Known universe

9a. Known universe

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

Towards the past

10a. Towards the past

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

Some computations

11a. Some computations

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

Goodbye physics

12a. Goodbye physics

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

Into the future

Saturday night, dance, I like the way you move
Pretty baby
It's party time and not one minute we can lose
Be my baby

CHAPTER 13

Time, evolution

13a. Time, evolution

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

Dark worlds

14a. Dark worlds

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

Life cycles

15a. Life cycles

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

The near future

16a. The near future

16b.

16c.

16d.

16e. Exercises

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

Bibliography

[1] A.A. Abrikosov, Fundamentals of the theory of metals, Dover (1988).

[2] A.A. Abrikosov, L.P. Gorkov and I.E. Dzyaloshinski, Methods of quantum field theory in statistical physics, Dover (1963).

[3] G.W. Anderson, A. Guionnet and O. Zeitouni, An introduction to random matrices, Cambridge Univ. Press (2010).

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

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

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

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

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

[9] N.W. Ashcroft and N.D. Mermin, Solid state physics, Saunders College Publ. (1976).

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

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

[12] T. Banica, Principles of thermodynamics (2024).

[13] T. Banica, The joys of relativity theory (2024).

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

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

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

[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] A.R. Choudhuri, Astrophysics for physicists, Cambridge Univ. Press (2012).

[20] D.D. Clayton, Principles of stellar evolution and nucleosynthesis, Univ. of Chicago Press (1968).

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

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

[23] P.A. Davidson, Introduction to magnetohydrodynamics, Cambridge Univ. Press (2001).

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

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

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

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

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

[29] S.K. Donaldson, Riemann surfaces, Oxford Univ. Press (2004).

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

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

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

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

[37] R.P. Feynman, Statistical mechanics: a set of lectures, CRC Press (1988).

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

[39] W. Fulton and J. Harris, Representation theory, Springer (1991).

[40] S. Galtier, Introduction to modern magnetohydrodynamics, Cambridge Univ. Press (2016).

[41] N. Goldenfeld, Lectures on phase transitions and the renormalization group, CRC Press (1992).

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

[43] D.L. Goodstein, States of matter, Dover (1975).

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

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

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

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

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

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

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

[51] A. Hatcher, Algebraic topology, Cambridge Univ. Press (2002).

[52] M.P. Hobson, G.P. Efstathiou and A.N. Lasenby, General relativity, Cambridge Univ. Press (2006).

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

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

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

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

[57] J.E. Humphreys, Introduction to Lie algebras and representation theory, Springer (1972).

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

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

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

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

[62] C. Kittel, Introduction to solid state physics, Wiley (1953).

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

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

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

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

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

[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] P. Petersen, Linear algebra, Springer (2012).

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

[79] B. Ryden and B.M. Peterson, Foundations of astrophysics, Cambridge Univ. Press (2010).

[80] B. Ryden and R.W. Pogge, Interstellar and intergalactic medium, Cambridge Univ. Press (2021).

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

[82] B. Schutz, A first course in general relativity, Cambridge Univ. Press (2009).

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

[84] J. Smit, Introduction to quantum fields on a lattice, Cambridge Univ. Press (2002).

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

[86] A.M. Steane, Relativity made relatively easy, Oxford Univ. Press (2012).

[87] C.H. Taubes, Differential geometry, Oxford Univ. Press (2011).

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

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

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

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

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

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

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

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

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

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

[98] J.M. Yeomans, Statistical mechanics of phase transitions, Oxford Univ. Press (1992).

[99] J. Zinn-Justin, Phase transitions and renormalization group, Oxford Univ. Press (2005).

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

Index