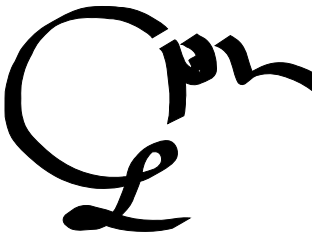
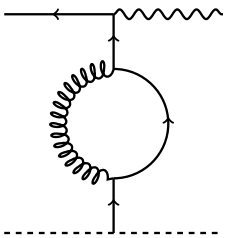
Preparatory Notes for the Aotearoa Fundamental Physics Summer School 2016



Sebastian Schuster

11 December 2016



These notes are supposed to provide people intending to go to the summer school with some useful background material and references. The intention is not to be intimidating, or that you have to know this material inside out – but you should feel perfectly comfortable to go through the material if the need arises. This should not be too hard as many of the topics discussed here should be part of undergraduate physics courses. That being said, if something from the list of contents or on the list given on the website seems to be unknown, please have a look at these notes to see if you can learn it. As always, if you are stuck it will be a good idea to have a quick look at the references given below. So think of these notes as a mix of cheat sheet, summary, introduction and self-test. Enjoy[1] !

Contents

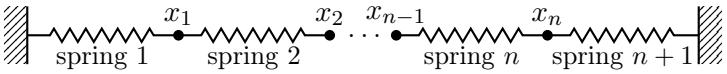
[1] Whoever finds errors may keep them.[2]

[2] Of course not. In all seriousness – send them to me (nice and friendly Sebastian) at `sebastian.schuster@msor.vuw.ac.nz`. I will then try to update these notes on the web page `http://sms.victoria.ac.nz/Events/AotearoaFP16/Prerequisites`.

1 Mathematics

1.1 Normal Modes – A Playground for Linear Algebra

Take a setup of n masses, all of mass m, coupled by $n+1$ springs[3], each with the same spring constant k, just as pictured below. Let us call their displacement from their equilibrium positions x_j, meaning $x_i = 0$ corresponds to the i-th equilibrium position. And the resulting, natural boundary conditions for the walls (also giving the names of the walls' coordinates) is $x_0 = x_{n+1} = 0$. The distance between two neighbouring equilibria shall be a – and independent of the mass considered.



[3]It may seem odd at first to call this a 'mathematics problem', but despite the physical setup, the discussion makes heaps of use of things like Jordan decompositions, eigenvalues and eigenvectors, orthogonality, . . . – so I find it fair to call it a mathematics problem.

An obvious question[4] is how these masses will move if they were to be moved out of equilibrium. Here is how one would go about this:

1. Write down Newton's law for each point. You should get something along the lines of
$$m\ddot{x}_j = kx_{j-1} - 2kx_j + kx_{j+1} \tag{1}$$
or something involving a matrix
$$\begin{pmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}.$$

2. To solve this set of ODE, it is a perfect opportunity to recap 'normal modes'. A more mathematical way of phrasing it is: We want to find a different set of coordinates (other than the already cleverly chosen x_j) such that the e.o.m. (equations of motion) become simpler. This is where the matrix picture comes in! This matrix can be diagonalized – think about it, try to figure out why! – and therefore provides new coordinates. So, the task now is to find the eigenvalues and eigenvectors of the matrix. These then are the normal frequencies and normal modes of the matrix.

   For example, try to guess a solution of the kind
$$x_j = A_j e^{i\omega t}.$$
   Insert in (1) and try to compare the result with what you know from a Jordan decomposition.

3. To actually *get* the normal frequencies can still look a bit daunting. So, forget – for a moment – about the boundary conditions $x_0 = x_{n+1}$. Now, the matrix will be infinitely large[5] and look like so:
$$\begin{pmatrix} \ddots & \ddots & \ddots & & \\ & 1 & -2 & 1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$
   Assume that the solution will take on the form of a travelling wave:
$$x_j(t,k) = C(k)e^{i(\ell ka - \omega(k)t)}, \tag{2}$$

[4] To anyone who likes physics...

[5] But part of the summer school will intend to get rid of fears of infinities, so buckle up!

where k is a wave number, and ℓ is an integer. Now: *Check that only frequencies $\omega(k)$ will provide a solution of the form (2) if they fulfil the dispersion relation*

$$\omega(k) = 2\sqrt{\frac{k}{m}}\sin\frac{ka}{2}. \tag{3}$$

4. Now, back to the finite chain – what will change for k and ℓ if we reintroduce the boundary conditions $x_0 = x_{n+1}$? Hint: Think standing wave! The allowed (and sought for!) frequencies should change to

$$\omega_\ell = 2\sqrt{\frac{k}{m}}\sin\left(\frac{\pi\ell}{2(n+1)}\right), \qquad \ell \in \{1,\ldots,n\}.$$

5. What does $x_j(t)$ look like now? Use that to find that the normal modes are, up to normalization constants, given by

$$a_\ell = C_\ell \begin{pmatrix} \sin\left(\frac{1}{n+1}\ell\pi\right) \\ \sin\left(\frac{2}{n+1}\ell\pi\right) \\ \vdots \\ \sin\left(\frac{n-1}{n+1}\ell\pi\right) \\ \sin\left(\frac{n}{n+1}\ell\pi\right) \end{pmatrix}.$$

6. That finally solves it. Unless you want to find the C_ℓ – for that, first check that these a_ℓ are orthogonal for different values of ℓ. Then demanding ortho*normality* you will find that

$$C_\ell = \sqrt{\frac{2}{n}}.$$

One (of the many) references you can use for looking this up, would be [JS12], p.187ff – at least for the physics side. Also the literature on numerical mathematics (e.g. [PTVF07], p.1024ff) will have plenty to say, this time looking at it as a discretization of the 1D wave equation $-u_{tt} + c^2 u_{xx} = 0$. Feel free to get that discretization yourself...

1.2 'Functional Analysis' – Bits and Pieces of Distributions

This section is more a summary of results then asking you to re-derive things or calculate things. Mostly. But some of it will be useful later in the section on complex analysis, so I will collect some things here.

1.2.1 General Things and Notation

First a bit on notation: While mathematicians usually like to highlight a distribution as a linear map from some function space to some field (normally either $\mathbb{R}$ or $\mathbb{C}$)[6], physicists

[6] And as a linear functional it therefore is contained in the *dual space* of the given function space.

prefer to view them as 'functions' with a bit weirder properties than normal functions. For example, what a physicist would call $\delta(x-x_0)$ (and demand a property under integration), a mathematician would call δ_{x_0} followed by the function it is acting on, e.g. $\delta_{x_0}[f]$. To get that function f into the picture, the physicist would write $\int_{\mathbb{R}} \delta(x-x_0)f(x)\mathrm{d}x$. What one prefers is personal taste – and sometimes different problems may be clearer in one notation or the other. For that specific, latter reason, I introduce the Kramers-Kronig relations in 1.3.5 in the mathematical notation. However, I will give the result in both notations and for the summer school we will be concerned (probably) only with the physicists' notation.

Another short note regarding language: Physicists look mainly at distributions as something coming with an integral that integrates some way or the other over the function the distribution is acting on. The bits inside the integral are called the *integral kernel*. The language is borrowed from integral transforms and while useful, becomes a bit subtle with the δ-distribution/function of Dirac. After all, it is far from apparent that the definition as given by the mathematicians (see below) gives rise to the thing (the kernel) that physicists love to throw around in a well-defined way[7]. So, in general this means that a distribution K acting on f has a kernel $K(x)$ if it can be written as:

$$K[f] = \int_{-\infty}^{\infty} K(x)f(x)\mathrm{d}x. \tag{4}$$

Common examples in physics are the Dirac δ, propagators, the exponential e^{ikx} of a Fourier transform, e^{-kx} for a Laplace transform, and many more...

1.2.2 Dirac's δ

Second, let us collect a few results for the Dirac delta function/distribution[8].

1. The definition, both notations:

$$\delta(x) \text{ s.t. } f(x_0) = \int_{-\infty}^{\infty} \delta(x-x_0)f(x)\mathrm{d}x, \qquad \delta_{x_0}[f] := f(x_0) \tag{5}$$

2. A straightforward result is what its Fourier transform[9] (see 1.3.4 below) is:

$$\mathcal{F}[\delta_{x_0}](k) = \int_{-\infty}^{\infty} \delta(x-x_0)e^{-ikx}\mathrm{d}x = e^{-ikx_0}. \tag{6}$$

   Put differently:

$$\int_{-\infty}^{\infty} e^{i(k-k_0)x}\mathrm{d}x = 2\pi\delta(k-k_0).$$

[7] Indeed, the Schwarz kernel theorem provides a justification for physicists' preference to write all kinds of things with a kernel – but let's not worry about these technicalities.

[8] It has plenty of names. Many more if one slightly changes the area of mathematics it is encountered in...

[9] Technically, one would have to be very careful in extending the Fourier transform's domain – after all, normally Fourier transforms are only defined for L^2 functions/Schwartz functions/...

Note the factor of 2π – this is due to our convention for the Fourier transform, again, see 1.3.4.

3. Obviously, we cannot expect to always have these nice arguments $(x - x_0)$ for δ. So it is important to (at least) know that

$$\delta(g(x)) = \sum_{\text{Zeroes } x_i \text{ of } g(x)} \frac{1}{|g'(x_i)|}\delta(x - x_i), \tag{7}$$

where it is important that $g(x)$ has only simple zeroes.

4. It is also possible to differentiate the δ itself. The idea is that you would use partial integration[10] to make sense of it. Neglecting boundary terms[11], of course. We get:

$$\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}\delta(x - x_0)\right) =: \delta^{(n)}(x - x_0) \text{ s.t. } (-1)^n \int_{-\infty}^{\infty} \delta^{(n)}(x - x_0)f(x)\mathrm{d}x = f^{(n)}(x_0). \tag{8}$$

Note that I will carefully distinguish between δ^n and $\delta^{(n)}$ to keep the n-dimensional distribution separated from its n-th derivative. Again, the mathematical notation would be a bit shorter:

$$\delta^{(n)}_{x_0}[f] := (-1)^n f^{(n)}(x_0).$$

5. The next step is to generalize this to n dimensions. We get:

$$\delta^n(\mathbf{x}) \text{ s.t. } f(\mathbf{x}_0) = \int_{\mathbb{R}^n} \delta^n(\mathbf{x} - \mathbf{x}_0)f(\mathbf{x})\mathrm{d}^n\mathbf{x}, \qquad \delta^n_{\mathbf{x}_0}[f] := f(\mathbf{x}_0), \tag{9}$$

$$\mathcal{F}[\delta^n_{\mathbf{x}_0}](\mathbf{k}) = \int_{\mathbb{R}^n} \delta^n(\mathbf{x} - \mathbf{x}_0)e^{-i\mathbf{k}\cdot\mathbf{x}}\mathrm{d}^n\mathbf{x} = e^{-i\mathbf{k}\cdot\mathbf{x}_0}, \tag{10}$$

$$\int_{\mathbb{R}^n} e^{i(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{x}}\mathrm{d}^n\mathbf{x} = (2\pi)^n\delta^n(\mathbf{k} - \mathbf{k}_0), \tag{11}$$

$$\delta^n(g(\mathbf{x})) = \sum_{\mathbf{x}_i \in g^{-1}(\mathbf{0})} \frac{1}{|J_g(\mathbf{x}_i)|}\delta^n(\mathbf{x} - \mathbf{x}_i), \tag{12}$$

where, now, $g : \mathbb{R}^n \to \mathbb{R}^n$ and J_g denotes the Jacobian determinant of g.

A good, though very mathematical and technical book on this would be [DK10]. More instrumentalist's approaches can be found in most physics/mathematics books that have to cover this. Examples would be: [Sha14], [Jac75], [Gri08], [Gri12], [SG10], [Zor04] or [BD97].

[10] Oh, thou lovely standard trick...

[11] Yet another standard trick. Here it means we are taking care with the domain (as a function space) of δ.

1.2.3 Green's Functions

Now that we know a bit about the δ-distribution, we can recap one of the most powerful tools for (linear[12]) partial differential equations.

Assume you have some linear partial differential operator D. If you want to solve an inhomogeneous equation of the type

$$Df = g \tag{13}$$

for f this looks daunting at first. But Green's functions[13] – if known – will tremendously help with this. So, what is the definition of a Green's function?

Definition 1. *A* Green's function $G(x,y)$ *for the linear partial differential operator* D *is a solution to the following inhomogeneous PDE:*

$$D_x G(x,x_0) = \delta(x - x_0). \tag{14}$$

In the following discussion I will assume that $G(x,y)$ can be written as $G(x-y)$ – which is the case for translational invariant (constant coefficient) D.

The trick now is to realize that a convolution[14] of our initial boundary condition g with this Green's function will guarantee a solution. In the following calculation of this I will try to be cavalier about the notation.

$$\begin{aligned}(D_x g * G)(x) &= D \int g(y)G(x-y)\mathrm{d}y &(15)\\ &= \int g(y)D_x G(x-y)\mathrm{d}y &(16)\\ &= \int g(y)\delta(x-y)\mathrm{d}y &(17)\\ &= g(x) &(18)\end{aligned}$$

Voilà[15]. Just what we wanted. A physical point of view on this is the following: Since our operator is linear, we expect it to fulfil some kind of superposition principle. And that is just what the Green's function provides us with. A physicist's interpretation of the mathematics going on might be to interpret D as a $\infty \times \infty$-matrix and $G(x,y)$ ends up as its 'matrix inverse'. You also could discretize D and then define a finite-dimensional

[12] In the summer school you will see how one can push this into uncomfortably non-linear regimes...

[13] There is some notational debate whether or not to include the apostrophe and s – so sometimes you will find them as 'Green functions'. At least, 'Green's functions' are oddballs when compared to Bessel functions, Taylor series or Hessian matrices.

[14] Remember that the convolution $f * g$ of two functions f and g (again forgetting about domains and function spaces for the moment) is defined as: $(f*g)(x) := \int_{-\infty}^{\infty} f(y)g(x-y)\mathrm{d}y$.

[15] A proof of this in the more general case when $G(x,y)$ cannot be written as $G(x-y)$ can be found at `=https://en.wikipedia.org/wiki/Fundamental_solution#Proof_that_the_convolution_is_a_solution`. Note that depending on the used literature, there will be slight differences in the definition of a 'fundamental solution' as used in this link and a Green's function, e.g. [Eva10].

matrix representation and consider $G(x,y)$ the corresponding limit. The latter can be done mathematically precise though it will be unnecessarily cumbersome (to a physicist).

Now that we have the general *use* of Green's functions, let's have a look at specific ones for the Laplace equation in different dimensions:

Dimension	D	Green's function G
1D:	∂_x^2	$x\Theta(x)$
2D:	$\partial_x^2 + \partial_y^2$	$\frac{1}{2\pi}\ln(\sqrt{x^2+y^2})$
3D:	$\partial_x^2 + \partial_y^2 + \partial_z^2$	$-\frac{1}{4\pi\sqrt{x^2+y^2+z^2}}$
nD ($n \geq 3$):	$\partial_i\partial^i$	$-\frac{1}{n(n-2)V(B_1^n)}\frac{1}{(x_i x^i)^{(n-2)/2}}$

$V(B_1^n)$ stands for the volume of the n-dimensional unit ball. The notation used in the last line is the Einstein sum convention. As we are working in Euclidean space, we do not have to worry about sign entering from lowering or raising indices. If you forgot how that works, look it up – Wikipedia[16] should be already sufficient, or any of the references that might go in that direction. Generally the most likely sources are books dealing with differential equations (e.g. [Hab04], [BD97], [Tay11]), electrodynamics (e.g. [Jac75]) or relativistic quantum theories (e.g. [Sre11] makes heavy use of it).

A last short note: Mathematicians usually like to define the Laplacian Δ as $-\partial_i\partial^i$ to have a positive spectrum. Physicists don't. So be prepared to occasionally find differing overall signs for things like Green's functions[17].

1.3 Complex Analysis

Here, I will follow mostly what I can find in either [DK10], [SG10] or [SLSS09]. Especially for sections 1.3.2 and 1.3.3 I suggest looking up your favourite exercise book and do a few of the corresponding calculations of residues, real integrals and the like. [SLSS09], for example, has a large amount of exercises for this.

1.3.1 Riemann-Lebesgue Lemma

The Riemann-Lebesgue lemma/theorem reads:

Theorem 1. *If $f(x)$ is Lebesgue integrable[18], its Fourier transform $\mathcal{F}[f](k)$ will be, too.*

[16] For example: `https://en.wikipedia.org/wiki/Einstein_notation`. And some of the Green's functions given in the table are adapted from `https://en.wikipedia.org/wiki/Green%27s_function#Table_of_Green.27s_functions`.

[17] In the wild. The summer school will follow physicists' convention.

[18] This technicality is actually important – take for example $f(x) = \sin(x^2)$, as in [SG10], p.785f.

In particular this means that

$$\lim_{k\to\infty} \mathcal{F}[f](k) = 0. \tag{19}$$

The physicist's interpretation is that the integrand in the Fourier transform oscillates more and more rapidly and as a result averages out.

1.3.2 Cauchy's Theorem, Residue Theorem and Co.

- **Theorem 2** (Cauchy's integral theorem)**.** *For U simply connected, $f : U \to \mathbb{C}$ holomorphic and γ a rectifiable, closed curve in U:*

$$\oint_\gamma f(z)\mathrm{d}z = 0. \tag{20}$$

- **Theorem 3** (Cauchy's integral formula)**.** *For U simply connected, $f : U \to \mathbb{C}$ holomorphic and γ a rectifiable, closed curve in U with winding number 1 around $a \in U$:*

$$f^{(n)}(a) = \frac{n!}{2\pi i} \oint_\gamma \frac{f(z)}{(z-a)^{n+1}}\mathrm{d}z. \tag{21}$$

- **Definition 2** (Residue of a function f)**.** *For finite a being a pole of order n of f, the residue of f at a is given by*

$$\mathrm{Res}(f,a) = \frac{1}{(n-1)!} \lim_{z\to a} \frac{\mathrm{d}^{n-1}}{\mathrm{d}z^{n-1}} \left((z-a)^n f(z)\right). \tag{22}$$

At infinity, the residue is given by

$$\mathrm{Res}(f,\infty) = -\mathrm{Res}\left(\frac{1}{z^2} f(1/z),0\right). \tag{23}$$

- **Theorem 4** (Residue theorem)**.** *For a positively oriented, simple closed curve with open interior U and f holomorphic on $U \setminus \{a_1,\ldots,a_n\}$:*

$$\oint_\gamma f(z)\mathrm{d}z = 2\pi i \sum_{i=1}^{n} \mathrm{Res}(f,a_i). \tag{24}$$

For poles at finite z and a function possible to write as a Laurent series around the pole, one can also take the coefficient of the (-1)-th term of the Laurent series.

It's good and easy enough to check the residue theorem by doing it for $\oint_{S^1} \frac{1}{z}\mathrm{d}z$.

1.3.3 Calculating Real Integrals with Complex Analysis

Two theorems helping in evaluating integrals[19] will be given here, based on what can be found in [SLSS09]. That's also, where you can find examples and exercises for this kind of thing. Also, I will describe a few typical functions giving off warning signs that you will end up having to discuss branch cuts.

- Let H_R be a stand-in for a half-circle of radius R in either the upper or the lower complex half-plane, closed along the real line.
- **Theorem 5.** *Let $\alpha > 1$ and $M > 0$. If*

$$|f(z)| \leq \frac{M}{R^\alpha} \quad \textit{for} \quad z = Re^{i\theta},\ \theta \in \mathbb{R} \tag{25}$$

then

$$\lim_{R\to\infty} \oint_{H_R} f(z)\mathrm{d}z = 0 \tag{26}$$

- **Theorem 6.** *Under the same conditions:*

$$\lim_{R\to\infty} \oint_{H_R} e^{ikz} f(z)\mathrm{d}z = 0 \tag{27}$$

Now for the branch cuts: *If* there are branch cuts, contours for using residue calculus/complex analysis have to be carefully chosen. For example, if a contour hits a branch cut, go – a distance ε away from the branch cut – all the way d to the nearest branching point, go around on a semicircle of radius ε and go back your distance d on the other side of the branch cut. *Then* you can finally continue with your original contour. Branch cuts make life difficult, so it is good to recognize when they happen:

- Roots $\sqrt[n]{z}$
- Logarithms; $\ln z$ for example has it along the non-positive real axis[21].
- As many integrals of inverse trig functions ($\arccos, \arcsin, \arctan, \dots$) have a relationship to the complex root (e.g. $\arccos, \arcsin$) or logarithm (e.g. $\arctan$), they do have, too.

[19] And series. Let us forget about series.[20]

[20] Here. In this notes. Maybe it will be useful. Who knows. One never knows. And you cannot know too much.

[21] Unless you want to use really unusual choices where to do the needed branch cut. There is a small amount of freedom if you want to mess with people and make your calculation obscure.

- As hyperbolic $(\cosh, \sinh, \dots)$ functions are trig functions with complex argument, their inverses[22] have them.
- More generally, functions as arguments of ln or $\sqrt[n]{\cdot}$ make things a lot more complicated.
- Example 1: $\ln\left(\frac{z+1}{z-1}\right)$ has a branch cut between -1 and 1
- Example 2: $\sqrt{z^2+1}$ has two branch points with separate cuts extending to infinity.

A **really** helpful example for this would be to calculate $\int_{-\infty}^{\infty} \frac{1}{-z^2+a^2 \mp i\varepsilon} \mathrm{d}z$, with our $\lim_{R\to\infty} H_R$ trick. This is not necessarily a good example to get started, so bear in mind my suggestion at the beginning of section 1.3.

1.3.4 Fourier Transformations

Let's summarize a few results on Fourier transforms!

- The Fourier transform $\mathcal{F}$ turns a function $f(x)$ into a function[23] $\hat{f}(k)$ according to the following scheme:

$$\mathcal{F}[f](k) = \hat{f}(k) := \int_{-\infty}^{\infty} e^{-ikx} f(x) \mathrm{d}x. \tag{28}$$

- It has an inverse $\mathcal{F}^{-1}$ turning $\hat{f}(k)$ back into $f(x)$ (under omission of technicalities) given by

$$\mathcal{F}[\hat{f}](x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} \hat{f}(k) \mathrm{d}k. \tag{29}$$

  If we were to generalize to higher dimensions n, the prefactor would have to have the dimension n as a power. You can see an example of this in section 1.2.2 for the Dirac δ.

- Gaussians are fixed points of $\mathcal{F}$:

$$\mathcal{F}[\exp(-\sigma x^2)](k) = \sqrt{\frac{\pi}{\sigma}} \exp\left(-\frac{k^2}{4\sigma}\right), \tag{30}$$

$$\text{i.e. Gaussian } \overset{\mathcal{F}}{\mapsto} \text{ Gaussian.} \tag{31}$$

[22] *I* call them arsinh, artanh, ... and so on, because they have nothing to do with 'arcs', but with 'areas'. This seems to be old-fashioned and slowly dying out. It's what I was taught and what my Latin knowledge makes me certain of. And Wikipedia agrees with me, see `https://en.wikipedia.org/wiki/Inverse_hyperbolic_function#Notation`. But as you can see, I need a lot of justification for this – at least 'down under' Latin is a minority background. All CAS I know of use arcsinh, arctanh etc. Bear with me here, I like being a special snowflake.

[23] At least in these notes it should be simple to distinguish a hat for 'operator' and a hat for 'Fourier transformed'.

- Taking the Fourier transform of a shifted function $f(x - x_0)$ introduces a phase factor.
- For complex conjugation we have

$$\mathcal{F}[\overline{f}](k) = \overline{\hat{f}(k)} \tag{32}$$

- The Fourier transform evaluated at $k = 0$ is the integral of the transformed function – this illustrates the need to put a lot of effort into characterizing the domain of $\mathcal{F}$. These notes will not cover this function space technicality with few exceptions, like this mentioning of it.
- Derivatives are turned into polynomials:

$$\mathcal{F}[\partial_x f](k) = ik\hat{f}(k). \tag{33}$$

  This can be used to turn differential equations into polynomial equations (as with Laplace transforms) and plays a role when defining fractional derivatives or more general pseudo-differential operators[24].
- Convolutions are turned into products:

$$\mathcal{F}[f * g](k) = \hat{f}(k) \cdot \hat{g}(k). \tag{34}$$

- Hermite functions form an orthonormal basis (for appropriate function spaces, here: $L^2(\mathbb{R})$) of eigenfunctions of $\mathcal{F}$.
- **Theorem 7** (Plancherel's Theorem)**.** *The Fourier transform is unitary and an isometry in* $L^2(\mathbb{R})$*:*

$$\int_{-\infty}^{\infty} f(x)\bar{g}(x)\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(k)\overline{\hat{g}}(k)\mathrm{d}k. \tag{35}$$

- If we call

$$\langle h\rangle_{0,\xi} = \int_{-\infty}^{\infty} h(\xi)\xi^2\bar{h}(\xi)\mathrm{d}\xi \tag{36}$$

  the dispersion of $h(\xi)$ around 0, we can formulate an uncertainty principle:

$$\langle h\rangle_{0,x}\langle\mathcal{F}[h]\rangle_{0,k} \geq \frac{1}{4}. \tag{37}$$

[24] A good introduction to this can be found in the first sections of chapter 3 in [LM94]. The rest of the book can be quite scary, but this introduction is really nice. Also, I love the abbreviation ΨDO for 'pseudo-differential operator'.

This can be generalized by defining

$$\langle h\rangle_{a,\xi} = \int_{-\infty}^{\infty} h(\xi)(\xi - a)^2 \bar{h}(\xi)\mathrm{d}\xi, \tag{38}$$

and if we then take the average $\bar{x}$ to be a, we can link this to the familiar

$$\langle x^2\rangle = \mathrm{Var}(x) + \bar{x}^2. \tag{39}$$

Apart from the physical constant $\hbar$, this is the source of Heisenberg's uncertainty principle. So from a mathematical point of view this principle is just an aspect of x and k being Fourier pairs. Similar results hold in signal processing, electrodynamics and any other physical theory using Fourier pairs! A proof (with a different definition of the Fourier transform!) can be found in [Pin02], p.131ff.

- If we use δ-distributions we can link the Fourier transform back to Fourier series – therefore it makes sense of using the same language in both formalisms.
- As already mentioned in 1.2.2, there exist differing conventions regarding the factors of $1/2\pi$ and how to distribute them among $\mathcal{F}$ and $\mathcal{F}^{-1}$. So, if you look things up involving Fourier transforms, make sure you know the convention!
- As an example of the last point *and* a table of functions and their Fourier transforms, have a look at `https://en.wikipedia.org/wiki/Fourier_transform#Tables_of_important_Fourier_transforms`.

Possible places to look at for more information regarding Fourier transforms include, besides the already linked Wikipedia article, a long list of books on analysis (e.g. [Zor04], [SG10], [Lan05]), complex analysis (e.g. [Lan99]), distribution theory (e.g. [DK10]), differential equations (e.g. [Eva10], [Hab04], [Tay11]), electrodynamics (e.g. [Jac75][25]), quantum mechanics (e.g. [Gri12], [CTDL77]) and much, much more. My examples are just what I could find – there's definitely more and probably better books for each subject area listed, at least for introductions to/summaries of Fourier transforms.

1.3.5 Cauchy Principal Value and Kramers-Kronig Relations

The Kramers-Kronig relations[26] link the Cauchy principal value distribution $\mathcal{P}[f]$ and the Dirac distribution $\delta_x[f]$. The Cauchy principal value distribution comes in many, many notations. Examples according to Wikipedia are: $\mathrm{PV}, \mathrm{P}, \mathrm{p.v.}, \mathcal{P}, \fint$ and more.

[25] Not necessarily a *good* reference for Fourier transforms.

[26] This goes by many names and which you have heard first strongly depends on your background. In physics, the most common one would be Kramers-Kronig relations, though mathematicians might have heard it as the Sokhotski-Plemelj[27]theorem.

[27] And depending on the language or age of your literature, the transcriptions will be all over the show. Depending on the literature used the different names are sometimes also associated with slightly different variations of the same result.

First, a bit more on the Cauchy principal value. For **real integrals** with an integrand $f(x)$ that *diverges* at c one can sometimes find a value for it by evaluating instead of

$$\int_a^b f(x)\mathrm{d}x$$

the integral

$$\int_a^{c-\varepsilon} f(x)\mathrm{d}x + \int_{c+\varepsilon}^b f(x)\mathrm{d}x. \tag{40}$$

The Cauchy principal value is then defined as

$$\mathcal{P}\int_a^b f(x)\mathrm{d}x := \lim_{\varepsilon\to 0^+}\int_a^{c-\varepsilon} f(x)\mathrm{d}x + \int_{c+\varepsilon}^b f(x)\mathrm{d}x. \tag{41}$$

For singularities at infinity one uses $1/\varepsilon$, for several singularities one separates the integral into more pieces. Important is to keep ε symmetric around the singularity.

In the complex plane, then, one exchanges the 2ε-interval by a ball of radius ε. So for a contour C with a pole of the integrand at z_0 on it, do the following:

1. The contour C is the boundary of a subset $S \subset \mathbb{C}$.
2. Take the ball of radius ε centred at z_0, let's call it $B_\varepsilon(z_0)$.
3. Instead of $\int_C = \int_{\partial S}$ use $\int_{\partial S\setminus B_\varepsilon(z_0)}$.
4. Take the limit of this integral. That is the Cauchy principal value.

We shall need the **distributional version** of the Cauchy principal value. Here, rather than $\mathcal{P}$, let's call the Cauchy principal value $\mathcal{P}(\frac{1}{x})$. Why? Because the idea is to take the singularity from $1/x$ and see what it does to a function. As a formula:

$$\mathcal{P}\left(\frac{1}{x}\right)[f] := \mathcal{P}\int_{-\infty}^{\infty}\frac{f(x)}{x}\mathrm{d}x. \tag{42}$$

The Cauchy principal value is then taken by cutting out a 'ball' of radius ε at zero.

So, now for the Kramers-Kronig relations:

Kramers-Kronig relations

For 'sufficiently nice' functions $f(x)$ we have

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty}\frac{f(x)}{x\mp i\varepsilon}\mathrm{d}x=\mathcal{P}\int_{-\infty}^{\infty}\frac{f(x)}{x}\mathrm{d}x\pm i\pi f(0). \tag{43}$$

It is common to rewrite this as

$$\int_{-\infty}^{\infty}\frac{f(x)}{x\mp i\varepsilon}\mathrm{d}x=\mathcal{P}\int_{-\infty}^{\infty}\frac{f(x)}{x}\mathrm{d}x\pm i\pi f(0)$$

or even

$$\int_{-\infty}^{\infty}\frac{f(x)}{x+i0^{\mp}}\mathrm{d}x=\mathcal{P}\int_{-\infty}^{\infty}\frac{f(x)}{x}\mathrm{d}x\pm i\pi f(0).$$

It is easy to shift this from x to $x-a$ and furthermore to realize that this is a statement about distributions acting on $f(x)$, so that we can summarize the result in the shorter formula

$$\frac{1}{(x-a)\mp i\varepsilon}=\mathcal{P}\left(\frac{1}{x-a}\right)\pm i\pi\delta(x-a) \tag{44}$$

$$=\frac{x-a}{(x-a)^2+\varepsilon^2}\pm\frac{i\varepsilon}{(x-a)^2+\varepsilon^2}. \tag{45}$$

A proof of (variations of) this can be found in [Jac75], [DK10] (several different proofs throughout the book), [SG10] or [Sha14] and many more books.

2 Physics

2.1 Classical (non-relativistic) Mechanics

2.1.1 Harmonic Oscillator(s) – Lagrangian and Hamiltonian Mechanics

1. Find the Lagrangian of a single harmonic oscillator.
2. Write down the Lagrangian for n independent harmonic oscillators.
3. Write down the Lagrangian for the problem at the very beginning of the mathematics section, 1.1.
4. Go through 1.1 again and try to find the connection to the Lagrangian of n independent harmonic oscillators. Remembering canonical transformations and the reason for the name 'normal mode' should give it already away...
5. Remind yourself what a Legendre transformation does and what condition needs to be fulfilled for it to work. Then, find the corresponding Hamiltonian for the Lagrangian of a single harmonic oscillator. What should[28] be checked before blindly

[28] Theoretically, at least, though it isn't always in practice...

performing a Legendre transformation?

6. Find a Hamiltonian for n independent harmonic oscillators.

7. Feel free to check how solving the ODEs changes, depending on whether you use the Hamiltonian or the Lagrangian approach. I.e., remember the trade-off between order of ODE and number of separate ODEs. Of course, the Lagrangian/Hamiltonian from the first question here will be the easiest and quickest approach.

Fun with simple Hamiltonians

If you want to see how much you can mangle a given Hamiltonian just with canonical transformations, have a look at the example from [JS12], p.246f. Using 'ladder coordinates'[a]

$$Q = \frac{m\omega q + ip}{\sqrt{2m\omega}}, \qquad P = iQ^*$$

and after some work to find a fitting generating function (here: $F(q,Q)$) for this particular canonical transformation, you will get the new Hamiltonian

$$\tilde{H} = -i\omega QP.$$

So why did I say 'mangle'? Try finding a corresponding Lagrangian in these canonical coordinates.....just looking at the Hessian should tell you whether that will work or not.

[a] After all, this is the classical mechanics part – not the quantum mechanics part, where no-one would be surprised to see ladder *operators*...

2.1.2 Classical Perturbation Theory

A harmonic oscillator of mass m has the potential energy

$$V = \frac{k}{2}x^2, \tag{46}$$

with k being the spring constant. Let's change this to include a small, cubic correction term:

$$V = \frac{k}{2}x^2 - \varepsilon C x^3, \tag{47}$$

where C takes care of units. The discussion of this small part I will base on [Kuy03] – I apologize for using a German reference, but [JS12] is my only English book on classical mechanics in 'classical' presentation[29]. More or less every classical mechanics book will discuss polynomial perturbations of the harmonic oscillator in some form, so I hope you will be able to find something in your favourite book that is easily adapted to the case

[29] Arnold's classical mechanics book will not help here...

of a cubic potential. In the case of [JS12], for example, you can find a discussion of the quartic potential in several sections.

In this section we will skip Lagrangian and Hamiltonian mechanics and start straight from Newtons second law, using

$$\mathbf{F} = -\nabla V.$$

1. From the things given above, derive the equations of motion for the systems as

$$\ddot{x} + \frac{k}{m}x - \varepsilon\frac{3C}{m}x^2 = 0. \tag{48}$$

2. Suppose, we can write our system like so:

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \dots \tag{49}$$

   This will be the starting point of our perturbation theory. It suffices to impose the initial conditions on $x_0(t)$ alone, i.e.

$$x_i = 0, \qquad \dot{x}_i = 0 \quad \forall i \neq 0. \tag{50}$$

   Show that – should this ansatz work – reduces in first order of ε to

$$\left(\ddot{x} + \frac{k}{m}x_0\right) + \varepsilon\left(\ddot{x}_1 + \frac{k}{m}x_1 - 3\frac{C}{m}x_0^2\right) + \mathcal{O}(\varepsilon^2) = 0. \tag{51}$$

3. For the initial conditions of elongation $x(0) = A$ with no initial velocity, solve the zero-th order to get

$$x_0(t) = A\cos\left(\sqrt{\frac{k}{m}}t\right) \tag{52}$$

   just as for the harmonic oscillator and as to be expected.

4. Now get the first order solution

$$x_1(t) = a\sin\left(\sqrt{\frac{k}{m}}t\right) + b\cos\left(\sqrt{\frac{k}{m}}t\right) + \frac{A^2C}{2k}\left(3 - \cos\left(2\sqrt{\frac{k}{m}}t\right)\right). \tag{53}$$

   The a and b are integration constants. If you run into trigonometric troubles to reproduce the factor of 2 in the argument of the last cosine, think a bit about trig identities...

5. Impose the boundary conditions on $x_1(t)$ to get

$$a = 0, \qquad b = -\frac{A^2C}{k}. \tag{54}$$

6. Now look at the full (first-order) solution

$$x(t) = x_0(t) + \varepsilon x_1(t).$$

What you can see here is that you get overtones. If you were to go to higher orders, you would get – besides the overtone with frequency[30] $2\sqrt{\frac{k}{m}} =: 2\omega_0$ – also frequencies of $n\omega_0$ for any natural number $n \geq 1$. If you add driving forces, you will get them even in this order!

7. As a small extension of just doing perturbation theory, let's look at the time-averaged position

$$\bar{x} = \frac{1}{T}\int_0^T x(t)\mathrm{d}t, \tag{55}$$

where

$$T = \frac{1}{\omega_0} = \sqrt{\frac{m}{k}}.$$

This should turn out to be

$$\bar{x} = \varepsilon\frac{3}{2}\frac{A^2C}{k} = \varepsilon\frac{3}{2}\frac{A^2Cm}{\omega_0^2}, \tag{56}$$

which is obviously different from the behaviour of the unperturbed harmonic oscillator.

8. What would happen if we were to go to higher orders in perturbation theory? To save time and space, let's just summarize one of the exercises given in [Kuy03]: You will find that the base frequency ω and its overtones change compared to ω_0. It will look like

$$\omega^2 = \omega_0^2\left(1 - \mathcal{O}(\varepsilon^2)\right). \tag{57}$$

In order to get this result, the perturbation must not get too large. If it is, the unboundedness of the cubic potential will give us a headache. Also, as the ODE are not linear anymore, there will not be any kind of superposition principle as with linear ODE.

9. If we wanted to do Lagrangian or Hamiltonian mechanics, the former would receive no big difference as our second order ODE would be the same. For the Hamiltonian case, note that the perturbing potential still is time-independent, so the Hamiltonian is just

$$H = T + V.$$

Neither would the canonical momentum change.

[30] I apologize, but here ω does *not* refer to angular frequency. It is the ordinary, or temporal frequency.

10. Let us reiterate the unboundedness:

Unboundedness of cubic potentials

Already in second order we get a need to keep the perturbation small. This becomes even more apparent in ever higher orders. Even without perturbation theory this can be seen: Were we to look at large negative values of x the resulting force of the cubic potential would drive the solution to ever smaller values of x. The solution will diverge.

The problem of vanishing superposition principles obviously holds for more general potentials than a cubic one. If you introduce driving forces with driving frequency Ω in these more general cases[31], you can get funky behaviour like discontinuous behaviour of the system frequency ω as you continuously change Ω. But let's stop here.

2.2 Special Relativity

2.2.1 Lorentz Invariance and Index Magic

Warning

This part will be dense. Play with it around if you have trouble with it. Manipulating indices is an incredibly useful skill to have and I will try to give away as much insight in as little pages as possible. Which (sadly?) means that I occasionally have to mix physics and maths lingo.

Without much ado, let us step right into the heart of it.

- Minkowski space $\mathbb{M}$ is flat 4D space with the difference from Euclidean 4D space that its metric is not diag(1,1,1,1) but $\eta = \text{diag}(-1,1,1,1)$ which is called the Minkowski metric. As in Euclidean spaces, this *of course* only is true for Cartesian coordinates. You can turn this constant, but non-diagonal ('null-coordinates'), non-constant but diagonal (spherical) or neither constant nor diagonal (null and spherical).

- The use of Minkowski space is to provide a unified treatment of space and time by setting its coordinates up in the following way:

$$X = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \in \mathbb{M}. \tag{58}$$

- Points in Minkowski space are called 'events' as they have a time coordinate telling us when they happen and space-coordinates telling us where.

[31] E.g. a quartic potential $\propto x^4$.

- I will use the index notation of [Sre11] and for the moment not be too concerned with factors of c, the speed of light in vacuum (though occasionally I might throw it in for good measure). As before, the Einstein sum convention will be used, see footnote 16. So let us collect bits about this:

  Here Be Dragons!

  If the following mathematical description is confusing technobabble for you, either *ignore this point* or look it up in [Gou13], [Nab10] or the first bits of [Kri01].

  ▼ I'll make this point first mathematically then physically.

  Maths: Events X in $\mathbb{M}$ can be identified with points in the tangent space $T_X\mathbb{M}$ of $\mathbb{M}$ by specifying an origin for the affine (think Euclidean vector space) coordinates on $\mathbb{M}$. Physically this corresponds to the choice of an inertial observer. More on this below. Again, don't be afraid of this differential geometric hullabaloo.

  Physics: Put differently, it distinguishes between events and how events move. In classical mechanics this corresponds to position and velocity/momentum – obviously, they are in different spaces, as $\vec{x}+\vec{v}$ looks just revolting after a few semesters of physics. However, if we take a point X and the origin $0 \in \mathbb{M}$ and then look at their distance $\overline{0X}$ it makes sense that this will be proportional to any velocity taking us from 0 to X.

  ▼ I will, as long as we are dealing with special relativity, be very nonchalant with the difference between things like $\overline{0X}$ and X – that way we can easily adapt the language below as we need.

  ▼ Events (in the sense of $X = \overline{0X}$, points in Minkowski space) and lines between events $\overline{XY}$ are called contravariant vectors (math-lingo: they are elements of the tangent space of some event, i.e. of some observer going through that event) and written in components with an index up like so:

  $$X = (x^\mu)_{\mu=0,1,2,3}\,. \tag{59}$$

  Greek indices run from 0 to 3, zero denoting the time-component, the others spatial components. It is also quite common to use Latin indices to run from 1 to 3, i.e. label only spatial components.

  ▼ The corresponding elements of the dual space/the cotangent space are called covariant vectors and have an index down like so:

  $$p = (p_\mu)_{\mu=0,1,2,3}\,. \tag{60}$$

  ▼ As a common abuse of notation vectors/covectors (i.e. contravariant/covariant vectors) are frequently identified with their components.

▼ A quick intro to tensors: Scalars are rank-0 tensors. Vectors are rank-1 tensors. Matrices are rank-2 tensors. So far it is just linear algebra. Matrix multiplication of a vector then reads

$$x^\mu = A_\mu{}^\nu y_\nu,$$

while scalar products with a bilinear, symmetric, positive definite matrix M would look like so:

$$x^\mu M_{\mu\nu} y^\nu = s \in \mathbb{R}.$$

Now what's with all the up and down of indices? For vectors I already covered that. But with the sum convention at hand, we can make sense of it in a more general case: Anything with n indices up and m indices down will be a linear mapping that takes n covectors/covariant vectors and m vectors/contravariant vectors and spit out a scalar, i.e. a real number. **This linear mappings are called tensors of rank** (n,m)**.** Anything with a total number of 2 indices can be represented as a matrix. Anything with one as a 'vector'. It is common to denote column vectors as contravariant vectors and row vectors as covariant vectors. For matrices it becomes irksome to distinguish one with two indices up from one with two indices down or even one up, one down... All the more reason to forget about the matrix representation and head straight to index notation and general tensors.

▼ That being said, there is an important thing that's easiest explained with matrices: Index *placement* is important! This can be made visible by transposing matrices:

$$\eta^T_{\mu\nu} = \eta_{\nu\mu}, \quad A^\mu{}_\nu{}^T = A_\nu{}^\mu.$$

Why? Transposing changes which argument of the bilinear map goes where, e.g. $A^T(X,Y) = A(Y,X)$. The same goes for more general tensors. So even though it gets lost in calling something, say, a (1,2) tensor, the following are different:

$$A^\mu{}_\nu{}^\alpha \neq A^{\mu\alpha}{}_\nu.$$

Why this care? Often one finds the abuse of notation of writing such a tensor as

$$A^{\mu\alpha}_\nu,$$

which utterly obscures what exactly goes where. While this is kinda tolerable in quick notes or as a conscious abuse of notation, it can occasionally make translating different notations unnecessarily difficult (e.g. index-free to index-based).

▼ Using the just introduced abuse of notation, we can then write the Minkowski metric as

$$\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1). \tag{61}$$

▼ As things with indices are just heaps of real numbers we can commute – the implicit sums of the Einstein convention will take care of the non-commutativity of matrix products or even tensor products.

- If we have two events $X,Y \in \mathbb{M}$, the vectorial difference between them tells us their causal relation:
  - ▼ If
    $$\eta(X,X) = \eta_{\mu\nu} X^\mu X^\nu < 0,$$
    we call X time-like.
  - ▼ If
    $$\eta(X,X) = \eta_{\mu\nu} X^\mu X^\nu > 0,$$
    we call X space-like.
  - ▼ If
    $$\eta(X,X) = \eta_{\mu\nu} X^\mu X^\nu = 0,$$
    we call X light-like or null.
  - ▼ Only two events X,Y such that $\overline{XY}$ is light-like or time-like can influence each other. Only time-like related events can be reached by a massive particle.
  - ▼ The set of events separated light-like from an event X is called the light-cone of X.
- An isometry L is a linear transformation of Minkowski space such that
  $$\forall X,Y \in \mathbb{M}: \quad \eta(X,Y) = \eta(LX, LY). \tag{62}$$
- In index notation (62) reads for the homogeneous Lorentz group:
  $$\forall X^\mu, Y^\mu: \quad \eta_{\mu\nu} X^\mu Y^\nu = \eta_{\mu\nu} L^\mu{}_\alpha X^\alpha L^\nu{}_\beta Y^\beta.$$
  Exercise: Try to rephrase this without explicitly mentioning X or Y, both in index notation and index-free.
- As linear transformations the isometries form a group called the Poincaré group, sometimes also named the inhomogeneous Lorentz group. Note: Here it is actually important to differentiate between X and $\overline{0X}$ – if we add a constant A to the former, we will change the metric $\eta(X,X) \neq \eta(X+A,X+A)$, but on the other hand (thanks to Minkowski space having an addition of elements) $\eta(\overline{0X},\overline{0X}) = \eta(\overline{(0+A)(X+A)},\overline{(0+A)(X+A)})$.
- **Nomenclature:** The subgroup with translations by A vanishing is called the Lorentz group $O(3,1)$ and contains as subgroup the rotations of 3-space while keeping time constant. $O(3,1)$ has four connected components and the component containing the identity element is called the proper, orthochronous Lorentz group. Orthochronous means, the transformation does not interchange past and future. As with the $O(n)$ groups, we call the elements with determinant $+1$ members of $SO(3,1)$.

- **Sign convention!** Some people prefer making space all negative and time positive in the Minkowski metric. This also leads to things like $SO(1,3)$ instead of $SO(3,1)$ and different definitions for space-like and time-like. Neither am I a fan of this different convention, nor is this the one used in [Sre11], so we will stick to the sign being in the time-component of the metric. The convention $-+++$ goes by the names '(general) relativity convention', 'mostly pluses', 'Pauli convention', 'East coast convention', 'space-like convention'. The opposite one has the names particle physics convention, mostly minuses, Landau-Lifshitz convention, time-like convention and West coast convention going. Obviously, East and West coast of the US already differed significantly before Bad Boy Records and Death Row Records.

- Besides the rotations in 3D, another famous element of (homogeneous) Lorentz transformations are the so called boosts corresponding to a mere change of relative velocity between to observers. If the boost is in x-direction it has the form

$$\begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{63}$$

  where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ and $\beta = \frac{v^2}{c^2}$. While two boosts in the same direction give another boost and the relativistic addition law for velocities, two boost in different directions are a boost *AND* a rotation. If you set $\gamma = \cosh\xi$ and $\beta\gamma = \sinh\xi$, you will realize that the law of addition of velocities is just a hyperbolic addition theorem. And please convince yourself that the link between (63) and the hyperbolic functions makes sense.

- Scalar quantities that don't change value under a Lorentz transformation are called Lorentz scalars. In general, a (p,q)-tensor is called Lorentz covariant if it transform under Lorentz transformations in the following way:

$$T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} \mapsto L^{\mu_1}{}_{\alpha_1}\cdots L^{\mu_p}{}_{\alpha_p} L^{\beta_1}{}_{\nu_1}\cdots L^{\beta_q}{}_{\nu_q} T^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q}. \tag{64}$$

- As the metric η is obviously invertible, we can write down an inverse for it in index notation as $\eta^{\mu\nu}$. Matrix-multiplying this inverse with η itself has to give the identity matrix. Let's see how this works:

$$\eta^{\mu\nu}\eta_{\nu\alpha} = \delta^{\mu}{}_{\alpha} \quad \Longleftrightarrow \quad \eta^{-1}\eta = \mathbb{1}.$$

  Now why the index placement like this? The identity matrix has to turn a vector into itself, but is a second-rank tensor, so it has to be of type (1,1). Then the LHS of the last equation is the way to go to write down an inverse matrix for $\eta_{\mu\nu}$.

- The metric $\eta_{\mu\nu}$ provides us with a unique mapping between space and dual space – or less technical, with a way to raise and lower indices. So if we start with a

vector V^μ or a covector ω_μ we can turn them into a covector or vector respectively by setting

$$V_\mu = \eta_{\mu\nu} V^\nu, \text{ and } \omega^\mu = \eta^{\mu\nu}\omega_\nu. \tag{65}$$

As an example, let us look at the coordinate vector $x^\mu = (ct, x^1, x^2, x^3)^T$ and see what it looks like as a covector:

$$x_\mu = \eta_{\mu\nu} x^\nu = (-ct, \vec{x}^T). \tag{66}$$

Note two things: First of all, the different sign in the zero-th component. Second, how the T for transpose changed position.

- Derivatives can be given indices, too!

$$\partial_\mu := \begin{pmatrix} \frac{\partial}{\partial x^0} & \frac{\partial}{\partial x^1} & \frac{\partial}{\partial x^2} & \frac{\partial}{\partial x^3} \end{pmatrix} \tag{67}$$

There are basically two ways to see why the index has to be down: Either, you can prove it by checking how it behaves under a Lorentz transformation (**Exercise!**), or you can handwave and just save the mnemonic that it's a bit like fractions. If you divide something by a fraction the numerator becomes the denominator and vice versa, and then just think of the index up as a numerator and the index down as a denominator[32].

- Two other important things with derivatives are the one with raised indices

$$\partial^\mu = \eta^{\mu\nu}\partial_\mu = \begin{pmatrix} -\frac{\partial}{\partial x^0} \\ \frac{\partial}{\partial x^1} \\ \frac{\partial}{\partial x^2} \\ \frac{\partial}{\partial x^3} \end{pmatrix} \tag{68}$$

and the D'Alembertian

$$\Box := \partial_\mu \partial^\mu = -\partial_t^2 + \Delta_x. \tag{69}$$

Note how the sign in the time-component appears this time in the contravariant version ∂^μ – for x^μ it appeared in the covariant one – and that $\partial^\mu \partial_\mu$ gives just the wave operator.

You can find more on this in maaaaaaany books. Both in mathematics (differential geometry, linear algebra, sometimes analysis, ...) and physics (electrodynamics, relativity, particle physics, mechanics, ... – all both theoretical and experimental), sometimes more, sometimes less depending on the corresponding authors' preferences. If you found my fairly mathematical and very short summary of results rather scary, try looking into your favourite book covering relativity and try to match it with what I mention. Feel free to fill in gaps in my exposition – this is by no means a complete summary of results.

[32] It's silly, but that always helped me remember this in the beginning.

2.2.2 Kinematics and Mandelstam variables

Having set all of this up, we are now in the position to do a bit more physics.

- A massive particle has an associated four-momentum p_μ given by

$$p_\mu := \begin{pmatrix} -\frac{E}{c} & p_x & p_y & p_z \end{pmatrix} \tag{70}$$

  or equivalently

$$p^\mu = \begin{pmatrix} \frac{E}{c} \\ \mathbf{p} \end{pmatrix},$$

  with

$$p^\mu p_\mu = -m^2 c^2. \tag{71}$$

  The minus sign on the right hand side is the reason why some people prefer $\eta = \mathrm{diag}(1, -1, -1, -1)$. I am not one of them.

- From equations (70) and (71) we also can get the expansion in v/c

$$E = mc^2 \left(1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \dots\right), \tag{72}$$

  which, for example, allows to calculate relativistic correction to quantum mechanical results via perturbation theory. I shall not do that here.

- A massless particle, on the other hand, has a four-momentum

$$p_\mu := \begin{pmatrix} -\frac{E}{c} & p_x & p_y & p_z \end{pmatrix}, \tag{73}$$

  but this time such that

$$E = |\mathbf{p}|\, c. \tag{74}$$

  Massless particles therefore have a null-vector as four-momentum.

- If we now have any kind of collision or decay or reaction with n in-going particles and m out-going ones, the energy conservation and momentum conservation of Newtonian kinematics is expressed as energy-momentum-conservation in terms of four-vectors or four-covectors:

$$\sum_{i=1}^{n} p^{\mathrm{in},i}_\mu p^\mu_{\mathrm{in},i} = \sum_{i=1}^{m} p^{\mathrm{out},i}_\mu p^\mu_{\mathrm{out},i} \tag{75}$$

  and

$$\sum_{i=1}^{n} p^\mu_{\mathrm{in},i} = \sum_{i=1}^{m} p^\mu_{\mathrm{out},i}. \tag{76}$$

- A rough classification of different n and m (adapted from [Gou13], p.294) would be

- ▼ De-excitation or decay: $n = 1$, $m \geq 2$
- ▼ Elastic collision: $n = m = 2$ and $m_i^{\text{in}} = m_i^{\text{out}}$ – the reason for not just saying $n = m$ is that you normally do not have interactions between more than two particles on a fundamental level. That is more a matter of experimental fact than of mathematical consideration and if I remember correctly there are at least 'effective' three-body-interactions.
- ▼ Annihilation/Inelastic collision/fusion: $n = 2$ but $m > 0$, 'as long as something different comes out' (see me wave hands vigorously).

- Other invariants that are incredibly important for describing $n = m = 2$ are the so-called *Mandelstam variables*:
  - ▼ $s := -(p_1 + p_2)^\mu (p_1 + p_2)_\mu = -(p_1' + p_2')^\mu (p_1' + p_2')_\mu$
  - ▼ $t := -(p_1 - p_1')^\mu (p_1 - p_1')_\mu = -(p_2 - p_2')^\mu (p_2 - p_2')_\mu$
  - ▼ $u := -(p_1 - p_2')^\mu (p_1 - p_2')_\mu = -(p_2 - p_1')^\mu (p_2 - p_1')_\mu$

  s has the interpretation of the *total mass squared* of the system.

- **Exercise:** You can actually check that

$$s + t + u = c^2(m_1^2 + m_2^2 + m_1'^2 + m_2'^2). \tag{77}$$

  This is a *good* exercise to do.

- Lastly, there are two common choices of frames to do calculations in. Their usefulness depends on the corresponding physical situation at hand.
  - ▼ The centre-of-momentum frame (CM or COM frame) where

$$\mathbf{p}_1 + \mathbf{p}_2 = 0 = \mathbf{p}_1' + \mathbf{p}_2', \tag{78}$$

    i.e. the two particles have equal and opposite three momenta before and, due to (76), also after the collision.
  - ▼ The lab frame where one of the two particles is initially at rest.

More on this can be found in your favourite particle/nuclear physics textbook (e.g. [HM84], [Gri04] or [Mar06][33]) or textbooks on special relativity (e.g. [Gou13]).

2.3 Electrodynamics

For this section, I will follow the units used in [Sre11], that is Heaviside-Lorentz and (sometimes) $c = 1$.

[33] These books are either thoroughly outdated, out-of-print or at least only old editions. Sorry about that. [HM84] I like very much despite its age, but the other ones probably benefit very much from a more modern edition as they are closer to current, experimental data, like the Higgs boson. But the relativistic kinematics don't change – so all are perfectly good references for this section.

2.3.1 3D versus 4D

First, let us recap the basics of electrodynamics – the Maxwell equations.

1. The Maxwell equations in vacuum (with sources!) are:

$$\nabla \cdot \mathbf{E} = \rho, \tag{79}$$
$$\nabla \times \mathbf{B} - \dot{\mathbf{E}} = \mathbf{J}, \tag{80}$$
$$\nabla \times \mathbf{E} + \dot{\mathbf{B}} = 0, \tag{81}$$
$$\nabla \cdot \mathbf{B} = 0. \tag{82}$$

2. While the symmetry group of the Maxwell equations *is* the Lorentz group, this is far from obvious in the three-dimensional notation. Therefore, it is beneficial to rewrite this in a manifestly four-dimensional way:

$$A^\mu = \begin{pmatrix} \phi \\ \mathbf{A} \end{pmatrix}, \tag{83}$$
$$J^\mu = \begin{pmatrix} \rho \\ \mathbf{J} \end{pmatrix}, \tag{84}$$

and lastly, collect the **E**- and **B**-field in an anti-symmetric second-rank tensor $F^{\mu\nu}$ like so:

$$F^{ij} = \epsilon^{ijk} B_k, \tag{85}$$
$$F^{0i} = \frac{E^i}{c}. \tag{86}$$

So, if we were to write the $F^{\mu\nu}$ as a matrix, we'd get

$$\begin{pmatrix} 0 & -\frac{E_x}{c} & -\frac{E_y}{c} & -\frac{E_z}{c} \\ \frac{E_x}{c} & 0 & -B_z & B_y \\ \frac{E_y}{c} & B_z & 0 & -B_x \\ \frac{E_z}{c} & -B_y & B_x & 0 \end{pmatrix} \tag{87}$$

Remember for this that for the 3D vectors $E_i = E^i$ as no signs come into play from the Minkowski metric used to raise and lower indices.

3. Using this rephrasing, we can rewrite the Maxwell equations in the more compact form

$$\partial_\mu F^{\mu\nu} = J^\nu, \qquad \epsilon_{\mu\nu\kappa\lambda} \partial^\nu F^{\kappa\lambda}. \tag{88}$$

2.3.2 Gauge Freedom and Solving Maxwell Equations

1. Solving the homogeneous equations (81) and (82) gives rise to the electrostatic potential ϕ and the vector potential **A**, such that

$$\mathbf{E} = -\nabla\phi - \dot{\mathbf{A}}, \tag{89}$$
$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{90}$$

2. In 4D, we can rearrange this as

$$A^\mu = \begin{pmatrix} \phi \\ \mathbf{A} \end{pmatrix}. \tag{91}$$

This also gives us the chance to make the definition of $F^{\mu\nu}$ more transparent:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{92}$$

3. From this follows immediately that

$$\partial_\mu J^\mu = 0, \tag{93}$$

which rewritten in 3D is nothing but the charge conservation/continuity equation. (**Exercise:** Check this!)

4. Now that we have introduced the gauge fields, we have to mention gauge freedom. While A^μ (or ϕ and $\mathbf{A}$) uniquely specify the values of $F^{\mu\nu}$ (or $\mathbf{E}$ and $\mathbf{B}$), the other way around doesn't work. Because changing

$$A^\mu \to A^\mu - \partial^\mu \chi, \tag{94}$$

where χ is some real-valued function, gives rise to the same $F^{\mu\nu}$ as the original A^μ did. Correspondingly, in 3D the changes

$$\phi \to \phi + \dot{\chi}, \qquad \mathbf{A} \to \mathbf{A} - \nabla\chi \tag{95}$$

give rise to the same $\mathbf{E}$ and $\mathbf{B}$ as the original ϕ and $\mathbf{A}$.

5. Now the trick is, to choose χ cleverly enough to simplify solving the Maxwell equations.

6. **Exercise:** Show that a convenient choice of gauge χ can be found so that the the inhomogeneous Maxwell equation is turned into

$$\Box A^\mu = J^\mu. \tag{96}$$

This choice of χ is the Lorenz gauge[34]. Make sure that this choice can always be made.

7. In Lorenz gauge, the solution to the inhomogeneous Maxwell equations can be given by the retarded/advanced potentials:

$$A^\mu_{ret/adv}(t,\mathbf{r}) = \frac{1}{4\pi} \int \frac{J^\mu(t \mp |\mathbf{r}-\mathbf{r}'|/c, \mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r}', \tag{97}$$

where the retarded solution describes how the past influenced an observer at position $\mathbf{r}$ and the advanced solution how a field at $\mathbf{r}$ will propagate to other points in space. This is actually a direct application of the Green's functions described in 1.2.3, now only slightly modified to include the speed limit of c imposed by special relativity.

[34] Lorenz. Not Lorentz.

8. The Lagrangian density of electrodynamics is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + J^{\mu}A_{\mu}. \tag{98}$$

   As an **exercise** you can try to convert this into a 3D-expression.

9. To get the Hamiltonian of a field theory, one needs to have a canonical momentum $\frac{\partial \mathcal{L}}{\partial_0 A^{\mu}}$ for every field degree of freedom A^{μ}. Otherwise, our Hessian condition for the Legendre transformation will break down. Now take a look again at $F^{\mu\nu}$ which appears in our Lagrangian. Can this work?[35]

2.4 Quantum Mechanics

Quick heads-up: I will (mostly) set $\hbar = 1$. For a good introduction on this abuse of notation I recommend the notes Jaffe wrote on this: `https://stuff.mit.edu/afs/athena/course/8/8.06/spring08/handouts/units.pdf`.

2.4.1 A Review of Dirac Notation

Dirac's notation for (generalized) states in quantum theory is tremendously useful, so let's quickly recap it. As long as our Hilbert space $\mathcal{H}$ is just finite-dimensional, we can salvage a lot of our understanding of finite vector spaces. There's the space $\mathcal{H}$ and its dual space $\mathcal{H}^*$. Since a Hilbert space is a complete inner product space, we can easily go from one to the other, construct orthonormal bases and the like. Let's call the inner product $\langle \cdot, \cdot \rangle$ An element of our Hilbert space would in Dirac notation be written with a 'ket', like so: $|\Psi\rangle$. Its dual is the corresponding $\langle \Psi|$ called a 'bra'. For two states $|\Phi\rangle, |\Psi\rangle$ we write their inner product in Dirac notation as $\langle \Psi|\Phi\rangle = \langle |\Psi\rangle, |\Psi\rangle \rangle$, called a braket. Therefore, the whole Dirac notation is also known as 'braket-notation'[36]. Remember that for a complex Hilbert space, the inner product is only sesquilinear, i.e. $\langle \Psi|\Phi\rangle = \overline{\langle \Phi|\Psi\rangle}$, where the overline denotes complex conjugation.

While this all works well and nice also for the naive elements of, say, L^2, the whole things crumbles a bit if you take plane waves as states[37]. They would not be square integrable over the whole $\mathbb{R}^n$. There are technical ways (called 'rigged Hilbert spaces', see [Bö93][38], [de 05] or [dBG02]) to make this work – but since there are, we will take the stance that what we do works, so why go down the rabbit hole?

As a result, the position eigenstates can fulfil orthogonality if one extends the orthogonality relation for two normal states $|i\rangle, |j\rangle$ s.t. $\langle i|j\rangle = \delta_{ij}$ to the following:

$$\langle x|y\rangle = \delta(x - y). \tag{99}$$

[35] Hint: Look at $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ – what contributes to the 'new' needed canonical momentum?

[36] Some people cringe at this very punny [sic!] use of English due to Dirac. I like it.

[37] L^2, like all Hilbert spaces, has the beautiful property that its dual is isomorphic to L^2 – thanks to Riesz' representation theorem. Different function spaces might not be 'reflexive', i.e. isomorphic to their dual. If you encounter a PDE whose solution is not easily put into a Hilbert space, you are in for a hard time.

[38] Depending on where you look, you will find this author's name as both 'Bohm' and 'Böhm'.

And as position and momentum are Fourier pairs, we can nicely write Fourier transforms[39]:

$$\langle x|k\rangle = e^{ikx} \quad \Longleftrightarrow \quad \langle k|x\rangle = e^{-ikx}. \tag{100}$$

Note that our convention for the Fourier transform means that

$$\int e^{i\mathbf{k}\cdot\mathbf{x}} \mathrm{d}^n x = (2\pi)^n \delta^n(k).$$

Operators are inserted between bra and ket, e.g. as $\langle\Psi|A|\Phi\rangle$. And while a braket itself is a complex number,

$$\langle\Psi|\Phi\rangle \in \mathbb{C},$$

a ketbra will be an operator. This makes sense insofar as a bra is from the dual space, a ket from the space itself, and therefore the bra is a map taking Hilbert space elements to complex numbers. In the braket the ket will be the argument of the operator given by the bra.

So where does the ketbra $|A\rangle\langle B|$ as an operator stem from? Think of it as a 'tensor' – you need to hit it from the left with a bra $\langle A'|$ and from the right with a ket $|B'\rangle$ to get the multiplication of two brakets, i.e. a multiplication of two complex numbers: $\langle A'|A\rangle \cdot \langle B|B'\rangle$. So the ketbra is a map $\mathcal{H}^* \times \mathcal{H} \to \mathbb{C}$.

A particularly useful operator in ketbra notation is the projection operator:

$$|\Psi\rangle\langle\Psi|\,. \tag{101}$$

This operator projects any state $|\Phi\rangle$ onto the state $|\Psi\rangle$ and therefore is perfect to expand states in a given basis. Let's say we have a (normalized) basis of states $\{|n\rangle\}_{n\in I}$, with I some index set. It does not matter whether the index set is finite, countable or even uncountable, so in order to sum over it, let's use the fancy symbol $\sum\!\!\!\!\!\!\int$. Then any state $|\Psi\rangle$ can be written with our projection operators as

$$|\Psi\rangle = \sum\!\!\!\!\!\!\int_n |n\rangle \underbrace{\langle n|\Psi\rangle}_{=:\Psi_n\in\mathbb{C}}\,. \tag{102}$$

If you have not used Dirac's notation so far, have a close look at the next few sections because I will employ this notation heavily.

2.4.2 Ladder Operators

I will not carefully derive the ladder operators discussed in this section. Rather, I redirect you for that to standard quantum mechanics texts like [Gri12], [Sha14], [Wei13], [Rae16] or [Boh89]. However, I will give you a nice one-sentence-catch-all:

> Ladder operators enable you to go from one state labelled by some quantum number with value n to a different, orthogonal state by changing only the value of that quantum number by ± 1.

[39] I hope I correctly adapted to the normalization convention for Fourier transforms in [Sre11]...

An interesting application of ladder operators to the hydrogen atom and the operator corresponding to the Runge-Lenz vector from classical mechanics can be found in [Sha14], p.422f.

Harmonic Oscillator

The harmonic oscillator is a particularly important and also simple application of ladder operators. This bit is incredibly important, so if you want to do some exercises, start here! The Hamiltonian in this case is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2\hat{x}^2}{2}. \tag{103}$$

However, introducing ladder operators significantly simplifies this Hamiltonian.

1. In this case, the ladder operators made out of $\hat{x}$ and $\hat{p}$ are given by

$$\hat{a} := \sqrt{\frac{m\omega}{2\hbar}}(\hat{x} + \frac{i}{m\omega}\hat{p}), \tag{104}$$

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}(\hat{x} - \frac{i}{m\omega}\hat{p}), \tag{105}$$

   or, inverted (**Exercise!**),

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^\dagger + \hat{a}), \tag{106}$$

$$\hat{p} = i\sqrt{\frac{\hbar m\omega}{2}}(\hat{a}^\dagger - \hat{a}). \tag{107}$$

2. They fulfil the canonical commutation relations (**Exercise!**):

$$[\hat{a},\hat{a}^\dagger] = 1. \tag{108}$$

3. Written in $\hat{a}$ and $\hat{a}^\dagger$, the Hamiltonian turns into (**Exercise!**)

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) = \hbar\omega\left(\hat{a}\hat{a}^\dagger - \frac{1}{2}\right). \tag{109}$$

4. From (108) and (109) one gets (**Exercise!**)

$$[\hat{a},\hat{H}] = \hat{a}, \tag{110}$$

   and

$$[\hat{a}^\dagger,\hat{H}] = -\hat{a}^\dagger. \tag{111}$$

5. Eigenstates of the Hamiltonian, i.e. energy eigenstates, are now written as $|n\rangle, n \in \mathbb{N}_0$ and have energy $(n + \frac{1}{2})\hbar\omega$. We assume that they are normalized. Using the commutation relations between $\hat{H}$ and the ladder operators we can show that they generate new eigenfunctions (**Exercise!**). Also, since it is easy to show that eigenvalues of $\hat{H}$ have to be non-negative, we know that there has to be a smallest value which then justifies our notation.

6. It now follows from taking normalized eigenstates that (**Exercise!**):

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle\,, \hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{112}$$

More can be found in your favourite quantum mechanics textbook (say, [Sha14] or [Rae16]) or even just Wikipedia.

Angular Momentum

In the case of angular momentum, look at the four operators L_x, L_y, L_z and $\mathbf{L}^2$. While $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$ commutes with the other three, they fulfil among each other

$$[L_l, L_m] = i\hbar\epsilon_{lmn}L_n. \tag{113}$$

It is possible to choose simultaneous eigenstates of *one* L_n and $\mathbf{L}^2$ but not for two different L_l and L_m. Conventionally, one chooses L_z and $\mathbf{L}^2$ to label eigenstates of angular momentum. Then one can built ladder operators out of the remaining two:

$$L_\pm = L_x \pm iL_y. \tag{114}$$

In the case of different bases for a state space of added angular momenta these methods provide the way to calculate Clebsch-Gordan coefficients – fancy words for a change of basis.

2.4.3 A Smorgasbord of Perturbation Theory

In this section, I will collect a variety of results both from time-independent and time-dependent perturbation theory in non-relativistic quantum mechanics. I also include an example of applying time-independent perturbation theory to the harmonic oscillator.

Time-independent Perturbation Theory

Assuming we have an unperturbed Hamiltonian H_0 whose (non-degenerate) energy eigenstates we know[40], let's look at a more complicated problem with Hamiltonian

$$H = H_0 + \varepsilon H_1. \tag{115}$$

[40] (a) I will not bother with non-degeneracy, but it can be solved by carefully choosing the basis for the given Hilbert space. (b) We want H_0 to be self-adjoint, so its eigenstates *do* form a basis for our Hilbert space. We don't care at this point about the rigorous way to achieve this and just assume that H_0 and the Hilbert space match up.

There are different preferences how to do the expansion now. I prefer the way it is done on the Wikipedia[41]: Take $\varepsilon \in [0,1]$. That way you have:

- ε to expand in.
- A continuous transition from full system to unperturbed system.
- For a given H_1 a simpler handle to check convergence radii[42].

The way to go is to expand both the energy eigenstates and the energy of the full system H in ε:

$$|n\rangle = |n^{(0)}\rangle + \varepsilon\,|n^{(1)}\rangle + \varepsilon^2\,|n^{(0)}\rangle + \dots, \tag{116}$$

$$E_n = E_n^{(0)} + \varepsilon E_n^{(1)} + \varepsilon^2 E_n^{(2)} + \dots \tag{117}$$

The (n) gives the n-th term of a Taylor series in ε for E or $|n\rangle$, respectively.

After a bit of calculation you get that the first-order correction to the energy eigenstate $|n^{(0)}\rangle$ will be

$$|n^{(1)}\rangle = \sum_{k\neq n} \frac{\langle k^{(0)}|H_1|n^{(0)}\rangle}{E_n^{(0)} - E_k^{(0)}}\,|k^{(0)}\rangle\,, \tag{118}$$

while the first-order energy correction is simply

$$E_n^{(1)} = \langle n^{(0)}|H_1|n^{(0)}\rangle\,. \tag{119}$$

Combining (119) and (118) we can get the second-order correction to the energy $E^{(2)}$ as

$$E_n^{(2)} = \langle n^{(0)}|H_1|n^{(1)}\rangle = \sum_{k\neq n} \frac{\left|\langle k^{(0)}|H_1|n^{(0)}\rangle\right|^2}{E_n^{(0)} - E_k^{(0)}}. \tag{120}$$

Let's not go through (other) higher orders or degenerate perturbation theory, even though the latter is definitely physically incredibly important. Rather, let's have a look at an example. If you want to rehearse your knowledge of the rest of perturbation theory, go to your favourite book on quantum mechanics. Just a few examples would be, again: [Gri12], [Sha14], [Flü08], [Wei13], [Rae16], [Boh89] or even the Wikipedia article cited in footnote 41.

Perturbing the Harmonic Oscillator with a Cubic Potential

This example is adapted from [FW04] and [Flü93]. As much of my personal library on QM happens to be in German, I would like to apologize for having to reference German books, but at least for the latter I know of an English translation, [Flü08], which – as it happens – contains a lot more than my edition does.

[41] `https://en.wikipedia.org/wiki/Perturbation_theory_(quantum_mechanics)#Time-independent_perturbation_theory`

[42] Guess on my side. Not that I would want to do that.

So, here we will take

$$H_1 = C\hat{x}^3, \tag{121}$$

and use C to turn the units into something sensible – after all, our definition of ε didn't have any. Now, before we start there is an important thing to note straight away: While H_0, the harmonic oscillator, has a energy spectrum that is bounded from below we cannot expect this to be the case for the perturbed Hamiltonian H. After all, just looking at the classical case discussed in 2.1.2, no matter how small εC, H_1 will make H unbounded from below. The perturbation series will not converge and there always is a small probability that the state will tunnel through the introduced potential barrier and 'go to' energy $-\infty$. The only thing we can do is to choose εC small enough that this tunnelling probability is low enough to expect at least on physical grounds a sensible result even from perturbation theory. Making sense of this physical intuition mathematically is *a lot* harder and apparently still ongoing research ([FS14]). For a discussion of hardcore mathematics to quartic perturbations, see [BO99].

So, how to go about this?

1. In order to calculate (119) and (118), we will make use of the ladder operators $\hat{a}$ and $\hat{a}^\dagger$ via equation (106). So, do that and try to find out which states $|n\rangle$ can go to which states $|m\rangle$ under the perturbation[43].

2. For the first energy correction it is simple. Think about which states could be reached and what that means for equation (119)...

3. For the correction $|n^{(1)}\rangle$ you then get

$$|n^{(1)}\rangle = \frac{C}{\hbar\omega}\left(\frac{\hbar}{2m\omega}\right)^{3/2}\left[-\frac{1}{3}\sqrt{\frac{(n+3)!}{n!}}\,|(n+3)^{(0)}\rangle - 3(n+1)^{3/2}\,|(n+1)^{(0)}\rangle \right.$$
$$\left. +3n^{3/2}\,|(n-1)^{(0)}\rangle + \frac{1}{3}\sqrt{\frac{n!}{(n-3)!}}\,|(n-3)^{(0)}\rangle\right]. \tag{122}$$

4. So if we are interested in any changes in the energy, at all, we better go one order higher[44]. Using equation (120) we then get after a not too long calculation:

$$E_n^{(2)} = -\frac{\hbar^2 C^2}{8m^3\omega}\left[30n^2 + 30n + 11\right]. \tag{123}$$

Please note, that written like this the order of ε is not explicitly visible. This finishes our little example.

[43] Hint: Try to go from $|n\rangle$ to $|n \pm 1\rangle$ or $|n \pm 3\rangle$.

[44] The only reason I bothered with a second order result...

Time-dependent Perturbation Theory

Here now a few famous results for time-dependent perturbations. I will not develop this in all detail and rather skip straight to the results – this notes are (too?) long, as they are... Your favourite quantum mechanics books will help you again, say [Rae16], [Sha14], [Gri12], [CTDL77], [Wei13], [Boh89] or [Bö93]. In the following, I will denote the different pictures used by superscripts S, H or I for Schrödinger, Heisenberg and Interaction picture, respectively.

- First a quick reminder for the different pictures of quantum mechanics. In the Schrödinger picture, the state vectors carry all the time-dependence and obey the Schrödinger equation
$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi\right\rangle = H^{\mathrm{S}}\left|\Psi\right\rangle \tag{124}$$
and operators in it are time-independent. In the Heisenberg picture, we push the time-evolution operator $\exp(-iHt/\hbar)$ onto the time-independent Schrödinger picture operators A^{S}, turning them into time-dependent Heisenberg picture operators
$$A^{\mathrm{H}}(t) = e^{iH^{\mathrm{S}}t/\hbar}A^{\mathrm{S}}e^{-iH^{\mathrm{S}}t/\hbar}. \tag{125}$$
The benefit is that our states are now time-independent. The Schrödinger equation describing the time-dependence of states now is exchanged for the Heisenberg equation describing the time-evolution of operators:
$$\frac{\mathrm{d}}{\mathrm{d}t}A^{\mathrm{H}}(t) = \frac{\partial}{\partial t}A^{\mathrm{H}}(t) + \frac{i}{\hbar}[H^{\mathrm{H}}(t), A^{\mathrm{H}}(t)]. \tag{126}$$
Note that you can easily find things that don't follow this simple rule-of-thumb which things are time-dependent in which picture.

- Say, the system is described by $H = H_0 + H_1$, where now H_1 is a time-dependent perturbation. The *interaction picture* (sometimes also called the Dirac picture) is used to get rid of 'boring' bits of time-evolution, i.e. the time-evolution of the unperturbed system described by the Hamiltonian H_0. In order to achieve that, one pushes the time-evolution of H_0 onto the operators. Then the interaction picture Hamiltonian is
$$H_1^{\mathrm{I}} := e^{iH_0t/\hbar}H_1^{\mathrm{S}}e^{-iH_0t/\hbar}. \tag{127}$$
Then the states obey the (for time-dependent perturbation theory) more useful Schrödinger equation
$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi\right\rangle = H_1^{\mathrm{I}}\left|\Psi^{\mathrm{I}}\right\rangle. \tag{128}$$

- The solution of this new Schrödinger equation can formally be done by introducing the time-ordered exponential $T\exp$:
$$\left|\Psi^{\mathrm{I}}(t)\right\rangle = T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t}H^{\mathrm{I}}(t')\mathrm{d}t'\right)\left|\Psi^{\mathrm{I}}(t_0)\right\rangle =: U(t_0,t)\left|\Psi^{\mathrm{I}}(t_0)\right\rangle, \tag{129}$$
where $U(t_0,t)$ is the time-evolution operator.

- Here,
$$TA(t_1)A(t_2) = \begin{cases} A(t_1)A(t_2) & t_1 > t_2 \\ A(t_2)A(t_1) & t_2 > t_1 \end{cases}. \tag{130}$$
For $t_1 = t_2$ it doesn't matter what you choose.

- With this, the time-ordered exponential $T\exp$ can be developed in the Dyson series:
$$T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} H(t')\mathrm{d}t'\right) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} H(t')\mathrm{d}t' - \frac{1}{\hbar^2}\int_{t_0}^{t}\int_{t_0}^{t'} H(t')H(t'')\mathrm{d}t''\mathrm{d}t' + \dots, \tag{131}$$
where the $1/n!$ from the exponential get cancelled by over-counting of the integrals. Also, I omitted the picture-superscripts for the interaction picture.

- The transition probability per time $\Gamma_{i\to f}$ of a transition from a given initial state $|i\rangle$ to a final state $|f\rangle$ *in an energy continuum* is given by Fermi's Golden Rule:
$$\Gamma_{i\to f} = \frac{2\pi}{\hbar}\left|\langle f|H_1|i\rangle\right|^2 \rho(E_f). \tag{132}$$
Here, $|f\rangle$ and $|i\rangle$ are eigenstates of H_0, thus the perturbation H_1 must be negligible both in the far past and the far future to have well-defined initial and final states w.r.t. H_0. $\rho(E)$ is the density of state, i.e. the number of states at energy E. Fermi's Golden Rule works for all kinds of situation and often its inverse provides a mean life-time of the initial state.

2.4.4 Scattering Theory

!!!WARNING!!!

I have to apologize, but I guess that there might be factors of $\sqrt{2\pi}$ off in the following equations. The problem is our asymmetric Fourier transform pre-factors – I missed that while translating results from [Sha14]. But I will try to sort this out if time before the school permits. Or if someone figures it out before me and tells me, they will have my eternal gratitude. Differing conventions are a nightmare.

Scattering theory is slightly related to time-dependent perturbation theory. For example, the Dyson series 131 will appear again. As usual, this will be a collection of results and for a more in-depth treatment look at the list of QM books given ~~here~~, there and everywhere in these notes.

- First, a bit of vocabulary. Again, consider an unperturbed Hamiltonian H_0, but this time call the perturbation 'scattering potential' V. We will denote eigenstates of the full Hamiltonian
$$H = H_0 + V \tag{133}$$
by $|\Psi\rangle$ and the eigenstates of the unperturbed Hamiltonian $|\Phi\rangle$.

- If we look at the $|\Psi(\mathbf{k})\rangle$, we are interested in those of the form

$$|\Psi(\mathbf{k})\rangle = |\Psi_{\text{in}}\rangle + |\Psi_{\text{scattered}}\rangle\,. \tag{134}$$

- If we align our coordinates along $\mathbf{k}$, we can then rewrite these states asymptotically as

$$\langle\mathbf{r}|\Psi(\mathbf{k})\rangle = e^{ikz} + \frac{f(\theta,\varphi)}{r}e^{ikr}. \tag{135}$$

  $f(\theta,\varphi)$ is the *scattering amplitude*.

- Then there is also the *scattering cross-section*:

$$\sigma = \frac{\text{Transition Probability } |i\rangle \to |\Psi\rangle}{\text{Incident probability per unit area } \mathrm{d}A}. \tag{136}$$

  Now, I am not a great fan of words in formulae. However, here it is often helpful, unless you want to go for something like it is done in chapter 3.d of [Tay72]. If we are just looking at the differential cross-section into a solid angle $\mathrm{d}\Omega$, we look at

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}. \tag{137}$$

- Combining everything (with a bit of work) we get

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta,\varphi)|^2\,. \tag{138}$$

- The *scattering matrix* S (a.k.a. S-matrix) is the limit of the time-evolution operator $U(t_0,t)$:

$$S := \lim_{t_0\to-\infty}\lim_{t\to\infty} U(t_0,t). \tag{139}$$

- Evaluating the Dyson series for S to first order is called the *Born approximation* and has the nice result of linking the differential cross-section to the Fourier transform of the scattering potential V:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left|\frac{m}{2\pi\hbar^2}\int e^{-i\mathbf{q}\cdot\mathbf{r}'/\hbar}V(\mathbf{r}')\mathrm{d}^3\mathbf{r}'\right|^2. \tag{140}$$

  Here, $\mathbf{q} := \mathbf{p}_{\text{out}} - \mathbf{p}_{\text{in}}$ is the *momentum transfer*. In this formula, my warning regarding factors of $\sqrt{2\pi}$ definitely applies! Sorry! Also, our approximation means that $|\mathbf{p}_{\text{out}}| = |\mathbf{p}_{\text{in}}|$.

- It is incredibly helpful to know that

$$\mathbf{q}^2 = 4\,|\mathbf{p}_{\text{in}}|^2\sin^2(\theta/2). \tag{141}$$

- The *optical theorem* reads
$$\sigma = \frac{4\pi}{|\mathbf{p}_{\text{in}}|}\Im f(0), \tag{142}$$
with $f(0)$ being the scattering amplitude in forward direction.

- If we have a retarded/advanced Green's function for H_0, we can try to solve $|\Psi\rangle(\mathbf{k})$ as
$$|\Psi\rangle_{\text{out/in}} = |\Phi\rangle + G_0^{\pm}|\Psi\rangle, \tag{143}$$
which is the *Lippmann-Schwinger equation*. Reinserting the RHS into the $|\Psi\rangle$ on the RHS gives an approach to approximating this. Exchanging Ψ with Φ on the RHS is the first step of this and again the Born approximation.

- Expanding $f(\theta, \varphi)$ in spherical harmonics gives the partial wave analysis.

- Poles in the S-matrix at negative energies correspond to bound states, poles close to the negative energy axis in the complex energy plane correspond to resonances. The imaginary part gives a sense of the stability of the resonance – the closer to the real axis, the more stable the resonance.

2.4.5 A Preview of Dirac's equation and the Klein-Gordon equation

Here, I will give a maybe slightly idiosyncratic approach. I *do* give the physical and historical origin of the Klein-Gordon and the Dirac equation, but do *not* treat it as a quantum theory. The reason being three-fold: First, the attempts at giving a language similar to that known in quantum mechanics will not work out (see below). Second, some of the interpretations used to salvage this state of affairs are highly dependent on what particle is looked at. You will find talk about a 'Dirac sea' – but that works only for fermions, so it will not really help for bosons. Third, further attempts to make everything work out then finally end up being just what this summer school wants to teach you and then sometimes uses language I am not too fond of. I consider 'second quantization' to be a confusing term. So *instead*, I will give you a 'classical' approach to both equations. Also, this exposition will differ slightly from what can, for example, be found in [Sre11][45].

But before we get started with each equation, let's have a look at where it originates. The idea is to make the Schrödinger equation
$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \underbrace{\left(\frac{\hat{\mathbf{p}}^2}{2m} + V\right)}_{\hat{H}}|\Psi\rangle \tag{144}$$
agree with the relativistic energy-momentum relation
$$E^2 = p^2c^2 + m^2c^4. \tag{145}$$

[45] There, m doesn't inherit the interpretation of being a mass. But it's just units – and this time we're not building the Mars Climate Orbiter...

Or, for that matter, even just with the idea that in relativity space and time need to have the same standing in a reasonable, relativistic equation. But just looking at (144) tells use that while time enters as a first derivative, space enters at least as a second derivative. Not a good starting point. One, naive, idea is to take $\hat{H}$ as being related to the energy (145), take the square root, and start approximating. Naive, because this makes the space-time-asymmetry of a Schrödinger equation worse, not better, once you insert $\hat{\mathbf{p}} = i\hbar\nabla$. The Klein-Gordon equation and the Dirac equation are now two different approaches getting around this.

Literature-wise, I'd like to refer you to the many quantum-related books and links mentioned earlier, but would also add [Gre00].

Klein-Gordon Equation

- Just square (144) and insert (145)!

$$\begin{aligned} &\Longrightarrow && \left(-\hbar^2\partial_t^2 + \hbar^2c^2\Delta - m^2c^4\right)\text{something} = 0 \\ &\Longleftrightarrow && \left(\Box - \frac{m^2c^2}{\hbar^2}\right)\text{something} = 0 \end{aligned}$$

  In order not to get bogged down by interpretational stuff *yet*, I just said that this new operator should act on 'something'. More on this later.

- Now, before talking about 'something', let's look at what we achieved: Space and time both enter as second order derivatives. Yay!

- But: Nay, it doesn't work as a quantum theory. Why? We do get a continuity equation (nice!)

$$i(\text{something})^\dagger \overset{\leftrightarrow}{\partial_\mu} (\text{something}) = 0, \tag{146}$$

  where $A\overset{\leftrightarrow}{\partial_\mu}B := A\partial_\mu B - (\partial_\mu A)B$, but if we try to give some meaning of probability or probability density to this, we run into a brick wall – no component is ever going to be positive definite. Not what we want for a quantum theory. #sadface

- Forgetting about the problems for a second – can we solve this? Yes! First of all, it is nothing but a variation on the scalar wave-equation, so we can easily adapt what we know from electrodynamics. Another way would be to vigorously wave one's hands, say that in the rest frame the momentum operator corresponds just to ∂_t^2, solve the resulting equation, and then boost to whatever frame of reference we want.

Dirac Equation

The Dirac[46] equation takes a more subtle approach.

[46] Fact of the day: Did you know people used his name as unit for one word per hour? The more you know!

- The idea is to find a linear combination of ∂_μ such that the square of this linear combination is then the Klein-Gordon equation. From the latter we know that it will fulfil the relativistic energy-momentum relation. So we are looking for γ^μ such that
$$\Box = (\gamma^\mu \partial_\mu)^2 \,. \tag{147}$$
- Now this seems rather odd, after all, we also have the m-bit in (145) and that isn't how numbers and roots work. Yes. Numbers. Think of γ^μ rather as matrices! You then get an interesting set of constraints (**Exercise:** Work them out! Take heed of the order, as for matrices AB is not always BA.[47] Also, if feel free to forget about c to make life simpler.)
- As γ^μ appears *very* often contracted with some index down μ, it is customary to introduce the (Feynman-)slash notation:
$$\gamma^\mu C_\mu =: \not{C}. \tag{148}$$
- One can find plenty of matrices that do the trick – especially given that we didn't say how large these matrices have to be. Things that can be done (all possible **exercises**):
  - ▼ Check out the traces of the γ^μ.
  - ▼ The γ^μ have to be even dimensional.
  - ▼ Why can't 2×2-matrices work? Think about what algebraic constraints you have and how many matrices of this size there are that do this.
  - ▼ Find a way to reintroduce m.
  - ▼ Find a way to rewrite (147) as
$$(i\hbar\gamma^\mu\partial_\mu - mc\mathbb{1})\,(\text{something else}) = 0. \tag{149}$$
    Feel free to abbreviate 'something else' as Ψ or whatever.
  - ▼ Check that
$$\gamma^0 = \begin{pmatrix} \mathbb{1}_{2\times 2} & \\ & -\mathbb{1}_{2\times 2} \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} & \sigma_i \\ -\sigma_i & \end{pmatrix} \tag{150}$$
    works. Here, σ_i are the Pauli matrices.
- What happens to the flaw of the Klein-Gordon equation? Set $\overline{\text{something else}} := \gamma^0((\text{something else}))^\dagger$. We can see now that
$$\partial_\mu \left(\overline{(\text{something else})} \gamma^\mu (\text{something else}) \right) = 0, \tag{151}$$
and this time the zero-th component
$$\overline{(\text{something else})}\gamma^0(\text{something else}) = (\text{something else})^\dagger(\text{something else}) \tag{152}$$
is positive definite! So that worked well!

[47] Hint: You should something like $\{\gamma^\mu, \gamma^\nu\} = 2\mathbb{1}_{4\times 4}$[48]

[48] $\{A,B\}$ denotes the anti-commutator, i.e.$\{A,B\} \equiv [A,B]_+ := AB + BA$.

- However, we do get a rather ugly energy spectrum for the solutions (see below) – from $-\infty$ up to ∞. Not being bounded in both directions is asking for trouble, see our cubic potentials from earlier. This is not good, either.#sadface
- While the Klein-Gordon equation was rather simple to solve, the Dirac equation is a bit more involved. For starters, we do not have a model equation like the wave equation that we can easily adapt to it. (Though we can guess that the Dirac equation in a sense has to be 'the root of the wave equation' – whatever that should mean precisely.) Nevertheless, the trick of using the Lorentz covariance to first boost to a rest frame can be adapted to the Dirac equation. It's just a bit more complicated thanks to the matrices γ^μ. For example, the 'four-vector of all matrices γ^μ' is not obviously a four-vector. More on this can be found in [Nik14].

Classical Interpretation as Relativistic Wave Equations

So, what to do about 'something' and 'something else'? My answer would be to think of it as a classical (complex) field. Does not necessarily help with negative energy spectra[49], but certainly helps with not caring about the extra bits of 'quantum philosophy of science' – less physical ideas, less meaning to force onto symbols[50]. Concretely, we do not have to worry about things like probability density interpretations if the field in question is purely classical. Also, a bit on the terminology side/history of science side, it helps avoiding confusion about 'second quantization' – we distinguish quantum field theory from quantum mechanics by saying once we quantize a field (Klein-Gordon, Dirac, Maxwell, Proca, ...) and once particles (harmonic oscillator, hydrogen atom, molecules, ...) rather than confusing people about what comes first and why something is first and something else is second. Maybe also a bit of an opinionated thing from my side.

3 Coda

That all ~~said~~written and done, I hope, I could provide you with a decent overview of what you have done in physics before delving into quantum field theory. If there are errors, don't hesitate to tell me at `sebastian.schuster@msor.vuw.ac.nz`. Quantum field theory is in my humble opinion not so difficult because it is more difficult than any single subject encountered before in physics, but rather, because it quite comfortably makes use of all of it with a flick of a wrist, so getting the basics on a solid foundation is incredibly helpful. I, personally, find this makes up the allure of the subject, too – finally it all comes together for the great finale. Of course, it is (probably) the first time to encounter renormalization techniques in depth, as well, even though they find their use already aplenty in other, much older parts of physics[51]. Hoping that I managed to

[49] For the Klein-Gordon equation it is not quite as bad as for the Dirac equation, see for example p.12ff in [Gre00]. However, this would lead us to far afield.

[50] Part of the reason why I think it useful for physicists to care about philosophy of science. Many physicists might disagree here and quite passionately at that.

[51] That it is an old problem can be seen by the Cauchy principal value which obviously has the goal of getting rid of pesky singularities!

help people with preparing for the school, let's all have a fun, insightful and interesting week full of quantum field theory.

Acknowledgements

This text is dedicated to my father, Manfred, who was at the receiving end of writing-induced forgetfulness. Shame on me. Of course, given the wacky writing, this text would have not been possible without the generous help of other people. For error finding and general feedback, I'd like to thank Matt, Finn, Del, Jessica, Emma and Vee-Liem. This list is likely to grow monotonically over time – all shame, embarrassment and guilt for errors, imprecisions and misunderstandings fall[52] on me.

References

[BD97] W.E. Boyce and R.C. DiPrima. *Elementary Differential Equations and Boundary Value Problems*. John Wiley & Sons, 6th edition, 1997.

[BO99] C.M. Bender and S.A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. Springer-Verlag, 1999.

[Boh89] D. Bohm. *Quantum Theory*. Dover, 1989.

[Bö93] A. Böhm. *Quantum Mechanics – Foundations and Applications*. Springer, third edition, 1993.

[CTDL77] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum Mechanics*. John Wiley & Sons, 1977. 2 Volumes.

[dBG02] R. de la Madrid, A. Bohm, and M. Gadella. Rigged Hilbert Space Treatment of Continuous Spectrum. *Fortschr. Phys.*, 50:185–216, 2002. URL: `https://arxiv.org/abs/quant-ph/0109154`, `doi:10.1002/1521-3978(200203)50:2<185::AID-PROP185>3.0.CO;2-S`.

[de 05] R. de la Madrid. The role of the rigged hilbert space in quantum mechanics. *Eur. J. Phys.*, 26:287–312, 2005. URL: `https://arxiv.org/abs/quant-ph/0502053`, `doi:10.1088/0143-0807/26/2/008`.

[DK10] J.J. Duistermaat and J.A.C. Kolk. *Distributions*. Birkhäuser, 2010.

[Eva10] L.C. Evans. *Partial differential equations*. Number 19 in Graduate studies in mathematics. American Math. Soc., second edition, 2010.

[52] Probably the last footnote[53]: I am sure this has to be plural, given the list, but it just sounds wrong. Is there a special rule that would make 'falls' correct? Fill me in on dirty special rules of English grammar if there is.

[53] Ha! No! There is one final footnote – accompanying alliterations.

[Flü93] S. Flügge. *Rechenmethoden der Quantentheorie*. Springer, fünfte edition, 1993.

[Flü08] S. Flügge. *Practical Quantum Mechanics*. Springer, 2008.

[FS14] E.M. Ferreira and J. Sesma. Global solution of the cubic oscillaton. *J. Phys. A: Math. Theor.*, 47(41):415306, 2014. URL: `https://arxiv.org/abs/1601.02786`, `doi:10.1088/1751-8113/47/41/415306`.

[FW04] T. Fließbach and H. Walliser. *Arbeitsbuch zur Theoretischen Physik*. Elsevier, 2004.

[Gou13] É. Gourgoulhon. *Special Relativity in General Frames*. Springer, 2013.

[Gre00] W. Greiner. *Relativistic Quantum Mechanics: Wave Equations*. Springer, 3rd edition, 2000.

[Gri04] D.J. Griffiths. *Introduction to Elementary Particles*. Wiley-VCH, 2004.

[Gri08] D.J. Griffiths. *Introduction to Electrodynamics*. Pearson, 3rd edition, 2008.

[Gri12] D.J. Griffiths. *Introduction to Quantum Mechanics*. Pearson, 2nd edition, 2012.

[Hab04] R. Haberman. *Applied Differential Equations*. Pearson, fourth edition, 2004.

[HM84] F. Halzen and A.D. Martin. *Quarks and Leptons*. Wiley, New York, 1984.

[Jac75] J.D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, 2nd edition, 1975.

[JS12] J.V. José and E.J. Saletan. *Classical Dynamics - A Contemporary Approach*. Cambridge Univ. Press, 2nd edition, 2012.

[Kri01] M. Kriele. *Spacetime*. Springer, 2001.

[Kuy03] F. Kuypers. *Klassische Mechanik*. Wiley-VCH, 6. edition, 2003.

[Lan99] S. Lang. *Complex Analysis*. Springer, fourth edition, 1999.

[Lan05] S. Lang. *Undergraduate Analysis*. Springer, second edition, 2005.

[LM94] H.B. Lawson and M.-L. Michelsohn. *Spin Geometry*. Princeton University Press, 1994.

[Mar06] B.R. Martin. *Nuclear and Particle Physics*. John Wiley & Sons, 2006.

[Nab10] G.L. Naber. *The Geometry of Minkowski Spacetime*. Springer, second edition, 2010.

[Nik14] H. Nikolic. How (not) to teach lorentz covariance of the dirac equation. *Eur. J. Phys.*, 35:035003, 2014. URL: `https://arxiv.org/abs/1309.7070`, `doi:10.1088/0143-0807/35/3/035003`.

[Pin02] M.A. Pinsky. *Introduction to Fourier Analysis and Wavelets*. AMS, 2002.

[PTVF07] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes*. Cambridge Univ. Press, 3rd edition, 2007.

[Rae16] A.I.M. Rae. *Quantum Mechanics*. CRC Press, 6th edition, 2016.

[SG10] M. Stone and P. Goldbart. *Mathematics for Physics*. Cambridge Univ. Press, 2010.

[Sha14] R. Shankar. *Principles of Quantum Mechanics*. Springer, 2nd edition, 2014. 14th printing.

[SLSS09] M.R. Spiegel, S. Lipschutz, J.J. Schiller, and D. Spellman. *Schaum's Outlines: Complex Variables*. McGraw Hill, 2nd edition, 2009.

[Sre11] M. Srednicki. *Quantum Field Theory*. Cambridge Univ. Press, 2011.

[Tay72] J.R. Taylor. *Scattering Theory: The Quantum Theory on Nonrelativistic Collisions*. John Wiley & Sons, 1972.

[Tay11] M.E. Taylor. *Partial Differential Equations I*. Springer, 2nd edition, 2011.

[Wei13] S. Weinberg. *Lectures on Quantum Mechanics*. Cambridge University Press, 2013.

[Zor04] V.A. Zorich. *Mathematical Analysis II*. Springer, 2004.