Some of the Big Ideas in Physics

Joseph Cunningham

Contents

Chapter 0

Preface

This text is meant to be a succinct and self contained overview of the topics I have come across during my physics studies. The aim is to give an overview of the most important ideas and theories in physics into a relatively short text. It is also my aim to provide a solid theoretical and conceptual basis for these theories.
that is more or less accessible even to people with no background in physics or mathematics. This will almost certainly prove impossible, but I will give it a shot.
It is my view that mathematical notation provides an invaluable tool to support the physical reasoning in a concise way. The notation is often however quite a large barrier for people who have not studied the subject. In an attempt to make this text accessible to everyone I will try to explain all of the mathematics assuming no previous knowledge whatsoever.
It must also be obvious that my explanations and derivations cannot be completely rigorous and fit in a text the size of this one. All I will attempt is to guide the reader through some of the reasoning processes that gave us modern physics, trying to make everything seem plausible along the way. The ultimate goal is to give the reader enough insight so that they could work out the rigorous details on their own.
I doubt very much I will be able to achieve this goal. I will almost certainly err on the side of terseness due to time constraints, but I would be very grateful if anyone reading this text (especially anyone without a mathematical background) would give me feedback on which parts are understandable and which are not. It is not always obvious as the author how clear one's explanations are.
A major motivation for writing this text is to put into words, as clearly as possible, the logical structure of physics. By that I mean present a framework that allows one to trace back the "why" as far as possible; until, that is, we reach the frontiers of mathematical abstraction. I would love to tackle an overview of some of the key concepts in mathematics, but that is most definitely a future project.
Because I am not an expert, I need to rely on people who are. I have therefore borrowed shamelessly from several textbooks, as referenced in the bibliography, and Wikipedia.
TODO! Integrate the following:
In the pages that follow there will be a lot of mathematics; many strange symbols, complex formulae and the like. They all represent ideas, ways of formulating and solving problems to do with physics. I will attempt to formulate here very briefly some of the guiding principles that have been useful for me to understand the subject matter (insofar as I actually understand it, that is) and I will try to expand on these ideas in the body of the text.
Firstly there is the question of how much one needs to know to adequately understand a topic.

I feel the professor's answer to this is slightly different from mine. I would generally like slightly more mathematical background. I have added some material to that effect, but it still remains a very cursory introduction to the topic of Lie groups, Lie algebras and representations. My guiding principle when adding the material has been the question of "why". Why introduce all these concepts? Why do the calculations we are doing? I have tried to make the story more or less a coherent one, to make the definitions seem not too unnatural and to make the theorems at least plausible. I do however sometimes not observe strict mathematical rigour in favour of a more intuitive approach, in the hopes that the intuition enables the reader to work out the mathematical details if necessary.
Another thing I would like to touch one briefly here is the difference between a definition and a notion in mathematics. I will introduce several definitions for Lie groups and Lie algebras. Which, while quite different, all refer to the same idea or notion of this concept we call Lie groups and algebras. Hopefully the different definitions and the different angles at which the topic is approached will serve to strengthen the intuition of anybody foolish enough to read this text.
TODO Carnappian reconstruction; axioms: In discussing axiomatic systems several properties are often focused on:[6]
The axioms of an axiomatic system are said to be consistent if no logical contradiction can be derived from them. Except in the simplest systems, consistency is a difficult property to establish in an axiomatic system. On the other hand, if a model exists for the axiomatic system, then any contradiction derivable in the system is also derivable in the model, and the axiomatic system is as consistent as any system in which the model belongs. This property (having a model) is referred to as relative consistency or model consistency. An axiom is called independent if it can not be proved or disproved from the other axioms of the axiomatic system. An axiomatic system is said to be independent if each of its axioms is independent. If a true statement is a logical consequence of an axiomatic system, then it will be a true statement in every model of that system. To prove that an axiom is independent of the remaining axioms of the system, it is sufficient to find two models of the remaining axioms, for which the axiom is a true statement in one and a false statement in the other. Independence is not always a desirable property from a pedagogical viewpoint. An axiomatic system is called complete if every statement expressible in the terms of the system is either provable or has a provable negation. Another way to state this is that no independent statement can be added to a complete axiomatic system which is consistent with axioms of that system. An axiomatic system is categorical if any two models of the system are isomorphic (essentially, there is only one model for the system). A categorical system is necessarily complete, but completeness does not imply categoricity. In some situations categoricity is not a desirable property, since categorical axiomatic systems can not be generalized. For instance, the value of the axiomatic system for group theory is that it is not categorical, so proving a result in group theory means that the result is valid in all the different models for group theory and one doesn't have to reprove the result in each of the non-isomorphic models.
TODO: model theory!!

Chapter 1

Introduction: why the big ideas?

TODO + counter examples

Chapter 2

What is physics?

Physics is of course a science. Generally the goal of science is to gain an understanding and make predictions about how the world works. This understanding will not just appear in our minds unprovoked and reason alone will not provide all the answers. Consequently we must interact with the real world in order to get information about it. Experiments are such interactions.
Having done an experiment, we have information about the world. A priori this information is not very useful. It says something about the world when the experiment took place. That world is no longer. We must hypothesise that the world is not completely random and thus that it does follow some underlying rules. Unfortunately we cannot directly verify any underlying rule because we can only do so many experiments and for any set of experimental results there are many rules that can explain them. For example, if we drop a ball a hundred times and it always falls down to the ground, this may indicate that it will always fall down. It may however be the case that the hundred and first time it falls upwards. We can easily prove this wrong however by dropping it another time and seeing that it falls down. We can still not be sure however that it will always fall down. This leads us to the important insight that a theory can be proven wrong, but never right. In fact a theory is *only really worth considering* if it makes predictions we can try to prove wrong. In that case we call it falsifiable.
These ruminations lead us to view the scientific process as roughly the following:

1. First we **observe** some phenomenon.
2. This inspires us to **hypothesise** an explanation (often based on and extending theories that have previously been hypothesised and tested).
3. We then **make predictions** based on our newly hypothesised theory.
4. Finally we **test** these predictions. After enough predictions have been tested we accept that the hypothesis is probably true.

Having talked a bit about science in general, we now turn to physics in particular. While it's mostly convention which parts of scientific inquiry fall under the label physics, we can generally say that physics is in the business of developing theories and principles that are true for everything in the world. So for example whether you are dropping a ball or a cat, physics provides laws of motion that explain both trajectories. This is in contrast to e.g. biology that can only say something about the behaviour of the cat and what they would do in such a situation.

2.1 What is a formula? Why mathematics?

We need a way to share our ideas about the world. Debate is necessary to further scientific knowledge. No one person can understand all of physics on their own; even if this was possible, that person would probably still need to write things down to order their own thoughts. We physiologically do not have the brain capacity to actively hold all physics concepts in our mind at the same time.
With mathematics a system has been developed that helps us write down our reasoning process in a clear and concise manner. It can be seen as a kind of shorthand that allows us to write down our reasoning in steps so we don't need to keep everything constantly in mind.
A formula is the mathematical formulation of some fact we believe to have uncovered about the universe. It is a crystallisation of our knowledge of the world formulated using mathematics.
In this way formulae can be seen almost as an illustration of a concept. In fact illustrations can provide much of the same illumination as formulae and can even serve to provide additional understanding if the standard notation is too unwieldy. Judicious use of illustrations in this document will hopefully aid in understanding.

Example

As an illustration of the utility of formulae, we will consider the following problem:
Farmer John has 200 metres of fencing that he can use to fence off an area of one of his fields. He would like to know what the biggest possible surface area is he can fence off, if the fenced off area has to be a rectangle.
We will use some mathematical notation and procedures without much justification here. They will be clarified later; this example will partially help motivate the future discussion.
As will often be the case, the first step is to give names to things. So, for instance, we might call the length of the field l and the width w. The main advantage of this is, when considering complex lines of reasoning, we can take in and process these single letter symbols much faster that always having to reference "the length of the field". Thus we can formulate the main constraint of the problem quite simply and very compactly as

$$2l + 2w = 200\,\mathrm{m}$$

Where we write $2l$ and $2w$ because the perimeter of a rectangle 2 times the length plus 2 times the width. This means that the width also equals 200 metres minus 2 times the length, all divided by two

$$w = \frac{200\,\mathrm{m} - 2l}{2} = 100\,\mathrm{m} - l$$

TODO: image x2
In order to get the surface area (which we call S) of the field, we multiply the length by the width

$$S = l \cdot w = (100\,\mathrm{m}) \cdot l - l^2$$

Now the power of the abstract notation and reasoning really gets to shine: We are in effect trying to maximise the value of S and if we had properly introduced the mathematical machinery of calculus we would know that we can do that by setting the derivative of S to zero.

$$100\,\mathrm{m} - 2l = 0 \qquad \text{from which follows that} \qquad l = 50\,\mathrm{m}$$

For now we will find a different way to reach the same conclusion, because we have not yet developed the requisite calculus.

There are undoubtedly many ways to solve this problem. Here we will use an approach that is in some sense variational in nature (in that we will look at an expression for the surface area around the maximum). In fact we will prove that regardless of the actual length of the fenching, for a problem like this a square is the best shape. A proof might go as follows:
For a given length of fencing, a square can always be constructed. This square has a particular length that we call a for convenience and a surface area a^2. If we now look at other rectangles with the same circumference, they will have a side that is longer (we can say its length is $a + x$ for some positive x) and a side that is shorter. The shorter side must be shorter by the same amount that the longer side is longer by (in order for the total length of fencing to stay constant). So the shorter side has length $a - x$. See figure ref TODO We can now just work out the surface area:

$$S = (a + x) \cdot (a - x) = a^2 - x^2$$

This means that for any nonzero x the surface area is smaller than that of the square and thus the square is the best shape.
A fortiori we conclude that in our particular case $l = 50\,\mathrm{m}$ is optimal as this gives us a square.

This example also exemplifies the general aesthetic we will be going for: Mostly text punctuated with formulae and calculations to provide clarity and set out facts for future reference.

Part I

Some experimental notes

Chapter 1

Units

TODO flow Now the (imperial) system of measurements used in the example above is not so very useful for more complex an precise measurements. A first obvious problem is that not all barleycorns have the same length. This can of course be easily remedied by choosing one particular barleycorn as the standard. If many people are supposed to use this system of measurement, they cannot all have access to this special barleycorn. One way to get around this is by letting people copy the length of the barleycorn, in effect making rulers based on this barleycorn.

Another solution is to describe the unit of measurement in such a way that anybody can make there own ruler. This idea appeared in the seventeenth century and was known as the universal measure or μέτρον καθολικόν (métron katholikón) in Greek. (Hence the name metre). One proposal for such a system was the seconds pendulum, which is a pendulum with a period of exactly two seconds. As we will see, pendulums with the same period have the same length if the effect of gravity stays constant, independent of the attached weight. Unfortunately gravity is not equally strong everywhere (and the second was not undisputed as unit of time at the time) and thus it was decided that the new measure should be one ten-millionth of the distance from the North Pole to the Equator, measured along the meridian through Paris. This required the distance to be carefully measured, which was not an easy task at the time. A platinum bar, the mètre des Archives, was manufactured based on the measurements to serve as the prototype for all other metres. In the following years new surveys revealed that the mètre des Archives was not exactly correct, but for practical reasons it was decided in 1867 that the official length of the should be the length of the mètre des Archives. This meant we were back to being stuck with a definition of our units based on an arbitrary object, except now it was a platinum-iridium bar instead of a barleycorn. It was only in the twentieth century that a true universal definition of the metre was adopted (first using the wavelength of light emitted by krypton-86 and later as the length traveled by light in one 299792458th of a second).

While the project of the metre initially failed to provide a universal measure, it did introduce another useful innovation: it was a decimal system. In previous systems if the lengths you were measuring were too big or too small to be easily measured in your unit of choice, you would pick another unit that was often a quite arbitrary multiple of the first unit. So for example you might measure the length of people in feet, but the length of voyages in miles (which is 5280 feet). With the metric system all units derived from the metre differ by multiplication of a power of ten. Thus, e.g., a kilometre (km) is a thousand (which is 10^3) metres. A more extensive list can be found in table 1.1. This made it much easier to quantify very big and very small things, because no matter the scale, there was a logical system of units to work with.

Prefix name	Prefix symbol	Power of Ten	Decimal factor
yotta	Y	10^{24}	1000000000000000000000000
zetta	Z	10^{21}	1000000000000000000000
exa	E	10^{18}	1000000000000000000
peta	P	10^{15}	1000000000000000
tera	T	10^{12}	1000000000000
giga	G	10^{9}	1000000000
mega	M	10^{6}	1000000
kilo	k	10^{3}	1000
hecto	h	10^{2}	100
deca	da	10^{1}	10
		10^{0}	1
deci	d	10^{-1}	0.1
centi	c	10^{-2}	0.01
milli	m	10^{-3}	0.001
micro	μ	10^{-6}	0.000001
nano	n	10^{-9}	0.000000001
pico	p	10^{-12}	0.000000000001
femto	f	10^{-15}	0.000000000000001
atto	a	10^{-18}	0.000000000000000001
zepto	z	10^{-21}	0.000000000000000000001
yocto	y	10^{-24}	0.000000000000000000000001

Table 1.1: Prefixes for the International System of Units

When writing these units down, we often do not want to write out the full unit. Every SI unit has an abbreviation, formed by concatenating the prefix symbol with the abbreviation of the unit, thus we can write pm instead of picometres or ms instead of milliseconds.

1.1 Base units

So far we have mainly talked about units of length, with a brief mention of units of time in passing. Of course we will want to measure things other than just length and time, like temperature, mass, speed, electric current, electric charge, force, pressure etc. It would seem that we need to define units for all these quantities, preferably with universal definitions independent of and physical reference object. Luckily we do not need universal definitions for all these units. For example speed can be expressed as a length traveled per unit of time, thus we can use the unit meters per second (m/s). It turns out we can express many units in terms of other units so that we only need universal definitions for 7 units, the so-called SI base units. These units are listed in table 1.2, along their symbols and symbols referring to the quantity they measure (the "dimension symbols").
As of May 2019 all the base units have universal definitions.
It is also important to recognise that the choice of which units are base units is quite arbitrary. We could just as well have chosen speed to be a base units and defined a unit of length based on speed times time. Particularly it might be strange that the ampere is a base units seeing as electric current is electric charge per unit of time. It might seem more intuitive to make coulomb (the SI unit of electric charge) a base unit and define ampere as coulomb per second. The reason this was not done is because historically electric current was easier to measure than

Unit name	Unit symbol	Dimension symbol	Quantity name
second	s	T	time
metre	m	L	length
kilogram	kg	M	mass
ampere	A	I	electric current
kelvin	K	Θ	temperature
mole	mol	N	amount of substance
candela	cd	J	luminous intensity

Table 1.2: SI base units

electric charge.
The many other SI units based on these base units will be introduced as necessary in this text.

1.2 Geometrized units

TODO $c = G = 1$ Naturalness of numbers and implications for multiplication
TODO See also Mandl, Shaw 6.1

1.2.1 Conventions: two fundamental constants

$$c = 2,957 \times 10^{8} \text{ms}^{-1} \qquad \rightarrow \qquad \text{speed of light}$$
$$\hbar \equiv \frac{h}{2\pi} = 6,582 \times 10^{-12} \text{MeV s} \qquad \rightarrow \qquad \text{Planck constant}$$

We set $\hbar = c = 1$ (adimensional), which means we will be working in natural units. This has the advantage of giving simpler expressions, but we lose explicit dimensionality.

1.2.2 Assumptions underlying natural units

$$c = 1 \quad \rightarrow \quad [V] = \frac{[L]}{[T]} = \text{dimensionless} \qquad \rightarrow \qquad [L] = [T]$$
$$\hbar = 1 \quad \rightarrow \quad [S] = \frac{[M][L]^{\cancel{2}}}{\cancel{[T]}} = \text{dimemsionless} \qquad \rightarrow \qquad [M] = [L]^{-1}$$
$$\Rightarrow [M] = [L]^{-1} = [T]^{-1}$$

1.2.3 Dimensions of energy and momentum in natural units

$$[E] = [M][V]^2 \qquad \rightarrow \qquad [E] = [M]$$
$$[P] = [M][V] \qquad \rightarrow \qquad [P] = [M]$$
$$[M] = c\text{eV}$$

Chapter 2

Rounding and precision

significant figures scientific notation accuracy vs precision order of magnitude dimensions and dimensional analysis (+unit conversion)

Part II

Newtonian mechanics

Chapter 1

Setting the stage

1.1 What is mechanics?

- Describe motion of objects.
- Traditionally kinematics, dynamics and statics (Under what condition is there no motion? Important when we do not want motion such as in constructions.).

Later different formalisms: energy, action. Derived from Newton.

1.2 Some experimental facts

1.2.1 Space and time

- continuous trajectories -¿ topology
- straight lines -¿ affine structure
- distance and time intervals -¿ metric structure

Absolute? Bucket experiment. (¿¡ Mach).

$$\mathbb{E}^3 \times \mathbb{R}$$

1.2.2 Galileo's principle of relativity

Ability to switch coordinate systems. If not absolute, then equivalence classes of coordinate systems.
We cannot give an example of an inertial frame, only approximants.
Alternatively: define inertial frames as frames where first law is true. (The law asserts their existence.)

1.2.3 Determinacy

Laplace's demon. Given position and velocity we can determine further evolution of the system.

1.2.4 Divisibility

1.3 Using point particles

1.4 Newton's laws

Posits absolute space and time. Let $\mathbf{r}(t)$ be a position vector.

$$\mathbf{a} = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}$$

1. Every body perseveres in its state either of rest or of uniform motion in a straight line, except insofar as it is compelled to change its state by impressed forces.
2. The change in motion is proportional to the impressed motive force and is made along the straight line on which the force is impressed.

$$\mathbf{F} = m\mathbf{a}.$$

3. Action / reaction: less fundamental.

In classical mechanics the motion of massive particles is determined by Newton's laws. They are as follows:

1. When all external influences on a particles are removed, the particle moves with constant velocity.
2. When a force $\mathbf{F}$ acts on a particle of mass m, the particle moves with instantaneous acceleration $\mathbf{a}$ given by the formula

$$\mathbf{F} = m\mathbf{a}.$$

3. When two particles exert forces upon each other, these forces are
   (a) equal in magnitude,
   (b) opposite in direction, and
   (c) parallel to the straight line joining the two particles.

TODO: inertia, mass and force
Inertial frame: first law holds
law of multiple interactions
principle of superposition!

Chapter 2

Setting up the mathematics

Remarks on intervals and integration. Over time. (What is $\mathrm{d}l(t)$?)

Chapter 3

Kinematics

TODO apples, cars and pendula
Before trying to predict motion, we first need to introduce the concepts necessary to *describe* motion. This is the domain of kinematics. To describe the motion, i.e. the trajectory, of an object (say an apple, for example) through space, it would seem quite natural to use a three dimensional curve. There is however a slight problem with this approach: a curve is infinitesimally thin, whereas an apple has a volume. This means if you track different parts of the apple, you will get different trajectories. This is especially obvious if the apple is spinning, see picture TODO ref. Intuitively we would like to call curve TODO the trajectory of the apple, but how do we know which part of the apple we need to track to get such a curve? In essence we want to track to point that the apple is spinning around, that way we have a trajectory of the apple (its *translation*) without any information about its *rotation*, which we can then consider separately. We will show how to find the point the apple is spinning around later, but for now in order to build up the theory, we will solve the problem in a different way: We model the motion of the apple with a particle with the same mass that is infinitely small. We define a particle as an idealised body that occupies only a single point in space. We call a particle with a mass a point mass. We now do not need to find the point the apple is rotating around, because the point mass is infinitely small and thus only has one point to rotate around.
TODO picture of cycloidal apple. TODO: change of frame:

$$\delta l = \delta u + \boldsymbol{\delta} \times (\mathbf{r}_A - \mathbf{r}_B)$$

TODO: change of frame to express constraints!!

3.1 Speed, velocity and acceleration

Many of the fundamental concepts of kinematics have been touched upon in the section on spatial curves. We have seen that we can take the derivative of a curve (at least if the curve is differentiable, but we can safely require that trajectories be modelled with differentiable curves). This derivative is, as we have seen, a vector that points in the direction tangential to the motion and its magnitude gives us an indication of how far the particle is travelling per unit of time. We call the derivative the velocity. Sometimes we are only interested in how far the particle is travelling per unit of time, not which direction it is going in. In that case we would only be interested in the magnitude of the derivative, which we call the speed.

Now we could also be interested in how the velocity changes over time. Velocity is a time-dependent vector; taking the derivative of it we get a new time-dependent vector, which we call the acceleration. This is the second derivative of the position vector.
We can also take higher order derivatives of the position vector. The third derivative is known as the jolt. The fourth, fifth and sixth as jounce (or snap), crackle and pop. These higher order derivatives rarely appear in any practical applications.
In full generality there is not much more to say on the topic of particle kinematics. If we can show that a trajectory is of a certain type (e.g. linear or circular), this also imposes limits on the form of the derivatives. The mathematical expressions are also often a lot easier to deal with. We give some important examples of such systems below.

3.2 Describing the motion of a particle

In classical mechanics we assume that space is Euclidean and three dimensional. Thus we can choose an origin O and three orthogonal vectors, $\hat{i}, \hat{j}, \hat{k}$. Together these form a reference frame. Locations can be specified with a position vector $\mathbf{r}$, relative to the reference frame. The motion of a particle can be specified as a curve in the Euclidean space, parametrised by time t. So we view $\mathbf{r} : \mathbb{R} \to \mathrm{E} : t \mapsto \mathbf{r}(t)$ as giving the location of a particle in function of time.
We then define the velocity $\mathbf{v}$ and acceleration $\mathbf{a}$ of a particle whose motion is described by $\mathbf{r}(t)$ as

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \qquad \text{and} \qquad \mathbf{a} = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}$$

3.2.1 Interpreting velocity and acceleration

Assume $s(t)$ is the arc length of a trajectory $\mathbf{r}(t)$, so the arc length parametrisation is given by $\mathbf{r}(s)$. Then by the chain rule,

$$\mathbf{v} = \tfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \tfrac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\tfrac{\mathrm{d}s}{\mathrm{d}t} = \mathbf{T}v$$

where $\mathbf{T}$ is the unit vector tangent to the trajectory and $v = \frac{\mathrm{d}s}{\mathrm{d}t}$ is the speed.
We can write the acceleration as follows:

$$\begin{aligned}\mathbf{a} &= \tfrac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \tfrac{\mathrm{d}v\mathbf{T}}{\mathrm{d}t} = \tfrac{\mathrm{d}v}{\mathrm{d}t}\mathbf{T} + v\tfrac{\mathrm{d}\mathbf{T}}{\mathrm{d}t} \\ &= \left(\tfrac{\mathrm{d}v}{\mathrm{d}t}\right)\mathbf{T} + \left(v^2\kappa\right)\mathbf{N} \\ &= \left(\tfrac{\mathrm{d}v}{\mathrm{d}t}\right)\mathbf{T} + \left(\frac{v^2}{R}\right)\mathbf{N}\end{aligned}$$

as per the Frenet-Serret formulae in three dimensions; R is the radius of the osculating circle. We have split the acceleration in a component tangent to the trajectory that shows the instantaneous change in speed and a component perpendicular to the local path direction that shows how much the trajectory is curving and points in the direction the trajectory is curving towards.

3.2.2 Polar coordinates

There are other choices of basis of our reference frame we can make. If the trajectory $\mathbf{r}(t)$ is in a plane, then we can choose two basis vectors in the plane $\hat{r}, \hat{\theta}$ such that at anytime $\hat{r}$ is the unit vector in the direction $\mathbf{r}(t)$. Obviously these basisvectors are not constant, in particular

they depend on the location $\mathbf{r}$. It is useful to express this location in polar coordinates (r, θ). Now we can write

$$\begin{cases} \frac{\mathrm{d}\hat{r}}{\mathrm{d}r} = 0 & \frac{\mathrm{d}\hat{r}}{\mathrm{d}\theta} = \hat{\theta} \\ \frac{\mathrm{d}\hat{\theta}}{\mathrm{d}r} = 0 & \frac{\mathrm{d}\hat{\theta}}{\mathrm{d}\theta} = -\hat{r} \end{cases}$$

In terms of these polar coordinates we can also write expressions for the velocity and acceleration:

$$\begin{cases} \mathbf{r} = r\hat{r} \\ \mathbf{v} = \dot{r}\hat{r} + (r\dot{\theta})\hat{\theta} \\ \mathbf{a} = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{r} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\hat{\theta} \end{cases}$$

3.3 Rectilinear motion

If the trajectory of a particle is restricted to a straight line, we do not even need vectors to describe locations, velocities or accelerations etc.
We can always choose a reference frame such that the position vector is of the form

$$\mathbf{r}(t) = x(t)\hat{i}.$$

Where x is a real number that depends on time.
We can define and describe some particular forms of rectilinear motion (a_0, v_0, x_0, t_0 are constants denoting the initial acceleration, velocity, position and time, respectively):

1. **Uniform rectilinear motion** is when the particle is moving at constant velocity.

$$\begin{cases} a(t) = 0 \\ v(t) = v_0 \\ x(t) = v_0(t - t_0) + x_0 \end{cases}$$

2. For rectilinear motion with **constant acceleration**, we get the following:

$$\begin{cases} a(t) = a_0 \\ v(t) = a_0(t - t_0) + v_0 \\ x(t) = \frac{a_0}{2}(t - t_0)^2 + v_0(t - t0) + x_0 \end{cases}$$

TODO: combining these cases (projectile motion).

3.4 Circular motion

3.4.1 Uniform circular motion

TODO image.
In this case a particle is traveling along a circular trajectory with radius R at a constant speed v. We choose a reference frame such that the origin is the centre of the circle and the trajectory is in the plane spanned by $\hat{i}, \hat{j}$. The arc length traveled in time t is vt. In terms of the angle θ the arc length is $R\theta$, so we have

$$\theta = \frac{vt}{R} = \omega t$$

where θ is called the angular velocity. In general it is defined for any circular motion as $\frac{\mathrm{d}\theta}{\mathrm{d}t}$. We can then write

$$\begin{cases} \mathbf{r} = R\cos(\omega t)\hat{i} + R\sin(\omega t)\hat{j} \\ \mathbf{v} = -v\sin(\omega t)\hat{i} + v\cos(\omega t)\hat{j} \\ \mathbf{a} = -\frac{v^2}{R}\cos(\omega t)\hat{i} - \frac{v^2}{R}\sin(\omega t)\hat{j} \end{cases}$$

Alternatively we can work in polar coordinates, in which case

$$\begin{cases} \mathbf{r} = R\hat{r} \\ \mathbf{v} = v\hat{\theta} \\ \mathbf{a} = -\left(\frac{v^2}{R}\right)\hat{r} \end{cases}$$

3.4.2 General circular motion

In general circular motion can be described using polar coordinates as follows:

$$\begin{cases} \mathbf{r} = R\hat{r} \\ \mathbf{v} = v\hat{\theta} \\ \mathbf{a} = -\left(\frac{v^2}{R}\right)\hat{r} + \dot{v}\hat{\theta} \end{cases}$$

TODO centripetal force and acceleration (radial component).

3.5 Rigid body in planar motion

TODO?

Chapter 4

Dynamics

4.1 Some forces to get us started

Some forces traditionally associated with mechanics.

4.1.1 Internal forces in a body

4.1.2 Gravity

4.1.2.1 Universal law of gravitation

Gravity of spherical shells

Gravity of a sphere

4.1.2.2 Approximation on the earth's surface.

4.1.3 Normal force

4.1.4 Elasticity: Hooke's law

4.1.5 Friction and drag

4.2 Net forces and free-body diagrams

4.3 General principles for solving problems in dynamics

Chapter 5

Statics

1. Force balance
2. Torque balance

Chapter 6

Conservation principles

- Homogeneity of spacetime → translational invariance → conservation of momentum (conservation of p^μ)
- Isotropicity of spacetime → rotational invariance → conservation of $\mathbf{L}$

6.1 The energy principle

6.1.1 Work and power

For now only definition.

$$W = \int \mathbf{F} \cdot \mathrm{d}\mathbf{l}$$

$$P(t) = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}}) \cdot \dot{\mathbf{r}}??$$

$$W = \int_I P(t)\,\mathrm{d}t$$

6.1.2 Conservative forces

Link work to potential energy.
A force is conservative if the work done by the force over a closed trajectory is zero. This is equivalent to saying there is some $U(\mathbf{r})$, called the potential, such that

$$\mathbf{F}(\mathbf{r}) = -\nabla \cdot U(\mathbf{r}).$$

In this case the work done by the force over a trajectory only depends on the initial and final point, not the whole trajectory:

$$W = \int -\nabla \cdot U \cdot \mathrm{d}\mathbf{l} = U(\mathbf{a}) - U(\mathbf{b}) = -\Delta U$$

6.1.3 Kinetic energy

We define the kinetic energy

$$K = \frac{m\mathbf{v} \cdot \mathbf{v}}{2} = \frac{mv^2}{2}.$$

Taking the derivative with respect to time, we get

$$\frac{\mathrm{d}K}{\mathrm{d}t} = m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \cdot \mathbf{v} = m\mathbf{a} \cdot \mathbf{v} = \mathbf{F} \cdot \mathbf{v}.$$

Integrating over the time interval

$$K_2 - K_1 = \int_{t_1}^{t_2} \mathbf{F} \cdot \mathbf{v}\, \mathrm{d}t = \int \mathbf{F} \cdot \mathrm{d}\mathbf{l} = W$$

Work-energy principle.

6.1.4 Conservation of energy

For conservative forces.

$$\begin{aligned} W &= -\Delta U \\ &= \Delta K \end{aligned}$$

Thus $\Delta K + \Delta U = 0$ or

$$E_1 \equiv K_1 + U_1 = K_2 + U_2 \equiv E_2$$

TODO Gravitational potential. Newton's theorem. (See Hartle); TODO multiple forces.

6.2 The linear momentum principle

$$\mathbf{F}_{\text{tot}} = \sum \mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$$

$$\mathbf{F}_{\text{ext}} = \sum \mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$$

6.3 The angular momentum principle

Chapter 7

Rigid body mechanics

7.1 Rigid body kinematics

7.2 Rigid body dynamics

TODO: torque alignment example

Chapter 8

Small oscillations

Chapter 9

Applications

9.1 Some simple systems

9.2 Harmonic (linear) oscillators

90 percent of physics.
Potential around minimum: constant term arbitrary, first order zero (min), second order most important.
Prototype for small vibrations about stable equilibrium

9.3 Non-linear oscillations and phase space

9.4 Orbits in central field

TODO some history + elaborate whole section
Newton proves Kepler's laws

Kepler's laws

First law Each of the planets moves on an elliptical path with the sun at one focus of the ellipse.

Second law For each of the planets, the straight line connecting the planet to the sun sweeps out equal areas in equal times.

Third law The squares of the periods of the planets are proportional to the cubes of the major axes of their orbits.

one-body problem: central field

A force field $\mathbf{F}(\mathbf{r})$ is said to be a central field with centre $\mathcal{O}$ if it has the form

$$\mathbf{F}(\mathbf{r}) = F(r)\hat{r}$$

where $r = |\mathbf{r}|$ and $\hat{r} = \mathbf{r}/r$. A central field is thus spherically symmetric about its centre.

9.4.1 Newton's equations in a central field

Firstly we must remark that any orbit of a particle P in a central field with centre $\mathcal{O}$ must lie in a plane through $\mathcal{O}$. This can easily be seen on grounds of symmetry. Mirroring the whole system across this plane, we see that it is unchanged. This means that were there any effect causing the particle to leave the plane one one side, a similar effect would exist to cause it to leave on the other side, clearly a contradiction.
This means all motion is two-dimensional and we can use the polar coordinates r, θ (centred on $\mathcal{O}$) to specify the location of P at all times.
Using the general form of the acceleration in polar coordinates, we can write the Newtonian equations of motion as follows:

$$m\left(\ddot{r} - r\dot{\theta}^2\right) = F(r)$$
$$m\left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right) = 0$$

9.4.1.1 Angular momentum conservation

The second of these equations leads to a conservation law, as it can be rewritten:

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}\left(mr^2\dot{\theta}\right) = 0$$

which can be integrated with respect to t to get

$$mr^2\dot{\theta} = \text{constant} \equiv L$$

We call this quantity the angular momentum. We will define it properly in a subsequent section. For now we just need to know that it is conserved.
In fact angular momentum conservation is equivalent to Kepler's second law. The area A referenced in the law can be calculated as follows (TODO: why):

$$A = \frac{1}{2}\int_0^\theta r^2\,\mathrm{d}\theta.$$

By the chain rule, we have

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\mathrm{d}A}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{1}{2}r^2\dot{\theta} = \frac{L}{2m}.$$

So, assuming the mass m is constant, this means the derivative of A with respect to t is constant and so A depends only on time intervals.

9.4.1.2 Newton's equations in specific form

It will prove useful to eliminate the mass m from the theory. To do that we define the outward force per unit mass $f(r)$, such that

$$F(r) = mf(r).$$

We also define L to be the angular momentum per unit mass:

$$l \equiv \frac{L}{m} = r^2\dot{\theta}.$$

Newton's equations in specific form are then

$$\ddot{r} - r\dot{\theta}^2 = f(r)$$
$$r^2\dot{\theta} = L$$

9.4.1.3 Energy conservation

Every central field is conservative. We define $V(r)$ to be the potential per unit mass:

$$f(r) = -\frac{\mathrm{d}V}{\mathrm{d}r}.$$

We can now write down an equation expressing conservation of energy per unit mass, E:

$$\frac{1}{2}\left(\dot{r}^2 + (r\dot{\theta})^2\right) + V(r) = E$$

This can be rewritten using angular momentum:

$$\begin{aligned} E &= \frac{1}{2}\dot{r}^2 + V(r) + \frac{l^2}{2r^2} \\ &= \frac{1}{2}\dot{r}^2 + V^*(r) \end{aligned}$$

Where $V^*(r) \equiv V(r) + \frac{l^2}{2r^2}$ is the effective potential. We call this equation the radial motion equation.

9.4.2 Describing orbits

The radial motion equation is not analytically solvable for an arbitrary potential, but we can still make some qualitative remarks.
Firstly, of course, we must have

$$V^*(r) \leq E,$$

with equality when $\dot{r} = 0$.
Now we can make a distinction between two types of motion (TODO image):

1. **Bounded motion** TODO

9.4.3 The path equation

9.4.4 Nearly circular orbits

9.4.5 The attractive inverse square field

9.4.6 Space travel - Hohmann transfer orbits

9.5 Rutherford scattering

9.5.1 The experiment

9.5.2 The repulsive inverse square field

9.5.3 Cross-sections

Part III

Analytical mechanics

Chapter 1

Motivation and overview

In this chapter we will develop a new way to do classical mechanics. This new method does not make use of forces and is simple to apply, even to quite complex mechanical problems. Conceptually it can be thought to be based on a generalisation of two tricks we have already encountered to simplify problems in mechanics: using conservation laws such as conservation of energy and momentum and changing coordinates to more suitable ones. These tricks have so far only been applied on an ad hoc basis. There was no guaranty either could be applied.
This new approach is interesting in particular for a couple of reasons. First, there is a natural way to deal with constraints constraints. Second this new way of doing things is much more easily extended to quantum mechanics and quantum field theory.
In fact there are two formulations of this new theory: one due to Lagrange and one due to Hamilton.
The chapter will be structured as follows:

- First we will derive the Lagrangian version of this new theory from Newtonian mechanics using D'Alembert's principle.
- Next the Lagrangian equations will be derived from a different principle: that of least action. What this essentially means, is that the systems evolve in the "easiest possible way". A similar idea will provide the basis for general relativity. This idea is also much more general than any particular set of rules determining how to manipulate vectorial forces, and thus is much more generally extended to the quantum realm.
- In order to do that some more mathematics, the calculus of variations, is needed.
- Next we will consider the second, Hamiltonian, formulation of the theory. This reformulates the theory, using a Legendre transform, such that we get a system of first order differential equations. For practical problems, this is usually not very useful, but it allows some theoretical results to be derived more easily. It is also this version of the theory that will be most relevant for quantum mechanics.
- Finally we will be in a position to explore one of the most beautiful ideas in physics: that there is a link between symmetries and conserved quantities. This is formally encapsulated in Noether's theorem.

Chapter 2

A first derivation of the Euler-Lagrange equation

2.1 Constraints and configuration space

In many interesting mechanics problems we have to deal with constraints. Examples include:

1. The bob of a pendulum must remain a fixed distance from the point of support;
2. The particles of a rigid body must maintain fixed distances from each other;
3. The contact particle of a body rolling on a fixed surface must be at rest.

When solving the Newtonian equations of motion for a particle

$$m\dot{\mathbf{v}} = \mathbf{F},$$

we enforced the constraints by adding a constraint force.

$$\mathbf{F} = \mathbf{F}^S + \mathbf{F}^C$$

With $\mathbf{F}^S$ the specified force (the force on the particle if there were no constraint) and $\mathbf{F}^C$ the constraint force. Unfortunately this constraint force is variable, being just enough to counteract any force pulling the particle out of its constraint. There is no easy way to calculate the constraint force, it gives additional conditions that need to be solved simultaneously with the dynamical equations.
We have already encountered one way to deal with constraints when we discussed circular motion: using polar coordinates. The constraint then becomes that the radial component of the position vector is constant. Thus we are only interested in the tangential component of all the quantities involved and the radial component is no longer relevant.
We only need one coordinate to specify the configuration of our system.

2.1.1 Generalised coordinates

TODO: after equations
N particles with index i n degrees of freedom $\mathcal{Q}$
Holonomic + see Goldstein

2.1.2 Generalised velocities

$$\mathbf{v}_i = \frac{\partial \mathbf{r}_i}{\partial q_1}\dot{q}_1 + \ldots + \frac{\partial \mathbf{r}_i}{\partial q_n}\dot{q}_n = \sum_{j=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_j}\dot{q}_j \tag{2.1}$$

(note on partial derivatives and formalism subsubsection?)

2.2 D'Alembert's principle

We can now specify any possible configuration of our system using generalised coordinates. We still do not know what the constraint forces are and how they impact the specified forces.
One thing we can say however, is that the constraint force must at all times be perpendicular to the paths that a particle is constrained to. If there were a component of the constraint force in a direction the particle could move in, it would feel a force in that direction. Thus the constraint can not be geometrical (or equivalently geometrical) in nature.
Because $\mathbf{v}$ points in the direction the particle is traveling in, we can write

$$\mathbf{F}^C \cdot \mathbf{v} = 0.$$

In other words, the constraint force does no work. Actually we can obtain an even stronger result: because $\mathbf{F}^C$ must be perpendicular to any possible trajectory, we can write

$$\mathbf{F}^C \cdot \mathbf{v}^* = 0$$

where $\mathbf{v}^*$ is any velocity that is kinematically possible. We call $\mathbf{v}^*$ virtual motion and thus the equation above expresses that constraint forces do no virtual work. This means all the (virtual) work must be done by the specified forces.
Summing over all the particles in the system, we get

D'Alembert's principle:

$$\sum_{i=1}^{N} \mathbf{F}_i \cdot {\mathbf{v}^*}_i = \sum_{i=1}^{N} m_i \dot{\mathbf{v}}_i \cdot {\mathbf{v}^*}_i = \sum_{i=1}^{N} \mathbf{F}_i^S \cdot {\mathbf{v}^*}_i$$

where $\{{\mathbf{v}^*}_i\}$ is any virtual motion of the system at a time t.

2.3 Lagrange's equations

Lagrange's equations are obtained from D'Alembert's principle by choosing interesting values for ${\mathbf{v}^*}_i$.
As a first example, say we vary only the first generalised coordinate q_1 and keep all the others fixed (i.e. $\dot{q}_2 = \ldots = \dot{q}_n = 0$). If we further take $\dot{q}_1 = 1$, we have according the equation 2.1

$${\mathbf{v}^*}_i = \frac{\partial \mathbf{r}_i}{\partial q_1}.$$

According to D'Alembert's principle, it follows that

$$\sum_{i=1}^{N} m_i \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_1} = \sum_{i=1}^{N} \mathbf{F}_i^S \cdot \frac{\partial \mathbf{r}_i}{\partial q_1}$$

Varying the other generalised coorinates in the same way, we obtain the system of equations

$$\sum_{i=1}^{N} m_i \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_{i=1}^{N} \mathbf{F}_i^S \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \qquad j = 1, \ldots, n \tag{2.2}$$

The right hand side of this equation is called the generalised force Q_j corresponding to the coordinate q_j:

$$Q_j \equiv \sum_i \mathbf{F}_i^S \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}.$$

The left hand side can be rewritten in terms of the kinetic energy $T = \sum_i \frac{m_i}{2} \mathbf{v}_i \cdot \mathbf{v}_i$ in the following way:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = \sum_{i=1}^{N} m_i \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \tag{2.3}$$

The interpretation of this formula is not entirely straightforward. In particular the partial derivative with respect to the generalised velocity $\dot{q}_j$ combined with the total derivative with respect to time is quite surprising.
On reviewing equation 2.1 (and recalling that $\mathbf{r}_i = \mathbf{r}_i(\mathbf{q})$ is a function of the configuration $\mathbf{q}$) , we see that in general the kinetic energy can be seen as a function of $\mathbf{q}$ and $\dot{\mathbf{q}}$. We can treat $\mathbf{q}$ and $\dot{\mathbf{q}}$ as independent variables, due to the fact that at any particular time t, the system can be found in any configuration $\mathbf{q}$ with any set of velocities $\dot{\mathbf{q}}$. In other words, $\mathbf{q}$ and $\dot{\mathbf{q}}$ are only dependent if we are considering the time evolution of the system.
When computing the partial derivatives of the kinetic energy, we disregard the time evolution of the system and assume $\mathbf{q}$ and $\dot{\mathbf{q}}$ are independent variables.
At this point it may not be obvious why we would not consider the time evolution of the system. The equations we have derived so far are not guaranteed to be true if we just go plugging in random values of $\mathbf{q}$ and $\dot{\mathbf{q}}$. It is however always *possible* to compute $\frac{\partial T}{\partial \dot{q}_j}$ while forgetting about any time evolution, regardless of whether it is useful or not.
It should also be noted that there is no ambiguity in notation; if we state that $T = T(\mathbf{q}, \dot{\mathbf{q}})$ and we take a partial derivative with respect to $\dot{\mathbf{q}}$, the definition of partial derivative forces us to keep $\mathbf{q}$ constant (and vice versa). Thus the variables must be independent.
Then there is the total derivative with respect to time. This only makes sense if we consider the time evolution of the system; in which case $T(\mathbf{q}, \dot{\mathbf{q}}) = T(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = T(t)$. Thus T is a function only of time and the notation $\frac{\mathrm{d}}{\mathrm{d}t}$ makes sense. In a strict mathematical sense $T(t)$ and $T(\mathbf{q}, \dot{\mathbf{q}})$ are different functions. The context of the derivative makes it clear which one we need to use.
In conclusion then, the expression

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j}$$

is quite strange. It is not at all obvious, to me at least, that computing it results in anything at all but useful, but we *can* compute it, given an expression for T in function of $\mathbf{q}$ and $\dot{\mathbf{q}}$ and it does result in something useful.
Now we are ready to tackle the proof that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = \sum_{i=1}^{N} m_i \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}.$$

We start with

$$\begin{aligned}\frac{\partial}{\partial \dot{q}_j}\left(\frac{1}{2}\mathbf{v}_i \cdot \mathbf{v}_i\right) &= \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} \\ &= \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}\end{aligned}$$

this follows from first the chain rule, and then $\mathbf{v}_i = \frac{\partial \mathbf{r}_i}{\partial q_1}\dot{q}_1 + \ldots + \frac{\partial \mathbf{r}_i}{\partial q_n}\dot{q}_n$ using the independence of q_j and $\dot{q}_j$.
Then a further application of the chain rule gives

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial \dot{q}_j}\left(\frac{1}{2}\mathbf{v}_i \cdot \mathbf{v}_i\right)\right] &= \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} + \mathbf{v}_i \cdot \tfrac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbf{r}_i}{\partial q_j}\right) \\ &= \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} + \mathbf{v}_i \cdot \sum_{k=1}^{n} \frac{\partial^2 \mathbf{r}_i}{\partial {q_k}^2} q_j \dot{q}_k.\end{aligned}$$

Similarly (and again assuming independence of $\mathbf{q}$ and $\dot{\mathbf{q}}$),

$$\frac{\partial}{\partial q_j}\left(\frac{1}{2}\mathbf{v}_i \cdot \mathbf{v}_i\right) = \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial q_j} = \mathbf{v}_i \cdot \frac{\partial}{\partial q_j}\left(\sum_{k=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_k}\dot{q}_k\right) = \mathbf{v}_i \cdot \sum_{k=1}^{n} \frac{\partial^2 \mathbf{r}_i}{\partial {q_k}^2} q_j \dot{q}_k.$$

Combining these two results, gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial \dot{q}_j}\left(\frac{1}{2}\mathbf{v}_i \cdot \mathbf{v}_i\right)\right] - \frac{\partial}{\partial q_j}\left(\frac{1}{2}\mathbf{v}_i \cdot \mathbf{v}_i\right) = \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

Multiplying by m_i and summing over i, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = \sum_{i=1}^{N} m_i \dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}.$$

Using this result, we can write

Lagrange's equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j \qquad (1 \leq j \leq n).$$

This is just a rewrite of equation 2.2. It turns out that in general it is much easier to compute in this form.

2.4 Lagrange's equations for a conservative system

If the system is conservative, then the specified force has a potential associated with it. The generalised force Q_j can then be written in terms of this potential.

So we assume $\mathbf{F}^S = -\nabla V$. Then, if $\mathbf{q}^A$ and $\mathbf{q}^B$ are any two points in configuration space that can be joined by a straight line parallel to the q_j-axis, we can write

$$\begin{aligned}\int_{\mathbf{q}^A}^{\mathbf{q}^B} Q_j \,\mathrm{d}q_j &= \int_{\mathbf{q}^A}^{\mathbf{q}^B} \left(\sum_i \mathbf{F}_i^S \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}\right) \mathrm{d}q_j \\ &= \sum_i \int_{C_i} \mathbf{F}_i^S \cdot \mathrm{d}\mathbf{r} = V(\mathbf{q}^A) - V(\mathbf{q}^B) \\ &= -\int_{\mathbf{q}^A}^{\mathbf{q}^B} \frac{\partial V}{\partial q_j} \,\mathrm{d}q_j\end{aligned}$$

This equality holds for all $\mathbf{q}^A$, $\mathbf{q}^B$ chosen as described, so the integrands must be equal:

$$Q_j = -\frac{\partial V}{\partial q_j}$$

This yields the following form for Lagrange's equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = -\frac{\partial V}{\partial q_j} \qquad (1 \le j \le n).$$

For solving practical mechanics problems this is usually the most useful form.
There is however one last way we can rewrite Lagrange's equations, to put them in their most famous and elegant form. Since the potential is only a function of $\mathbf{q}$, $\frac{\partial V}{\partial \dot{q}_j}$ must be zero. So the equations can be written

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{q}_j}\right) - \frac{\partial V}{\partial q_j} \qquad (1 \le j \le n).$$

We now introduce $L(\mathbf{q}, \dot{\mathbf{q}}) \equiv T(\mathbf{q}, \dot{\mathbf{q}}) - V(\mathbf{q})$ which is called the Lagrangian of the system. Rewriting using the Lagrangian we get

Lagrange's equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad (1 \le j \le n).$$

This new way of writing the equations is mainly interesting from a theoretical point of view. It is this form of the equations that we will derive from the least action principle and it is this form of the equations that is most suitable to be adapted to quantum mechanics.

2.4.1 Velocity dependent potential

There are systems whose specified forces are not conservative, but their equations of motion can still be written in the Lagrangian form. This is possible if the generalised forces happen to be able to be written in the form

$$Q_j = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial U}{\partial \dot{q}_j}\right) - \frac{\partial U}{\partial q_j}$$

for some function $U(\mathbf{q}, \dot{\mathbf{q}})$ which is called the velocity dependent potential.
In practice there is really only one important case: that of a charged particle moving in electromagnetic fields. This case will be treated in the part on electromagnetism.

2.4.2 Sufficiency of the Lagrange equations

TODO

2.5 Lagrange's equations for systems with moving constraints

The theory expounded so far can (somewhat surprisingly) be extended to the class of problems in which the constraints are time dependent. Example include driven oscillators that are driven at a certain *prescribed* frequency and contraints that rotate in a *prescribed* manner.
TODO refer to introduction of generalised coordinates
The result is that now $\mathbf{r}_i$ depends on both $\mathbf{q}$ and time t.

$$\mathbf{r}_i = \mathbf{r}_i(\mathbf{q}, t) \qquad (1 \leq i \leq N)$$

The particle velocities are now given by

$$\mathbf{v}_i = \frac{\partial \mathbf{r}_i}{\partial q_1}\dot{q}_1 + \ldots + \frac{\partial \mathbf{r}_i}{\partial q_n}\dot{q}_n + \frac{\partial \mathbf{r}_i}{\partial t}.$$

Notice that we are treating t as an independent variable.
Obviously the constraint forces of moving constraints do work (just look at the driven oscillator), so initially there is no reason to think that systems with moving constraints satisfy Lagrange's equations, but it turns out they do. We motivate this claim in three steps:

1. Obviously D'Alembert's principle can no longer be true, but the conclusion (equation 2.2) we reached based on the principle *is*. In order for this to be true, we just need to verify that
$$\sum_{i=1}^{N} \mathbf{F}_i^C \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = 0.$$
In essence what we are trying to prove is that, while the constraint forces in general can do both real and virtual work, they do *not* do any virtual work if only the coordinate q_j is allowed to vary.
This is because, by the definition of the partial derivative, $\frac{\partial \mathbf{r}_i}{\partial q_j}$ is calculated keeping all the other coordinates *and time* constant. Now of course $\mathbf{F}_i^C$ does depend on time, but at any particular time t, the expression $\sum_{i=1}^{N} \mathbf{F}_i^C \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$ is the same as if the constraints were fixed in time, meaning it is still zero.

2. Next we need to verify equation 2.3. One may think that the extra term $\frac{\partial \mathbf{r}_i}{\partial t}$ in the expression for the particle velocity would throw sand in the works. It does not, mainly because the partial derivatives get rid of it. It is straightforward to redo the proof with this extra term.

3. Lastly, for a conservative system, is

$$\sum_{i=1}^{N} \mathbf{F}_i^S \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\frac{\partial V}{\partial q_j}$$

still true? The answer is yes. The proof is exactly the same because the partial derivatives are calculated at constant time.

So to recap: Lagranges's equations still hold when moving constraints are present, provided that t is regarded as an independent variable.

Chapter 3

Action and Hamilton's principle

3.1 A brief overview of the principle of least action

- stationary Hero of Alexandria Maupertuis and Hamilton

3.2 Variation of functionals

3.3 The Euler-Lagrange equation

Chapter 4

Hamilton's formulation

Chapter 5

Symmetry and conservation principles

Chapter 6

Non-holonomic systems

Chapter 7

The second variation

Part IV

Thermodynamics

not wrong (einstein). Arguments based on general thermodynamic arguments are very strong.
Maximises disorder.
Chemical kinetmatics

$$J = K \exp(-E/k_B T)$$

ideal gas law
Gibbs phase rule
`https://en.wikipedia.org/wiki/Category:Thermodynamic_equations`
(TODO quantum later)

Chapter 1

Purview and postulates

Few relevant degrees of freedom (of the 10^{23}):

- Slow on atomistic time scale 10^{-15}s
- Coarse on atomistic length scale: only to within 10^9 atoms

Only those few that are essentially time independent are macroscopically observable. (Although up to 10^{-7}s). Thermodynamics limiting case.

Chapter 2

Zeroth law, temperature and equilibrium

Chapter 3

First law and energy

Chapter 4

Second law, entropy and engines

4.1 Third law and absolute zero

4.2 Nernst's postulate

Chapter 5

Thermodynamic potentials

Chapter 6

Maxwell relations

Chapter 7

Phase transitions

7.1 First-order phase transitions

Chapter 8

Stability of thermodynamic systems

Chapter 9

Critical phenomena

Chapter 10

Irreversible thermodynamics

Part V

Statistical mechanics

Chapter 1

Random walks

1.1 Diffusion

1.2 Bownian motion

1.3 Polymers

Chapter 2

Ensemble theory

Chapter 3

Entropy

Chapter 4

Interacting systems

4.1 Virial theorem

4.2 Virial expansion

Chapter 5

Some models

5.1 Van de Waals model

5.2 Ising model

Chapter 6

Mean field theories

Chapter 7

Scaling theory

Chapter 8

The renormalisation group

Chapter 9

Non-equilibrium statistical mechanics

TODO Fick's law

Part VI

Fluid mechanics

Chapter 1

The continuum model

1.1 Hypotheses and definition

TODO + link mathematics - physics
Mesoscale + image
Description with vector field

$$\mathbf{v}(\mathbf{x}, t)$$

totally general
State variables:

$$\mathbf{v}, \rho, T, p$$

1.2 Observables

Based on the flow field $\mathbf{v}(\mathbf{x}, t)$ a number of quantities can be calculated that can be observed directly in an experimental setup.

1.2.1 Particle paths

For a given flow field $\mathbf{v}(\mathbf{x}, t)$, the paths of fluid elements can be calculated by solving

$$\boxed{\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{v}(\mathbf{x}, t)}.$$

Particle paths can be observed experimentally by tracking an object floating in the fluid or an ink dot. The tracking may be done by recording the placement at different times or taking a picture with a long exposure time.

1.2.2 Streamlines

A streamline is the path a fluid element would take if the flow field was static, meaning it did not change over time. Flows described by a static flow field are called steady flows.
Streamlines can be observed experimentally by taking very short snapshots, for example by injecting several ink markers, taking a photo with an exposure time just long enough to get short streaks.

A path $\gamma : \mathbb{R} \to \mathrm{E}^3 : u \mapsto \gamma(u)$ is a streamline, if its tangent vectors are proportional to vectors of $\mathbf{v}(\mathbf{x}, t_0)$:

$$\frac{\mathrm{d}\gamma(u)}{\mathrm{d}u} = \lambda \cdot \mathbf{v}(\mathbf{x}, t_0).$$

Fixing a basis, splitting γ into its components $\gamma(u) = (x(u), y(u), z(u))$ and writing $\mathbf{v} = (v_x, v_y, v_z)$, the differential equation can be written as

$$\frac{\frac{\mathrm{d}x}{\mathrm{d}u}}{v_x} = \frac{\frac{\mathrm{d}y}{\mathrm{d}u}}{v_y} = \frac{\frac{\mathrm{d}z}{\mathrm{d}u}}{v_z}$$

1.2.3 Streaklines

A streakline is made up of the current location of all fluid elements that passed through a fixed position $\mathbf{x}_0$ during the time interval $[t_0, t]$ from a time t_0 until now.
Experimentally this can be realised by continuously releasing ink or smoke (depending of the type of fluid we are dealing with) at a fixed location.
Because streaklines are composed of locations of fluid elements, the tangent vectors of a streakline $\gamma(t)$ must match the vectors of the flow field, $\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \mathbf{v}(\gamma(t), t)$. Additionally the streakline must originate at a location $\mathbf{x}_0$ at a time t_0. The problem then becomes to find a curve

$$\gamma : [t_0, t] \to \mathrm{E}^3 : t \mapsto \gamma(t)$$

such that

$$\begin{cases} \frac{\mathrm{d}\gamma}{\mathrm{d}t} = \mathbf{v}(\gamma(t), t) \\ \gamma(t_0) = \mathbf{x}_0. \end{cases}$$

1.3 Mathematical description

1.3.1 Eulerian and Lagrangian descriptions

The flow of a fluid can be studied by considering the velocity, density, temperature, pressure at different points in space. In other words, for each point $\mathbf{x}$ and time t, each of the above variables assumes a certain value. This is known as the Eulerian description of flow.
The Lagrangian description assigns values of these variables to (infinitesimal) fluid elements, not to points in space. A particular fluid element my be referenced by supplying its location $\mathbf{a}$ at $t = 0$. In the Lagrangian description the independent variables are $\mathbf{a}$ and t, not $\mathbf{x}$ and t.

$$\begin{cases} \mathbf{v}(\mathbf{x}, t) & \text{Eulerian description} \\ \mathbf{v}(\mathbf{a}, t) & \text{Lagrangian description} \end{cases}$$

One can move from the Eulerian description to the Lagrangian description by filling in a particular path $\mathbf{x}(\mathbf{a}, t)$. Such a particle path can be obtained by solving the equation cited above ($\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{v}$) with the boundary condition

$$\mathbf{x} = \mathbf{a} \qquad \text{at} \qquad t = 0$$

1.3.2 Material derivative

The two descriptions are completely equivalent, but often the Eulerian description will lead to simpler expressions. For this reason we will mostly use this description. It will however often prove useful to consider the partial derivative with respect to time in the Lagrangian description, which means keeping the initial position constant, i.e. differentiating following a fluid particle. This is as opposed to differentiating with respect to time keeping the Eulerian spatial coordinates constant.
When working in the Eulerian description, the latter is just the usual partial derivative, $\frac{\partial}{\partial t}$. The former derivative, following the fluid, is sometimes called the material derivative, the convective derivative or the Lagrangian derivative. It will be denoted $\frac{\mathrm{d}}{\mathrm{d}t}$.
Say $f(\mathbf{x},t) = f(x,y,z;t)$ is a scalar quantity of interest in the fluid (in the Eulerian description). The notation of the material derivative comes from the equality

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{x}(t),t) = \frac{\mathrm{d}}{\mathrm{d}t} f(x(t),y(t),z(t);t)$$

where $\mathbf{x}(t) = (x(t),y(t),z(t))$ is the path of a fluid element. The left hand derivative is material while the right hand one is an ordinary derivative seeing as $f(\mathbf{x}(t),t)$ only depend on t.
The material derivative can be calculated using the chain rule

$$\begin{aligned}\tfrac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\tfrac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\tfrac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial f}{\partial z}\tfrac{\mathrm{d}z}{\mathrm{d}t} \\ &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}v_x + \frac{\partial f}{\partial y}v_y + \frac{\partial f}{\partial z}v_z \\ &= \frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f\end{aligned}$$

where $\mathbf{v} = (v_x, v_y, v_z)$ is the value of the flow field at the point $\mathbf{x}(t)$.
This expression for the material derivative is sometimes taken as the definition.
The material derivative is readily generalised to vectors $\mathbf{u}$, by viewing the derivative $\mathbf{v}\cdot\nabla$ as the streamline directional derivative. So we may write

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{u}.$$

This works for arbitrary tensors as well.
Another possible generalisation is to the tensor derivative $\mathbf{v}\cdot(\nabla T)$. With T a tensor. This gives the same result.

1.3.2.1 Confusions in terminology

Many different names are sometimes used for the material derivative, such as

- advective derivative
- convective derivative
- derivative following the motion
- hydrodynamic derivative
- Lagrangian derivative

- particle derivative
- substantial derivative
- substantive derivative
- Stokes derivative
- total derivative

Some of these, such as the total derivative, are commonly used to refer to something else entirely. Often the term convective derivative is used to refer to only the spatial term $\mathbf{v} \cdot \nabla f$. Some authors then make the further distinction between advection for scalars and convection for tensors.

1.3.2.2 Link with streamlines

For streamlines we assume we are in a steady flow, i.e. $\frac{\partial f}{\partial t} = 0$. This means that

$$\begin{aligned}
\tfrac{\mathrm{d}f}{\mathrm{d}t} &= \mathbf{v} \cdot \nabla f \\
&= v\, \mathbf{t} \cdot \nabla f && \text{where } \mathbf{t} \text{ is the unit tangent vector along the streamline} \\
&= v \frac{\partial f}{\partial s} && \text{where } s \text{ is the distance along the streamline.}
\end{aligned}$$

Thus $\mathbf{v} \cdot \nabla f = 0$ means f is constant along streamlines.

1.3.2.3 From Lagrangian to Eulerian descriptions at a certain time

The transformation $\mathbf{a} \to \mathbf{x}$ is a transformation of Euclidean space with time t as a parameter. Assuming the fluid is a continuum, this transformation is continuous. On physical grounds, it must be one to one and have an inverse.
The transformation has a Jacobian

$$J(\mathbf{x}, t) = \frac{\partial(x_1, x_2, x_3)}{\partial(a_1, a_2, a_3)}.$$

Since the transformation is invertible and continuous, J is bounded and non-zero and we can find the convective derivative of J:

$$\begin{aligned}
\tfrac{\mathrm{d}J}{\mathrm{d}t} &= \frac{\partial(\frac{\mathrm{d}x_1}{\mathrm{d}t}, x_2, x_3)}{\partial(a_1, a_2, a_3)} + \frac{\partial(x_1, \frac{\mathrm{d}x_2}{\mathrm{d}t}, x_3)}{\partial(a_1, a_2, a_3)} + \frac{\partial(x_1, x_2, \frac{\mathrm{d}x_3}{\mathrm{d}t})}{\partial(a_1, a_2, a_3)} \\
&= \frac{\partial(v_1, x_2, x_3)}{\partial(a_1, a_2, a_3)} + \frac{\partial(x_1, v_2, x_3)}{\partial(a_1, a_2, a_3)} + \frac{\partial(x_1, x_2, v_3)}{\partial(a_1, a_2, a_3)}.
\end{aligned}$$

From

$$\begin{aligned}
\frac{\partial(v_1, x_2, x_3)}{\partial(a_1, a_2, a_3)} &= \det \begin{pmatrix} \frac{\partial v_1}{\partial x_1}\frac{\partial x_1}{\partial a_1} & \frac{\partial v_1}{\partial x_1}\frac{\partial x_1}{\partial a_2} & \frac{\partial v_1}{\partial x_1}\frac{\partial x_1}{\partial a_3} \\ \frac{\partial x_2}{\partial a_1} & \frac{\partial x_2}{\partial a_2} & \frac{\partial x_2}{\partial a_3} \\ \frac{\partial x_3}{\partial a_1} & \frac{\partial x_3}{\partial a_2} & \frac{\partial x_3}{\partial a_3} \end{pmatrix} \\
&= \frac{\partial v_1}{\partial x_1} J
\end{aligned}$$

we get that

$$\frac{\mathrm{d}J}{\mathrm{d}t} = J \nabla \cdot \mathbf{v}. \tag{1.1}$$

1.4 Reynolds' transport theorem

Sometimes we want to follow a finite (i.e. not infinitesimal) fluid volume, called a material fluid volume. This volume will expand, compress and deform as it moves, but always contains the same particles.
Let $L(\mathbf{x}, t)$ be any property of the fluid. We can then associate the property

$$\int_{V(t)} L(\mathbf{x}, t)\, \mathrm{d}V(\mathbf{x})$$

with the material volume $V(\mathbf{x})$. We are now interested in how this changes in time, i.e. the time derivative. We calculate the time derivative by transforming to Lagrangian coordinates, which makes the infinitesimal volume element $\mathrm{d}V(\mathbf{a})$ no longer depend on time:

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}t}\left[\int_{V(t)} L(\mathbf{x}, t)\, \mathrm{d}V(\mathbf{x})\right] &= \tfrac{\mathrm{d}}{\mathrm{d}t}\left[\int_{V(0)} L(\mathbf{x}(\mathbf{a}, t), t) J\, \mathrm{d}V(\mathbf{a})\right] \\
&= \int_{V(0)} \tfrac{\mathrm{d}}{\mathrm{d}t}\left[L(\mathbf{x}(\mathbf{a}, t), t) J\right] \mathrm{d}V(\mathbf{a}) \\
&= \int_{V(0)} \left[J \tfrac{\mathrm{d}L}{\mathrm{d}t} + L \tfrac{\mathrm{d}J}{\mathrm{d}t}\right] \mathrm{d}V(\mathbf{a}) \\
&= \int_{V(0)} \left[\tfrac{\mathrm{d}L}{\mathrm{d}t} + L\, \nabla \cdot \mathbf{v}\right] J\, \mathrm{d}V(\mathbf{a}) \qquad \text{using equation (1.1)} \\
&= \int_{V(t)} \left[\tfrac{\mathrm{d}L}{\mathrm{d}t} + L\, \nabla \cdot \mathbf{v}\right] \mathrm{d}V(\mathbf{x})
\end{aligned}$$

The last equality gives the transport theorem

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{V(t)} L(\mathbf{x}, t)\, \mathrm{d}V(\mathbf{x})\right] = \int_{V(t)} \left[\tfrac{\mathrm{d}L}{\mathrm{d}t} + L\, \nabla \cdot \mathbf{v}\right] \mathrm{d}V(\mathbf{x}) \tag{1.2}$$

1.5 Cauchy-Stokes decomposition theorem

The Cauchy-Stokes decomposition shows that any flow can be fully characterised by its divergence and vorticity. This is to be expected as

$$\begin{cases} \nabla \cdot \mathbf{v} = f(\mathbf{x}, t) \\ \nabla \times \mathbf{v} = \mathbf{g}(\mathbf{x}, t) \end{cases}$$

gives 3 independent partial differential equations for three components v_x, v_y, v_z. (Only 2 of the 3 components of $\mathbf{g}$ are independent as they have to satisfy $\nabla \cdot \mathbf{g} = 0$.) These should determine $\mathbf{v}$, if boundary conditions are given.

1.5.1 Divergence

We would like to give a physical interpretation to the divergence $\nabla \cdot \mathbf{v}$. It turns out that the divergence is a measure for the rate of variation of the volume of a material fluid element. To see how, we make use of the transport theorem (1.2). If we set $L(\mathbf{x}, t) = 1$, then

$$\int_{V(t)} \mathrm{d}V(t) = \mathcal{V}(t) = \text{volume of the material fluid element.}$$

The transport theorem then gives

$$\frac{\mathrm{d}\mathcal{V}(t)}{\mathrm{d}t} = \int_{V(t)} (\nabla \cdot \mathbf{v})\, \mathrm{d}V(\mathbf{x}).$$

Taking the limit of $\mathcal{V} \to 0$ and using the mean value theorem for real functions of several variables we get

$$\lim_{\mathcal{V}\to 0} \frac{1}{\mathcal{V}(t)} \frac{\mathrm{d}\mathcal{V}(t)}{\mathrm{d}t} = \nabla \cdot \mathbf{v}$$

1.5.1.1 Incompressible flow

A flow is incompressible if no material fluid volume can change its volume as it moves. Since

$$\int \nabla \cdot \mathbf{v}\, \mathrm{d}V = 0$$

for all possible material volumes, it follows that

$$\nabla \cdot \mathbf{v} = 0$$

everywhere in an incompressible flow.

1.5.2 Vorticity

We define the vorticity $\boldsymbol{\omega}$ as the curl of the velocity field.

$$\boldsymbol{\omega} \equiv \nabla \times \mathbf{v}$$

The vorticity is a measure for the local rotation of the fluid.
TODO Consider the material derivative of the velocity field.

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}.$$

Using the vector identity

$$\mathbf{a} \times (\nabla \times \mathbf{a}) = \nabla \left(\frac{a^2}{2} \right) - (\mathbf{a} \cdot \nabla)\mathbf{a}$$

we write

$$\begin{aligned}(\mathbf{v} \cdot \nabla)\mathbf{v} &= \nabla \left(\frac{v^2}{2} \right) - \mathbf{v} \times (\nabla \times \mathbf{v}) \\ &= \nabla \left(\frac{v^2}{2} \right) - \mathbf{v} \times (\omega).\end{aligned}$$

This contains the kinetic energy $v^2/2$ and the vorticity explicitly.

1.5.3 Decomposition

We shall now show that in the neighbourhood of each point of the fluid $\mathbf{x}_0$, the velocity is the sum of a translation, a rigid rotation and a deformation.

We start by doing a Taylor expansion around a point $\mathbf{x}_0$:

$$\begin{aligned}\mathbf{v}(\mathbf{x}) = \mathbf{v}(\mathbf{x}_0) &= \begin{pmatrix} \frac{\partial v_x}{\partial x} & \frac{\partial v_x}{\partial y} & \frac{\partial v_x}{\partial z} \\ \frac{\partial v_y}{\partial x} & \frac{\partial v_y}{\partial y} & \frac{\partial v_y}{\partial z} \\ \frac{\partial v_z}{\partial x} & \frac{\partial v_z}{\partial y} & \frac{\partial v_z}{\partial z} \end{pmatrix} \begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} + O(h^2) \\ &= \mathbf{v}(\mathbf{x}_0) + M\mathbf{h} + O(h^2)\end{aligned}$$

The we write M as the sum of a symmetric and an antisymmetric matrix:

$$\begin{aligned}M &= \frac{1}{2}(M + M^\intercal) + \frac{1}{2}(M - M^\intercal) \\ &= D + R\end{aligned}$$

The matrix D will turn out to be associated with the deformation and R with the rotation.

1.5.3.1 2D flows

In two dimensions the elements of the decomposition are simply

- $\boldsymbol{\omega} = \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right)\hat{z}$
- $D = \begin{pmatrix} \frac{\partial v_x}{\partial x} & \frac{1}{2}\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right) \\ \frac{1}{2}\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right) & \frac{\partial v_y}{\partial y} \end{pmatrix}$
- $\frac{(\boldsymbol{\omega}_0 \times \mathbf{x})}{2} = \frac{1}{2}(-\omega_0 y, \omega_0 x)$

1.6 Examples of flow models

1.6.1 In two dimensions

1.6.1.1 Point vortex

1.6.1.2 Rankine vortex

1.6.2 In three dimensions

Chapter 2

Some naive formulae

Hydrostatic pressure

Archimedes' principle

Bernoulli's equation

Chapter 3

Navier-Stokes equations

3.1 Conservation of mass

3.2 Conservation of momentum

3.3 Conservation of energy

surface energy and tension
Young–Laplace equation

3.4 Navier-Stokes equations and Euler equations

Chapter 4

Properties of fluid flows

4.1 Compressibility

4.2 Viscosity

4.2.1 Reynolds number

4.3 Vorticity

4.3.1 Crocco's theorem

4.3.2 Vortex tubes

4.3.3 Kelvins's theorem for isentropic flow

Chapter 5

Calculations under certain assumptions

5.1 Fluid at rest: hydrostatic pressure

5.2 Irrotational: potential flows

5.3 Bernoulli's equations

5.4 Examples of low Reynolds' number flows

5.4.1 Plane Couette flow

5.4.2 Plane Poiseuille flow

5.4.3 Poiseuille flow (blood flow)

Chapter 6

Airflow around a wing

6.1 Potential flow

6.2 Flow around a cylinder

6.3 Flow around a wing

Chapter 7

Waves

TODO phase / group velocity

7.1 Surface gravity waves

7.2 Internal gravity waves

7.3 Sound waves

7.3.1 Intensity of sound: Decibels

7.3.2 The Doppler effect

7.3.3 Shock Waves and the sonic boom

Part VII

Classical electromagnetism

Chapter 1

Introduction

TODO light and electromagnetic radiation
energy depends only on intensity?? notation: ν, λ
electronics: LRC
Classical electrodynamics studies the effects of charges and currents within the context of classical Newtonian mechanics. Now if we know what force charged particles exert on each other, we can in principle solve the second order differential equation that is Newton's second law and we can find out all we want to know.
Say we have a charged particle we want to follow in particular (we call this a test charge). We want to know what effect other charged particles (called source charges) in the neighbourhood have on the test charge. Luckily we are aided by the principle of superposition baked into Newtonian mechanics: Say we have source charges $q_1, q_2, \ldots, q_n$ and we call the forces due to each source charge on its own $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ respectively (i.e. if we only had source charge q_1, the test charge would feel a force $\mathbf{F}_1$). Now the principle of superposition asserts that, for the total force $\mathbf{F}$ on the test charge, the following holds:

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \ldots \mathbf{F}_n$$

So we only need to find a formula for the force between two charged particles.
This is the good news. The bad news is that writing down such a formula in full generality is quite difficult. We will approach it in steps. This also allows us to explore some of the interesting features and useful mathematics of electromagnetic theory.

1.1 Notation

Before we start, we will define some notation used in this chapter.

- In general we denote the test charge Q and use lower case q to reference source charges, indexed as necessary.
- Because the distance between points is so important, we introduce the script $\mathscr{r}$ to denote it. In particular if are considering two points with at locations $\mathbf{r}$ and $\mathbf{r'}$, we define

  $$\boldsymbol{\mathscr{r}} \equiv \mathbf{r} - \mathbf{r'}$$

  We also introduce related notation for the unit vector in the direction of $\boldsymbol{\mathscr{r}}$:

  $$\hat{\boldsymbol{\mathscr{r}}} \equiv \frac{\mathbf{r} - \mathbf{r'}}{|\mathbf{r} - \mathbf{r'}|}$$

and magnitude r of $\boldsymbol{r}$.

It bears stressing that $\boldsymbol{r}$ is not the same as the displacement vector $\mathbf{r}$.

Chapter 2

Electrostatics

Our first step will be electrostatics, where the source charges are stationary (although the test charge may move).
TODO summary triangle

2.1 Coulomb's law

In this case the important formula is Coulomb's law:

$$\boxed{\mathbf{F} = \frac{1}{4\pi\epsilon_0}\frac{qQ}{r^2}\hat{\boldsymbol{r}}}$$

We call ϵ_0 the permitivity of free space.

$$\epsilon_0 = 8.85 \times 10^{12}\frac{\mathrm{C}^2}{\mathrm{N}\cdot\mathrm{m}^2}$$

This is an experimental law. The factor $\frac{1}{4\pi\epsilon_0}$ may at this stage be viewed as just that, an experimental factor. We will later see why it is written in this form.
Notice that the force is positive (which means repulsive) if both charges are either positive or negative; the force is negative (meaning attractive) if one charge is positive and one is negative. So we have been able to give an equation for the force in the electrostatic case. This contains all the physics. In the remainder of this section we will see some consequences of this law and elaborate some useful mathematical techniques.

2.2 The electric field E

2.2.1 Discrete source charges

Assume we have source charges $q_1, q_2, \ldots, q_n$ at distances $r_1, r_2, \ldots, r_n$ from the test charge Q. We can write

$$\begin{aligned}\mathbf{F} = \mathbf{F_1} + \mathbf{F}_2 + \ldots &= \frac{1}{4\pi\epsilon_0}\left(\frac{q_1 Q}{r_1^2}\hat{\boldsymbol{r}}_1 + \frac{q_2 Q}{r_2^2}\hat{\boldsymbol{r}}_2 + \ldots\right)\\ &= \frac{Q}{4\pi\epsilon_0}\left(\frac{q_1\hat{\boldsymbol{r}}_1}{r_1^2} + \frac{q_2\hat{\boldsymbol{r}}_2}{r_2^2} + \ldots\right)\end{aligned}$$

Thus we can the force as

$$\mathbf{F} = Q\mathbf{E}$$

Where we have introduced the electric field E. This means we can talk about electromagnetic effect without having to reference a particular test charge in a particular location, instead we define the electric field that has a (vectorial) value in every point in space. For discrete source charges, the electric field is

$$\mathbf{E}(\mathbf{r}) \equiv \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n} \frac{q_i}{\mathscr{r}_i^2} \hat{\boldsymbol{\mathscr{r}}}_i,$$

where the $\mathscr{r}_i$'s are the distances between $\mathbf{r}$ and the source charge q_i.

2.2.2 Continuous charge distributions

The above formula can be readily extended to a continuous charge distribution:

$$\mathbf{E}(\mathbf{r}) \equiv \frac{1}{4\pi\epsilon_0} \int \frac{1}{\mathscr{r}^2} \hat{\boldsymbol{\mathscr{r}}} \, \mathrm{d}q.$$

We can reformulate this as a slightly more tractable volume integral using the charge density ρ:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{\mathcal{V}} \frac{\rho(\mathbf{r}')}{\mathscr{r}^2} \hat{\boldsymbol{\mathscr{r}}} \, \mathrm{d}\tau',$$

where $\boxed{\mathrm{d}q = \rho \, \mathrm{d}\tau'}$. Similarly we can calculate the electric field of a charged surface using the charge per unit area σ (now $\mathrm{d}q = \sigma \, \mathrm{d}a'$):

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{\mathcal{S}} \frac{\sigma(\mathbf{r}')}{\mathscr{r}^2} \hat{\boldsymbol{\mathscr{r}}} \, \mathrm{d}a',$$

and of a linear charge distribution (with charge per unit length λ and $\mathrm{d}q = \lambda \, \mathrm{d}l'$):

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{\mathcal{P}} \frac{\lambda(\mathbf{r}')}{\mathscr{r}^2} \hat{\boldsymbol{\mathscr{r}}} \, \mathrm{d}l'.$$

These integrals can be very difficult to solve.

2.3 Field lines and flux

We can represent a vectorial field using field lines. For a given vector field (that is smooth enough) a field line is a (directed) line whose tangent vectors in every point are the vectors of the vector field. [TODO image]
This way field lines convey information about the direction of the vectors in the vector field. Coincidentally for the electric field we can do even better and express the (relative) strength of the field with field lines as well!
To see why, consider the field generated by a single source charge. We can now draw some field lines. At a certain distance r from the source charge, the density of lines is the total number of lines divided by the area of the sphere: $\frac{n}{4\pi r^2} \propto 1/r^2$.
[TODO image; all images using point sources here!]
The strength of the electric field

$$E = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2}$$

is also proportional to $1/r^2$. This means we can use the density of the field lines to represent the relative strength of the electric field (i.e. the strength of the electric field up to a certain constant that depends on how many field lines we can be bothered to draw, among other things).
We can also use field line diagrams to represent more complicated fields [TODO images of examples]. Due to the fact that the electric field and the field line density behave similarly under addition, we can keep using field line density to represent the relative strength of the electric field.
Some important remarks about field lines:

- Field lines start in positive charges and end in negative ones. The direction is always positive to negative.
- Field lines start and end either at charges of at infinity. They can never start or end in the middle of space.
- Field lines never cross. If they did cross, it would mean that they had different tangent vectors at the point of crossing which is definitionally impossible.

It is importance at this stage to give a more rigourous way to conceive of the density of field lines. To do that we define the flux of $\mathbf{E}$ through a surface $\mathcal{S}$ as

$$\Phi_E \equiv \int_{\mathcal{S}} \mathbf{E} \cdot \mathrm{d}\mathbf{a}$$

which gives a measure of the "number of field lines" going through the surface $\mathcal{S}$. TODO: why?

2.4 Gauss's law

Say we now take the flux through a closed surface. This has the following properties:

1. Any charges outside the closed surface do not contribute to the flux, because any field lines originating from them must both enter and exit the surface as field lines cannot stop in the middle of space. This means these field lines are both added to and subtracted from the flux.
2. If there is a charge enclosed in the surface, its location within the surface does not impact the total flux. For any field line that starts at the charge, there are two options:
   (a) It ends at another charge within the surface, in which case it either does not exit the surface or exits an reenters. Either way it does not contribute to the total flux.
   (b) It ends at a charge outside the surface or at infinity, in which case it contributes to the total flux by exiting, but it can only exit once (without reentering) and is thus counted once, regardless of its location relative to the surface.

Thus, using the superposition principle for the electric field, the flux depends only on the total enclosed charge (times a constant)! This is known as Gauss's law.

$$\oint_{\mathcal{S}} \mathbf{E} \cdot \mathrm{d}\mathbf{a} = \frac{Q_{\text{enc}}}{\epsilon_0}$$

We can easily see that the constant must be ϵ_0 by explicitly working out the flux for a simple setup, like the flux from a single charge through a concentric sphere. In fact this is the reason the constant in Coulomb's law was chosen the way it was.

Using the divergence theorem from vector calculus, we can obtain a differential form of Gauss's law:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

2.4.1 Applying Gauss's law

Gauss's law is incredibly powerful and can make calculating the electric field very easy, if the system is symmetric enough. Otherwise it is not useful. There are three kinds of symmetry that make Gauss's law useful:

1. Spherical symmetry: make the Gaussian surface a concentric sphere.
2. Cylindrical symmetry: make the Gaussian surface a coaxial cylinder.
3. Planar symmetry: make the Gaussian surface a "pillbox" that intersects the surface.

We can also calculate the electric field of more complex arrangements by decomposing them into elements with such symmetry and using the superposition principle.

2.4.2 The curl of the electric field

We can view the electric field as being composed of the fields of each charge in space. Each of those fields radiates straight out and is thus irrotational. This means that the total electric field is also irrotational:

$$\nabla \times \mathbf{E} = 0 \tag{2.1}$$

so the path integral around any closed path is zero

$$\oint \mathbf{E} \cdot \mathrm{d}\mathbf{l} = 0. \tag{2.2}$$

2.5 Electric potential

We can define the scalar field called the electric potential V:

$$V(\mathbf{r}) \equiv -\int_{\mathcal{O}}^{\mathbf{r}} \mathbf{E} \cdot \mathrm{d}\mathbf{l}.$$

Using the fact that the curl of $\mathbf{E}$ is zero, we can see that

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

Some remarks:

1. It has been claimed the the name potential is a misnomer, because it's too much like potential energy. This is possible. It is important to remember that the potential is not potential energy (it will turn out it is more akin to potential energy per unit charge).
2. A surface in space over which the potential is constant is called and equipotential (surface).
3. The advantage of the potential is that it has some of the characteristics of the electric field baked in. The electric field is a vector field which has the property of being irrotational. The potential is a scalar field which has the irrotationality of the electric field built in.

4. The potential is defined with respect to a reference point $\mathcal{O}$. This point is arbitrary, if we chose a different one it would just add a constant to the potential:

$$V'(\mathbf{r}) = -\int_{\mathcal{O}'}^{\mathbf{r}} \mathbf{E} \cdot \mathrm{d}\mathbf{l} = -\int_{\mathcal{O}'}^{\mathcal{O}} \mathbf{E} \cdot \mathrm{d}\mathbf{l} - \int_{\mathcal{O}}^{\mathbf{r}} \mathbf{E} \cdot \mathrm{d}\mathbf{l} = K + V(\mathbf{r}),$$

where $V'(\mathbf{r})$ is the potential with respect to the reference point $\mathcal{O}'$. Because adding a constant does not change the gradient ($\nabla V' = \nabla V$), both are equivalent. In other words only potential differences are physically relevant.

It is often natural to put the reference point at infinity (i.e. far away from the charges).

5. Positive test charges want to move to areas of low potential (meaning the source charges generate a force pulling them towards areas of low potential). Negative test charges want to move in the opposite direction.

6. The potential obeys the superposition principle.

$$V = V_1 + V_2 + \dots$$

7. **Units**: Force is measured in newtons and charge in coulombs, so the electric field is measured in newtons per coulomb. The potential is then measured in newton-meters per coulomb, or joules per coulomb, also known as volts.

2.5.1 Relationship with the charge density ρ

Combining $\mathbf{E} = -\nabla V$ and $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$, we get

$$\nabla^2 V = -\frac{\rho}{\epsilon_0}.$$

This is known as Poisson's equation. In regions where there is no charge, this is known a Laplace's equation

$$\nabla^2 V = 0.$$

In order to work out the potential from the charge distribution, we can use the following formula which is not too difficult to derive:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{\mathscr{r}} \, \mathrm{d}\tau'$$

This is quite a bit easier to work with than the equivalent formula for the electric field because it is scalar.

2.6 Charged surface

We quickly make some observations about passing through a charged surface with surface charge density σ.
TODO image
Using Gauss's law, we obtain

$$E^{\perp}_{\text{above}} - E^{\perp}_{\text{below}} = \frac{1}{\epsilon_0}\sigma.$$

Using $\oint \mathbf{E} \cdot \mathrm{d}\mathbf{l} = 0$, we also obtain

$$E^{\parallel}_{\text{above}} - E^{\parallel}_{\text{below}} = 0.$$

Combining them we get

$$\mathbf{E}_{\text{above}} - \mathbf{E}_{\text{below}} = \frac{\sigma}{\epsilon_0}\hat{n}$$

The potential, however, is continuous:

$$V_{\text{above}} = V_{\text{below}}$$

2.7 Work and energy

We can calculate the work necessary to move a charged particle from a point $\mathbf{a}$ to a point $\mathbf{b}$.

$$W = \int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F} \cdot \mathrm{d}\mathbf{l} = -Q \int_{\mathbf{a}}^{\mathbf{b}} \mathbf{E} \cdot \mathrm{d}\mathbf{l} = Q[V(\mathbf{b}) - V(\mathbf{a})].$$

In other words

$$V(\mathbf{b}) - V(\mathbf{a}) = \frac{W}{Q}.$$

So in this sense potential is just potential energy per unit charge.

2.7.1 Energy of a point charge distribution

TODO: superposition for potential energy: sum over all pairs of point charges.
In this section we are interested in knowing how much work it would take to assemble a collection of point charges. We can imagine bringing in the charges in one by one from infinity, which we choose as our reference point. This means that to bring a charge in from infinity we need to do the work

$$W = q_i V_{<i}(\mathbf{r}_i).$$

Where $\mathbf{r}_i$ is the place we are taking charge q_i to, and $V_{<i}$ is the potential generated by all the other charges that are already assembled (i.e. charges labeled up to i).
The first takes no work, because it is not fighting against any field. For the second charge we need to do the work

$$W = q_2 V_{<2}(\mathbf{r}_2) = q_2 \frac{1}{4\pi\epsilon_0}\left(\frac{q_1}{\imath_{12}}\right)$$

For the third charge, we have 2 such terms. Due to the superposition principle, we know that there is one for the first charge and one for the second.
Continuing that chain of logic, we can see that the total work is

$$W = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j<i}}^{n} \frac{q_i q_j}{\imath_{ij}}.$$

If we change the $j < i$ to $j \neq i$, two things happen:

1. Each pair is counted twice, so we need to add a factor $\frac{1}{2}$.

2. The second summation effectively becomes an expression for the potential when all the the charges are already assembled (dropping the $j < i$ means we count all of the charges is the final configuration). We can then just call this potential $V_i(\mathbf{r}_i)$, meaning the potential generated by all charges other than q_i.

So the expression for the work becomes

$$W = \frac{1}{2}\sum_{i=1}^{n} q_i V_i(\mathbf{r}_i) \tag{2.3}$$

2.7.2 Energy of a continuous charge distribution

By taking the individual changes to be infinitesimal, we can transition from a discrete sum, to an integral:

$$W = \frac{1}{2}\int \rho V \,\mathrm{d}\tau$$

for a volume charge density ρ. We can write similar expressions for line and surface charge densities. Notice we can write V and not V_i because the difference between the two is infinitesimal.
Now using $\rho = \epsilon_0 \nabla \cdot \mathbf{E}$ and integrating by parts, we get

$$W = \frac{\epsilon_0}{2}\left(\int_{\mathcal{V}} E^2 \,\mathrm{d}\tau + \oint_{\mathcal{S}} V\mathbf{E}\cdot \mathrm{d}\mathbf{a}\right).$$

Just like the previous one, this expression is valid if we integrate over any volume that contains all the charges. In particular we can integrate over all of space, which makes the surface integral go to zero.

$$W = \frac{\epsilon_0}{2}\int_{\text{all space}} E^2 \,\mathrm{d}\tau \tag{2.4}$$

Again there are some remarks to be made:

1. In deriving equation (2.4) for W, we explicitly made use of the fact that we are working with a charge distribution when we set $V_i = V$. The final equation does not reference the charge distribution however and so we may wonder whether this equation also works for point charges. Unfortunately applying this equation to a point charge gives us

$$W = \frac{\epsilon_0}{2(4\pi\epsilon_0)^2}\int \left(\frac{q^2}{r^4}\right)(r^2 \sin\theta \,\mathrm{d}r \,\mathrm{d}\theta \,\mathrm{d}\phi) = \frac{q^2}{8\pi\epsilon_0}\int_0^\infty \frac{1}{r^2}\,\mathrm{d}r = \infty.$$

   In fact this is because the equations (2.3) and (2.4) refer to slightly different energies: for the latter we move infinitesimal charges in from infinity, for the former we have ready-made point charges we can move in from infinity. What we see here is that bringing together enough infinitesimal charge close enough together to create a finite point charge, requires an infinite amount of work. This is a singularity that is also present in quantum electrodynamics and is still an open problem.

2. In the context of radiation theory it is useful and in the context of general relativity it is essential to view the equation (2.4) as the more fundamental one. In other words we view the energy as being stored in the field, not the charge. In electrostatics it does not make much difference.

3. The work depends quadratically on the electric field and thus the superposition principle does not apply.

2.8 Multipole expansion

Say we have a localised charge distribution that we are viewing from far away, i.e. the distance r between us and the charge distribution is large. Being far away we can approximate the potential as the one generated by a point with the same charge as the net charge of the distribution:

$$V_{\text{mon}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{Q}{r}.$$

The subscript mon refers to the fact that we are approximating the potential with that of a point charge, called a monopole in this context.
This means that the potential will go to zero at least as $1/r$, and maybe quicker if the net charge is zero. This motivates us to expand the potential in powers of $1/r$.
There is a nice correspondence between the powers of $1/r$ and point charge distributions. The $1/r$ term is called the monopole term. If we take two monopoles of opposite charge and put them at a distance d from each other, we have what is called a (physical) dipole. (TODO image with $d, \theta, q, \mathbf{r}$ and $\mathbf{p}$). The potential generated by this setup can be approximated at large distances by

$$V(\mathbf{r}) \approx \frac{1}{4\pi\epsilon_0}\frac{qd\cos\theta}{r^2}$$

This expression depends on the angle θ, but for a fixed θ the potential drops off as $1/r^2$. We call the $1/r^2$ term of our expansion the dipole term.
If we take a second identical dipole, flip it and put it next to the first one at a distance d, we have what is known as a quadrupole. Now the potential drops off as $1/r^3$ and consequently we call the $1/r^3$ term the quadrupole term. The $1/r^4$ term corresponds to an octopole (TODO image of multipoles).
Now let us explicitly do the expansion. We start with the formula to calculate the potential from an arbitrary charge distribution

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int \frac{1}{\mathscr{r}}\rho(\mathbf{r'})\,\mathrm{d}\tau'$$

Refering to figure TODO and using the law of cosines, we can write

$$\mathscr{r} = r\sqrt{1+\epsilon} \qquad \text{with} \qquad \epsilon \equiv \left(\frac{r'}{r}\right)\left(\frac{r'}{r} - 2\cos\theta'\right).$$

We can now use a binomial expansion:

$$\begin{aligned}
\frac{1}{\mathscr{r}} &= \frac{1}{r}(1+\epsilon)^{-1/2} = \frac{1}{r}\left(1 - \frac{1}{2}\epsilon + \frac{3}{8}\epsilon^2 - \frac{5}{16}\epsilon^3\right) \\
&= \frac{1}{r}\left[1 + \left(\frac{r'}{r}\right)\cos\theta' + \left(\frac{r'}{r}\right)^2(3\cos^2\theta' - 1)/2 + \left(\frac{r'}{r}\right)^3(5\cos^3\theta' - 3\cos\theta')/2 + \dots\right] \\
&= \frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{r'}{r}\right)^n P_n(\cos\theta')
\end{aligned}$$

where P_n are the Legendre polynomials.
Plugging this back in the expression for V, and noting that r stays constant during integration, we get

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\sum_{n=0}^{\infty}\frac{1}{r^{n+1}}\int (r')^n P_n(\cos\theta')\rho(\mathbf{r'})\,\mathrm{d}\tau'.$$

It must be remarked that the multipole expansion depends on the chosen coordinate system.

2.8.1 The monopole and dipole terms

Usually the multipole expansion is dominated by the first term

$$V_{\text{mon}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{1}{r}\int \rho(\mathbf{r'})\,\mathrm{d}\tau' = \frac{1}{4\pi\epsilon_0}\frac{Q}{r}$$

This is the potential of a point charge, which we expected given the discussion above. Thus for a point charge centered at the origin, the monopole term is exact and all higher multipole terms vanish. If the point charge is somewhere else, the monopole term stays the same, but we also have higher order corrections.
The dipole term is

$$V_{\text{dip}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{1}{r^2}\int r'\cos\theta'\rho(\mathbf{r'})\,\mathrm{d}\tau'.$$

Using the fact that $r'\cos\theta' = \hat{r}\cdot\mathbf{r'}$, we can write

$$\begin{aligned} V_{\text{dip}}(\mathbf{r}) &= \frac{1}{4\pi\epsilon_0}\frac{1}{r^2}\hat{r}\cdot\int \mathbf{r'}\rho(\mathbf{r'})\,\mathrm{d}\tau' \\ &= \frac{1}{4\pi\epsilon_0}\frac{\hat{r}\cdot\mathbf{p}}{r^2} \qquad \text{with} \qquad \mathbf{p} \equiv \int \mathbf{r'}\rho(\mathbf{r'})\,\mathrm{d}\tau' \end{aligned}$$

where we have introduced the dipole moment $\mathbf{p}$. We can remark on the following properties:

1. The dipole moment is a vector and can be manipulated accordingly. For example, if we add the dipole moments of the two dipoles that make up a quadrupole, we see that the dipole moment of the quadrupole vanishes.

2. For point charge distributions the integral becomes a sum in the usual way

   $$\mathbf{p} = \sum_{i=0}^{n} q_i \mathbf{r'}_i.$$

   If we apply this to the physical dipole, we get

   $$\mathbf{p} = q\mathbf{d}.$$

   Using this the dipole term calculated here is the same as the approximate field of the physical dipole quoted before.

3. For a physical dipole the dipole terms gives only an approximation of the potential. There are higher multipole terms. If we however shrink the distance d keeping the dipole moment $p = qd$ constant, we get what is called a pure dipole in the limit. For this distribution, all higher multipole terms vanish.

4. The dipole moment is independent of the chosen coordinate system, if and only if the monopole moment (i.e. the net charge) is zero.

2.8.1.1 A more geometric derivation of the dipole term

Consider figure (TODO image, add $\mathbf{d}$ between A and B) of a physical dipole. Let A be a point far away from the dipole with displacement $\mathbf{r}_+$ from the positive charge and $\mathbf{r}_-$ from the negative charge. Let B be the point such that $B = A + \mathbf{d}$. Let V_+ be the potential due the positive charge and V_- due to the negative charge.

Now we note

$$V(A) = V_+(A) + V_-(A) = \frac{q}{r_+} - \frac{q}{r_-} \qquad \text{and} \qquad V_-(A) = -\frac{q}{r_-} = -V_+(B)$$

From which we get

$$V(A) = V_+(A) - V_+(B) \approx -(\nabla V_+) \cdot \mathbf{d} = -\left(\nabla \frac{1}{r}\right) \cdot \mathbf{p} = \frac{\mathbf{p} \cdot \mathbf{r}}{r^3}$$

which is exactly the dipole term we got before.

2.8.2 The electric field of a dipole

We can simply calculate

$$\begin{aligned} \mathbf{E}_{\text{dip}}(r, \theta) &= -\nabla V_{\text{dip}} \\ &= \frac{p}{4\pi\epsilon_0 r^3}(2\cos\theta \hat{r} + \sin\theta\hat{\theta}). \end{aligned}$$

We know that to monopole electric field falls off as $1/r^2$. The dipole electric field we now see falls off as $1/r^3$. The quadrupole field will go like $1/r^4$, the octopole as $1/r^5$ etc.

2.9 Miscellaneous results in electrostatics

In this section a collection of useful and / or interesting results from electrostatics will be presented, typically pertaining to the electric field or potential for a certain configuration.

2.9.1 Average electric field inside a sphere.

We can split the average field into two components: the average field due to charges inside the sphere and due to charges outside the sphere.

- We start with the field **due to charges inside the sphere**. If there was only a single charge q at point $\mathbf{r}$ inside the sphere, the average field would be

$$\langle \mathbf{E} \rangle = \frac{1}{\mathcal{V}_{\text{sphere}}} \int \mathbf{E}(\mathbf{r'})\, \mathrm{d}\tau' = \frac{1}{\frac{4}{3}\pi R^3} \frac{1}{4\pi\epsilon_0} \int \frac{q}{\boldsymbol{\imath}^2} \hat{\boldsymbol{\imath}}\, \mathrm{d}\tau'$$

  where R is the radius of the sphere and $\mathcal{V}_{\text{sphere}}$ is the volume. The dipole moment of this point charge is

$$\mathbf{p} = q\mathbf{r}.$$

  The integral for the average field is the same as the integral we would obtain if we were calculating the field in $\mathbf{r}$ due to a uniform charge distribution $\rho = -q/(\frac{3}{4}\pi R^3)$ in the entire sphere. To solve this second scenario we can make use of Gauss' law because we know that the direction of $\mathbf{E}$ is $\hat{r}$.

$$\begin{aligned} \oint \mathbf{E} \cdot \mathrm{d}\mathbf{a} &= \int E\, \mathrm{d}a \\ = E4\pi r^2 \ &= \frac{Q_{\text{enc}}}{\epsilon_0} \\ &= \rho \frac{4}{3}\pi r^3 \end{aligned}$$

So we get

$$\mathbf{E} = \frac{\rho}{3}\mathbf{r} = -\frac{\mathbf{p}}{4\pi\epsilon_0 R^3}$$

We can split any arbitrary charge configuration up into point charges and calculate the average field for each one. Using the superposition principle, the total average field is just the sum of all those contributions, and the total dipole moment is the sum of all the individual dipole moments. Thus for an arbitrary charge configuration inside a sphere we have

$$\langle \mathbf{E} \rangle = -\frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}}{R^3}.$$

- Now for the average field **due to external charges**. We follow the same procedure, except now the charge q is at a point $\mathbf{r}$ outside the sphere and the charge enclosed in the gaussian surface is simply $-q$.

  Therefore the average field due to a single point charge is

  $$-\frac{q}{4\pi\epsilon_0 r^2}\hat{r}.$$

  This is the same as the field in the centre of the sphere. Again we can use the superposition principle to generalise. In general *the average field over the sphere due to all charges outside is the same as the field they produce in the centre.*

2.9.2 Earnshaw's theorem

TODO

Chapter 3

Electric fields in matter

In this section we will be studying electric fields in matter of which there are many forms. Most matter however can be approximated as either a conductor or an insulator.

- In a conductor the electrons are free to move wherever they like within the conductor. In an ideal conductor there is an unlimited supply of charged particles that can move.
- In an insulator, also called a dielectric, the charges are not free to move wherever they like.

  A dielectric consists of a collection of electrically neutral units, which can be atoms or molecules. These units however can contain charges that can move relative to each other, *within the confines of the unit.*

3.1 Conductors

We can deduce the following properties of (ideal) conductors:

1. $\mathbf{E} = 0$ **inside a conductor**. If this were not the case, then the electrons would feel a force and move to counteract the field. We cannot use electrostatics to describe this process, but we do know that once it settles down into an equilibrium condition, the field inside the conductor must be zero.
2. $\rho = 0$ **inside a conductor**. This follows from Gauss' law: $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho/\epsilon_0$. This is not, however, the same in one or two dimensions.
3. **Any net charge resides on the surface.** If the field outside the conductor is not zero, the charges must gather at the edges to counteract the effect of the external field. TODO Image.
4. **A conductor is an equipotential**:

   $$V(\mathbf{b}) - V(\mathbf{a}) = -\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{E} \cdot \mathrm{d}\mathbf{l} = 0$$

   for any two points $\mathbf{a}$ and $\mathbf{b}$ inside the conductor.
5. **Just outside a conductor, E is perpendicular to the surface**. Otherwise the tangential component would cause charges to flow along the surface, meaning we weren't in an electrostatic equilibrium.

3.1.1 Faraday cages

Assume we have a conductor with a cavity that does not contain any charges. The field inside the cavity must be zero: every field line must start and end at the cavity wall, as there are no charges from which the field lines can originate. If we follow a loop as in figure TODO, we see that the path integral $\oint \mathbf{E} \cdot \mathrm{d}\mathbf{l}$ is a sum of a positive part from inside the cavity and a vanishing part from inside the conductor. Thus the loop integral would be positive, but from equation (2.2), we know that it must be zero. Thus the field inside the cavity must be zero.
We derived this fact without knowing what charges there are outside the conductor. Consequently, no matter the electrical charges outside the conductor, inside the cavity you are shielded. In fact in practice the enclosing conductor does not even have to be solid. A mesh conductor also works. This is the principle behind a Faraday cage.

3.1.2 Capacitors

Suppose we have two (randomly shaped) conductors and we put a charge $+Q$ on one and a charge $-Q$ on the other. The potential is constant over a conductor (it's an equipotential), so we can speak unambiguously of the potential difference V between them.
Because $\mathbf{E}$ is proportional to ρ (which is proportional to Q) and V is proportional to $\mathbf{E}$, V is proportional to Q. The proportionality constant is called the capacitance C.

$$C \equiv \frac{Q}{V}$$

The capacitance is a purely geometric quantity, dependent on the sizes, shapes and separation of the two conductors. In SI units, it is measured in farads (F), which are coulomb-per-volts. This unit is quite large. More practical units are micro- and picofarads.
We give expressions for the capacitance of two common geometries.

1. A parallel-plate capacitor consists of two metal plates of area A, held a distance d apart. The capacitance is given by
$$C = \frac{A\epsilon_0}{d}$$
2. The capacitance of two concentric spherical metal shells with radii a and b is given by
$$C = 4\pi\epsilon \frac{ab}{(b-a)}.$$

Finally we also would like to have an idea of the energy stored in a capacitor. We charge the capacitor by moving electrons from the positive conductor to the negative conductor. Summing over all these infinitesimal bits of charge we are moving, we get

$$W = \int_0^Q V(q)\,\mathrm{d}q = \int_0^Q \left(\frac{q}{C}\right)\mathrm{d}q = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}CV^2.$$

In the last expression V is the final potential of the capacitor.

3.2 Dielectrics

Even though charges cannot flow though dielectrics, the constituent chunks can be given a dipole moment by putting them in an electric field.

1. The external field can induce a dipole by moving the positive charges to one side and the negative charges to the other side of the molecule or atom.

2. The external electric field can also rotate the dipoles so they align.

3.2.1 Induced dipoles

3.2.1.1 Atoms.

If we take a neutral atom and place it in an electric field, the charges inside it will move: the electrons will tend to be more on one side and the positive nucleus will tend to favour the other one. In effect we are *polerising* the atom and giving it a dipole moment in the same direction as $\mathbf{E}$. Typically the induced dipole moment depends linearly on the external field

$$\mathbf{p} = \alpha \mathbf{E}.$$

We call α the atomic polarizability. We can think of this as being a bit like Hooke's law. It is an approximation which breaks down if the electric field is strong.

3.2.1.2 Molecules.

Atoms are spherically symmetric. And so the induced dipole will be the same, whatever the direction of the electric field. The same cannot be said for molecules. So we cannot take α to be a scalar. Instead it is a map that maps the electric field vector onto the polarisation of the molecule. We can still assume the map to be linear though. In other words, we have a (1,1)-tensor: α is the polarisability tensor. (TODO use chosen tensor notation)

3.2.2 Alignment of polar molecules

Atoms, being spherically symmetric do not have an inherent dipole moment. Any induced dipole is naturally aligned with the electric field. Molecules may have an inherent dipole moment, in which case they will feel a torque unless they align themselves with the field. Such molecules are called polar molecules.
We can calculate the torque about the centre of the molecule, assuming $\mathbf{E}$ is more or less uniform over the length of the molecule (TODO image clarifying symbols):

$$\begin{aligned}\mathbf{M} &= [(\mathbf{d}/2) \times (q\mathbf{E}) + (-\mathbf{d}/2) \times (-q\mathbf{E})] \\ &= q\mathbf{d} \times \mathbf{E} \\ &= \mathbf{p} \times \mathbf{E}\end{aligned}$$

If the electric field is not uniform, there is a net force on the molecule, given by

$$\mathbf{F} = (\mathbf{p} \cdot \boldsymbol{\nabla})\mathbf{E}.$$

3.2.3 The field inside a dielectric

If a dielectric is placed in an electric field, the effects described above cause its atoms and / or molecules to polarise. This is a messy and complicated process that is counteracted by thermal effects. The net result is captured in a quantity called the polarisation P which is the dipole moment per unit volume.

3.2.3.1 Microscopic and macroscopic field

At first glance, the introduction of the polarisation may seem quite natural, but on closer inspection there are some subtleties.
The dipoles induced in the material are physical dipoles, not pure ones. This means they have higher multipole moments. If we are considering the field far away from the material, this is not a problem. The higher multipole terms become vanishingly small in comparison. Unfortunately we want to describe the electric field inside the dielectric as well and there these higher multipole moments most definitely do matter.
Thinking about it, we may realise that the electric field inside a dielectric must be extremely complicated. Close to an electron or a nucleus the field may be extremely strong, but just a short distance away the field may be completely different. This **microscopic** field fluctuates a lot and is incredibly difficult to calculate.
We want to average this field over volumes large enough that the fluctuations get smoothed out, but small enough that we do not lose all spatial variation and that the average field does not depend on the size of the volumes we average over. The field averaged in this way is called the **macroscopic field**.
Picture TODO qualitatively shows the average electric field over a sphere in function of the radius of the sphere. When the radius is very small the average field fluctuates a lot, because the inclusion of an extra electron or positive nucleus makes a lot of difference. At slightly larger radii the average field does not depend on the radius: if we increase the the radius a bit the added volume is close enough to all the other parts of the sphere that the external field will induce an average field that is approximately the same as the average field in the rest of the sphere. At even larger radii the average field again depends on the radius. This time it is because if we increase the radius slightly, the volume added to the sphere is so far away from the centre that the external field may be different and it may have a significantly different average field than other parts of the sphere.
In fact we do exactly the same averaging when we talk about the density of a material. This process will be brought up in that context when we will discuss fluid mechanics (TODO: check location).
Now we need to show that this macroscopic field is the field we obtain from the polarisation $\mathbf{P}$. Suppose we want to calculate the macroscopic field at a point $\mathbf{r}$ inside the dielectric. We take a sphere of the right size, as discussed above, about $\mathbf{r}$. The average field inside the sphere then consists of two parts: then average field due to all the charges outside plus the average field due to all the charges inside the sphere:

$$\mathbf{E} = \mathbf{E}_{\text{out}} + \mathbf{E}_{\text{in}}$$

Referring to section 2.9.1, we see that

- The average field due to the charges outside the sphere is equal to the field strength in the centre of the sphere. These charges are far enough away that we can safely use the dipole approximation:
$$V_{\text{out}} = \frac{1}{4\pi\epsilon_0} \int_{\text{outside}} \frac{\hat{\boldsymbol{\imath}} \cdot \mathbf{P}(\mathbf{r}')}{\imath^2} \, \mathrm{d}\tau'. \tag{3.1}$$
- The average field due to the charges inside the sphere is, regardless of their distribution, the same as that of a uniformly polarised sphere:
$$\mathbf{E}_{\text{in}} = -\frac{1}{4\pi\epsilon_0} \frac{\mathbf{p}}{R^3}$$

where $\mathbf{p}$ is the total dipole moment of the sphere: $\mathbf{p} = (\frac{4}{3}\pi R^3)\mathbf{P}$.

This is exactly the term that would be added if the integral in equation (3.1) were extended over all space.

In conclusion the macroscopic field is given by

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int_{\mathcal{V}} \frac{\hat{\boldsymbol{\imath}}\cdot\mathbf{P}(\mathbf{r}')}{\imath^2}\,\mathrm{d}\tau'$$

where the integral runs over the entire volume of the dielectric. This is exactly the expression we would get if we assumed the electric field inside a dielectric was due to each volume element $\mathrm{d}\tau$ having a dipole moment $\mathbf{P}\,\mathrm{d}\tau$.

3.2.3.2 Bound charges

With a bit of manipulation, we can write the expression for the potential of the macroscopic field as the sum of a potential due to a surface charge and a volume charge.
We start by observing that

$$\frac{\hat{\boldsymbol{\imath}}}{\imath^2} = \nabla'\left(\frac{1}{\imath}\right)$$

where the prime means that the differentiation is with respect to the source coordinates ($\mathbf{r}'$). Then the manipulation is as follows:

$$\begin{aligned}
V &= \frac{1}{4\pi\epsilon_0}\int_{\mathcal{V}} \frac{\hat{\boldsymbol{\imath}}\cdot\mathbf{P}(\mathbf{r}')}{\imath^2}\,\mathrm{d}\tau' \\
&= \frac{1}{4\pi\epsilon_0}\int_{\mathcal{V}} \mathbf{P}\cdot\boldsymbol{\nabla}'\left(\frac{1}{\imath}\right)\mathrm{d}\tau' \\
&= \frac{1}{4\pi\epsilon_0}\left[\int_{\mathcal{V}} \boldsymbol{\nabla}'\cdot\left(\frac{\mathbf{P}}{\imath}\right)\mathrm{d}\tau' - \int_{\mathcal{V}}\frac{1}{\imath}\left(\boldsymbol{\nabla}'\cdot\mathbf{P}\right)\mathrm{d}\tau'\right] \\
&= \frac{1}{4\pi\epsilon_0}\oint_{\mathcal{S}}\frac{1}{\imath}\mathbf{P}\cdot\mathrm{d}\mathbf{a}' - \frac{1}{4\pi\epsilon_0}\int_{\mathcal{V}}\frac{1}{\imath}\left(\boldsymbol{\nabla}'\cdot\mathbf{P}\right)\mathrm{d}\tau'
\end{aligned}$$

where we have first integrated by parts and then used the divergence theorem.
The first term looks like the potential of a surface charge

$$\sigma_b \equiv \mathbf{P}\cdot\hat{n}$$

(where $\hat{n}$ is the unit normal vector) and the second term looks like the potential of a volume charge

$$\rho_b \equiv -\nabla\cdot\mathbf{P}.$$

We can treat the field caused by the polarisation of matter as being generated by the bound charges ρ_b and σ_b.

3.2.3.3 Physical interpretation

The above manipulations are quite abstract, but there is also a much more physical (if less rigorous) way to derive the expressions for the bound charges.
First we calculate the surface charge σ_b by considering a tube through the dielectric made up of strings of back-to-back infinitesimal dipoles, as illustrated in figure TODO 4.11. The net

result is the accumulation of charge at the ends of the strings. For a small section of tube (in which $\mathbf{P}$ constant) with faces perpendicular to the strings of dipoles, the dipole moment can be written as both $P \cdot A \cdot d$, with d the length of the tube section, and $q \cdot d$ with q the charge accumulation at one end. Equating those two expressions, we get $q = PA$, so

$$\sigma_b = \frac{q}{A} = P.$$

Allowing the face to be oblique, we get

$$\sigma_b = P \cos\theta = \mathbf{P} \cdot \hat{n}.$$

Considering tube element at the surface, we see that this must also be an expression for the surface charge of the dielectric, and indeed it corresponds to the expression we have already derived.
If the polarisation is nonuniform there must also be accumulations of bound charges within the dielectric, because adjacent tube elements will have slightly different surface charges that do not quite cancel.
The net bound charge in a given volume $\int \rho_b \, \mathrm{d}\tau$ is equal and opposite to the amount that has been pushed out through the surface, which we have already reasoned to be $\mathbf{P} \cdot \hat{n}$ per unit area, so

$$\int_{\mathcal{V}} \rho_b \, \mathrm{d}\tau = - \oint_{\mathcal{S}} \mathbf{P} \cdot \mathrm{d}\mathbf{a} = - \int_{\mathcal{V}} (\nabla \cdot \mathbf{P}) \, \mathrm{d}\tau.$$

Since this is true for any volume, we have

$$\rho_b = - \nabla \cdot \mathbf{P}.$$

Luckily this again confirms our previous findings.

3.2.4 The electric displacement field D

Now we consider both the bound charge due to the polarised dielectric, and the charge that caused the external electric field in the first place, which we call the free charge ρ_f . This is basically any charge that is not a result of the polarisation of the dielectric. The total charge density can then be written

$$\rho = \rho_b + \rho_f$$

and Gauss's law reads

$$\epsilon_0 \nabla \cdot \mathbf{E} = \rho = \rho_b + \rho_f = - \nabla \cdot \mathbf{P} + \rho_f.$$

Combining the divergence terms gives

$$\rho_f = \nabla \cdot (\epsilon_0 \mathbf{E} + \mathbf{P}) = \nabla \cdot \mathbf{D}$$

where

$$\mathbf{D} \equiv \epsilon_0 \mathbf{E} + \mathbf{P}$$

is known as the electric displacement. In integral form Gauss's law now reads

$$\oint \mathbf{D} \cdot \mathrm{d}\mathbf{a} = Q_{f,\mathrm{enc}}$$

where $Q_{f,\mathrm{enc}}$ is the total free charge enclosed in the volume.
At this point one may think that $\mathbf{D}$ is like $\mathbf{E}$, and try to use other equations containing $\mathbf{E}$ with $\mathbf{D}/\epsilon_0$ instead, substituting ρ_f for ρ. This does in general not work for equations other than Gauss's law. In particular the curl of $\mathbf{E}$ is always zero, but the curl of $\mathbf{D}$ is not:

$$\nabla \times \mathbf{D} = \epsilon_0 (\nabla \times \mathbf{E}) + (\nabla \times \mathbf{P}) = \nabla \times \mathbf{P}.$$

3.2.5 Linear dielectrics

Provided $\mathbf{E}$ is not too strong, the polarisation is often proportional to the electric field:

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}$$

where the constant of proportionality χ_e is called the electric susceptibility. (Extracting the factor ϵ_0 makes χ_e dimensionless). Dielectrics for which this is the case are called linear dielectrics. It is important to note that $\mathbf{E}$ is the total field and thus *itself depends on* $\mathbf{P}$. This makes it difficult to use this formula to use this formula to calculate $\mathbf{P}$. It is usually simpler to start with the displacement field $\mathbf{D}$.
In linear media we have

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon_0 (1 + \chi_e)\mathbf{E} = \epsilon \mathbf{D}.$$

Thus for linear dielectrics $\mathbf{D}$ is also proportional to $\mathbf{E}$ with a proportionality constant

$$\epsilon \equiv \epsilon_0 (1 + \chi_e)$$

we call the permittivity of the material. (In a vacuum, where there is no matter to polarise, $\chi_e = 0$ and thus the permittivity is ϵ_0, which is why it is called the permittivity of vacuum).
We also define the dimensionless constant

$$\epsilon_r \equiv \frac{\epsilon}{\epsilon_0} = 1 + \chi_e$$

which is called the relative permittivity or dielectric constant. In later sections we will sometimes drop the subscript and refer to the relative permittivity simply as ϵ.
In crystals, some directions are easier to polarise than others, but in any one direction the relation is still linear. In this case the susceptibility is a $(1,1)$-tensor. TODO tensor notation.

3.2.5.1 Boundary value problems with linear dielectrics

3.2.5.2 Energy in dielectric systems

3.2.5.3 Forces on dielectrics

Chapter 4

Magnetostatics

TODO summary triangle

4.1 Experimental evidence

So far we have developed all of our equations based on Coulomb's law, which is only valid for stationary source charges. If we allow the charges to move, interesting new phenomena occur. For example, say we have a piece of wire made out of some conductive material, like copper. In the section on electrostatics, we said that a conductor must be an equipotential, because otherwise charges would flow in order to make it one. Now say we hold the two ends of the wire at different potentials (for example, by connecting them to a battery), then obviously charges must flow and continue flowing. How they flow will be discussed in more detail in the section on electrodynamics.
Now something interesting occurs when we put two such wire next to each other. We observe that the attract or repel each other (depending on whether the charges are flowing in the same direction or not). TODO wires neutral TODO explain Lorentz law + historical evidence for B. TODO: magnet-magnet interaction?? Also no work?

4.2 The Lorentz force law

Combining the electric and magnetic forces on a charged particle, we get the Lorentz force law

$$\mathbf{F} = Q[\mathbf{E} + (\mathbf{v} \times \mathbf{B})].$$

Notice that this formula is not really complete. It assumes we know how to calculate the fields $\mathbf{E}$ and $\mathbf{B}$. In electrostatics we can calculate $\mathbf{E}$ from Coulomb's law. We need an analogue for the field $\mathbf{B}$. As with electrostatics our task is simpler if we consider a steady state. In magnetostatics we study such steady states.
Once we have found $\mathbf{E}$ and $\mathbf{B}$, the Lorentz force law is generally valid, even outside the electro- or magnetostatic régimes.
(TODO: $\mathbf{B}$ pseudovector)

4.2.1 Cyclotron motion

Assume there is a uniform magnetic field of magnitude $\mathbf{B}$. Now a charged particle of charge Q enters the field with speed $v_\perp$ perpendicular to $\mathbf{B}$. The Lorentz force is then perpendicular to

both the velocity and the magnetic field. It thus functions as a centripetal force satifying the cyclotron formula:

$$Qv_\perp B = m\frac{v_\perp^2}{R} \qquad \text{or} \qquad p_\perp = QBR$$

where R is the radius of curvature, m the particles mass and $p_\perp = mv_\perp$ its momentum.

4.2.2 Magnetic forces do no work

The magnetic force on a charged particle is always perpendicular to the direction of motion, due to the cross product. Consequently magnetic forces do no work:

$$\mathrm{d}W_{\text{mag}} = \mathbf{F}_{\text{mag}} \cdot \mathrm{d}\mathbf{l} = Q(\mathbf{v} \times \mathbf{B}) \cdot \mathbf{v}\,\mathrm{d}t = 0$$

because $(\mathbf{v} \times \mathbf{B}) \cdot \mathbf{v}$ is zero.

4.3 Currents

TODO superposition of currents
The current in a wire is the *charge per unit time* passing a given point. Current is measured in coulombs-per-second, or amperes:

$$1\text{A} = 1\text{C/s}$$

Line current. A line charge λ traveling along a wire at a speed $\mathbf{v}$ constitutes a current

$$\mathbf{I} = \lambda\mathbf{v}.$$

Current is actually a vector, even if its vectorial character is often not important (cfr. circuit diagrams). The magnetic force on a segment of current-carrying wire is then

$$\mathbf{F}_{\text{mag}} = \int (\mathbf{v} \times \mathbf{B})\,\mathrm{d}q = \int (\mathbf{v} \times \mathbf{B})\lambda\,\mathrm{d}l = \int (\mathbf{I} \times \mathbf{B})\,\mathrm{d}l$$

which we can rewrite as

$$\mathbf{F}_{\text{mag}} = I\int (\mathrm{d}\mathbf{I} \times \mathbf{B}).$$

Surface current. Imagine a ribbon of current of infinitesimal width $\mathrm{d}l_\perp$ flowing along a surface. If the current in this ribbon is $\mathrm{d}\mathbf{I}$, the surface current density $\mathbf{K}$ (i.e. current per unit width) is

$$\mathbf{K} \equiv \frac{\mathrm{d}\mathbf{I}}{\mathrm{d}l_\perp}.$$

If a surface current density σ is moving at a velocity $\mathbf{v}$, then

$$\mathbf{K} = \sigma\mathbf{v}.$$

The magnetic force on the surface current is

$$\mathbf{F}_{\text{mag}} = \int (\mathbf{v} \times \mathbf{B})\sigma\,\mathrm{d}a = \int (\mathbf{K} \times \mathbf{B})\,\mathrm{d}a.$$

(TODO caveat: Just as $\mathbf{E}$ suffers a discontinuity at a surface charge, so $\mathbf{B}$ is discontinuous at a surface current. So here the average field must be used.)

Volume current density. Consider a tube of infinitesimal cross section $\mathrm{d}a_\perp$ running parallel to the flow. If the current in the tube is $\mathrm{d}\mathbf{I}$, the volume current density $\mathbf{J}$ (i.e. current per unit area) is

$$\mathbf{J} \equiv \frac{\mathrm{d}\mathbf{I}}{\mathrm{d}a_\perp}$$

If a volume current density ρ is moving at a velocity $\mathbf{v}$, then

$$\mathbf{J} = \rho\mathbf{v}.$$

The magnetic force on the surface current is

$$\mathbf{F}_{\mathrm{mag}} = \int (\mathbf{v} \times \mathbf{B})\rho \, \mathrm{d}\tau = \int (\mathbf{J} \times \mathbf{B}) \, \mathrm{d}\tau.$$

4.3.1 Charge conservation

Local charge conservation is expressed by the continuity equation

$$\boxed{\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}}$$

4.3.2 Steady currents

This section is called magnetostatics, but the Lorentz force law is generally true. So far there has been nothing particularly static, in fact we can only consider magnetic effects if there are *moving* charges. In magnetostatics we study the effects of currents that do not change in time. We can call them steady currents. Formally steady currents are defined by the condition

$$\frac{\partial \rho}{\partial t} = 0, \qquad \frac{\partial \mathbf{J}}{\partial t} = 0$$

at all places and all times. From charge conservation we also get

$$\nabla \cdot \mathbf{J}$$

which means in a wire I mus be the same all along the wire; otherwise, charge would be piling up somewhere and it would not be a steady current.
Point charges can never constitute as steady current.

4.3.3 The Biot-Savart law

The magnetic field of a steady line current is given by the Biot-Savart law

$$\boxed{\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{I} \times \hat{\boldsymbol{\imath}}}{\imath^2} \, \mathrm{d}l' = \frac{\mu_0}{4\pi} I \int \frac{\mathrm{d}\mathbf{l}' \times \hat{\boldsymbol{\imath}}}{\imath^2}}$$

For surface and volume currents, the Biot-Savart law becomes

$$\boxed{\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{K}(\mathbf{r}') \times \hat{\boldsymbol{\imath}}}{\imath^2} \, \mathrm{d}a'} \qquad \text{and} \qquad \boxed{\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}') \times \hat{\boldsymbol{\imath}}}{\imath^2} \, \mathrm{d}\tau'}.$$

The constant μ_0 is called the permeability of free space:

$$\mu_0 = 4\pi \times 10^{-7} \mathrm{N/A}^2.$$

4.4 The divergence and curl of B

In the following calculations, the integration is over the *primed* coordinates; the divergence and curl are with respect to the *unprimed* coordinates. Let $\mathbf{J}(\mathbf{r'})$ depend only on primed coordinates, such that divergences and curls of $\mathbf{J}$ vanish.
We start with the Biot-Savart law

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r'}) \times \hat{\boldsymbol{\imath}}}{\imath^2} \, \mathrm{d}\tau'$$

to which we apply the divergence:

$$\nabla \cdot \mathbf{B} = \frac{\mu_0}{4\pi} \int \nabla \cdot \left(\mathbf{J} \times \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \right) \mathrm{d}\tau'.$$

Invoking the product rule,

$$\begin{aligned} \nabla \cdot \left(\mathbf{J} \times \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \right) &= \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \cdot (\nabla \times \mathbf{J}) - \mathbf{J} \cdot \left(\nabla \times \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \right) \\ &= 0. \end{aligned}$$

So the divergence of the magnetic field is zero

$$\boxed{\nabla \cdot \mathbf{B} = 0}$$

Starting again with the Biot-Savart law

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r'}) \times \hat{\boldsymbol{\imath}}}{\imath^2} \, \mathrm{d}\tau'$$

we apply the curl:

$$\nabla \cdot \mathbf{B} = \frac{\mu_0}{4\pi} \int \nabla \times \left(\mathbf{J} \times \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \right) \mathrm{d}\tau'.$$

Again using the relevant product rule:

$$\begin{aligned} \nabla \times \left(\mathbf{J} \times \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \right) &= \left(\frac{\hat{\boldsymbol{\imath}}}{\imath^2} \cdot \boldsymbol{\nabla} \right) \mathbf{J} - (\mathbf{J} \cdot \boldsymbol{\nabla}) \frac{\hat{\boldsymbol{\imath}}}{\imath^2} + \mathbf{J} \left(\nabla \cdot \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \right) - \frac{\hat{\boldsymbol{\imath}}}{\imath^2} (\nabla \cdot \mathbf{J}) \\ &= - (\mathbf{J} \cdot \boldsymbol{\nabla}) \frac{\hat{\boldsymbol{\imath}}}{\imath^2} + \mathbf{J} \left(\nabla \cdot \frac{\hat{\boldsymbol{\imath}}}{\imath^2} \right) \end{aligned} \tag{4.1}$$

The first term integrates to zero. Because $\boldsymbol{\imath}$ only depends on the difference between coordinates and $(\frac{\partial}{\partial x}) f(x - x') = -(\frac{\partial}{\partial x'}) f(x - x')$, we can write

$$- (\mathbf{J} \cdot \boldsymbol{\nabla}) \frac{\hat{\boldsymbol{\imath}}}{\imath^2} = \left(\mathbf{J} \cdot \boldsymbol{\nabla}' \right) \frac{\hat{\boldsymbol{\imath}}}{\imath^2}.$$

Now taking the x component $\left[\left(\mathbf{J} \cdot \boldsymbol{\nabla}' \right) \left(\frac{x - x'}{\imath^3} \right) \right]$, the product rule for $\boldsymbol{\nabla}' \cdot \left[\frac{(x - x')}{\imath^3} \mathbf{J} \right]$ can be rearranged to give

$$\left(\mathbf{J} \cdot \boldsymbol{\nabla}' \right) \left(\frac{x - x'}{\imath^3} \right) = \boldsymbol{\nabla}' \cdot \left[\frac{(x - x')}{\imath^3} \mathbf{J} \right] - \left(\frac{x - x'}{\imath^3} \right) \left(\boldsymbol{\nabla}' \cdot \mathbf{J} \right).$$

For steady currents the divergence of $\mathbf{J}$ is zero, so

$$\left[-\left(\mathbf{J}\cdot\boldsymbol{\nabla}\right)\frac{\hat{\boldsymbol{\imath}}}{\imath^2}\right]_x = \boldsymbol{\nabla}'\cdot\left[\frac{(x-x')}{\imath^3}\mathbf{J}\right].$$

Integrating the first term of equation (4.1), yields

$$\int_{\mathcal{V}}\boldsymbol{\nabla}'\cdot\left[\frac{(x-x')}{\imath^3}\mathbf{J}\right]\mathrm{d}\tau' = \oint_{\mathcal{S}}\frac{(x-x')}{\imath^3}\mathbf{J}\cdot\mathrm{d}\mathbf{a}'.$$

If we integrate over a volume large enough such that all the current is inside the volume and $\mathbf{J}$ is zero on the boundary, the integral vanishes. (This typically still holds if $\mathbf{J}$ extends to infinity, as in the case of an infinite straight wire).
Thus the curl of $\mathbf{B}$ is obtained by integrating over the second term of equation (4.1):

$$\begin{aligned}\nabla\times\mathbf{B} &= \frac{\mu_0}{4\pi}\int\mathbf{J}(\mathbf{r}')\left(\nabla\cdot\frac{\hat{\boldsymbol{\imath}}}{\imath^2}\right)\mathrm{d}\tau' \\ &= \frac{\mu_0}{4\pi}\int\mathbf{J}(\mathbf{r}')4\pi\delta^3(\mathbf{r}-\mathbf{r}')\,\mathrm{d}\tau' \\ &= \mu_0\mathbf{J}(\mathbf{r})\end{aligned}$$

4.4.1 Ampère's law

The equation for the curl of $\mathbf{B}$,

$$\boxed{\nabla\times\mathbf{B}} = \mu_0\mathbf{J}$$

is called Ampère's law (in differential form). It can be written in integral form by integrating over a surface of your choosing and using Stokes' theorem:

$$\int(\nabla\times\mathbf{B})\cdot\mathrm{d}\mathbf{a} = \oint\mathbf{B}\cdot\mathrm{d}\mathbf{l} = \mu_0\int\mathbf{J}\cdot\mathrm{d}\mathbf{a}.$$

Now, $\int\mathbf{J}\cdot\mathrm{d}\mathbf{a}$ is the total current passing through the surface, which we call I_{enc}. It is the current enclosed by the Amperian loop. Thus

$$\boxed{\oint\mathbf{B}\cdot\mathrm{d}\mathbf{l} = \mu_0 I_{\mathrm{enc}}}$$

Like Gauss's law, Ampère's law is always *true* (for steady currents), but not always *useful*. It is only useful if there is enough symmetry. Some situations where it is useful:

1. Infinite straight lines: disk perpendicular to line, centered on line.
2. Infinite planes: rectangle perpendicular to plane and to $\mathbf{K}$.
3. Infinite solenoids: rectangle with sides parallel to direction of solenoid; one outside and one half in, half out.
4. Toroids: the magnetic field of a toroid is circumferential at all points, both inside and outside the coil; take a circle about the axis of the toroid.

4.5 Magnetic vector potential

Just as $\nabla \times \mathbf{E}$ permitted us to introduce the scalar potential V in electrostatics, so $\nabla \cdot \mathbf{B}$ allows the introduction of a vector potential $\mathbf{A}$ in magnetostatics:

$$\boxed{\mathbf{B} = \nabla \times \mathbf{A}}$$

There is some freedom in determining $\mathbf{A}$: you can add any function whose curl is zero (i.e. the gradient of any scalar). This freedom can be used to eliminate the divergecne of $\mathbf{A}$:

$$\nabla \cdot \mathbf{A} = 0.$$

In terms of the vector potential, Ampère's law becomes

$$\boxed{\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J}.}$$

This is just three Poisson's equations. Assuming $\mathbf{J}$ goes to zero at infinity, the solution is

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r'})}{\mathscr{r}} \, \mathrm{d}\tau'.$$

For line and surface currents,

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{I}}{\mathscr{r}} \, \mathrm{d}l' = \frac{\mu_0 I}{4\pi} \int \frac{1}{\mathscr{r}} \, \mathrm{d}\mathbf{l'}; \qquad \mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{K}}{\mathscr{r}} \, \mathrm{d}a'. \tag{4.2}$$

4.5.1 Multipole expansion

As before we wish to write the vector potential as a power series in $1/r$. To the end we recall

$$\frac{1}{\mathscr{r}} = \frac{1}{\sqrt{r^2 + (r')^2 - 2rr'\cos\theta'}} = \frac{1}{r} \sum_{n=0}^{\infty} \left(\frac{r'}{r}\right)^n P_n(\cos\theta').$$

Using this, we can write the vector potential of a current loop as

$$\begin{aligned}
\mathbf{A}(\mathbf{r}) &= \frac{\mu_0 I}{4\pi} \oint \frac{1}{\mathscr{r}} \, \mathrm{d}\mathbf{l'} = \frac{\mu_0 I}{4\pi} \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \oint (r')^n P_n(\cos\theta') \, \mathrm{d}\mathbf{l'} \\
&= \frac{\mu_0 I}{4\pi} \left[\frac{1}{r} \oint \mathrm{d}\mathbf{l'} + \frac{1}{r^2} \oint r' \cos\theta' \, \mathrm{d}\mathbf{l'} + \frac{1}{r^3} \oint (r')^2 \left(\frac{3}{2}\cos^2\theta' - \frac{1}{2} \right) \mathrm{d}\mathbf{l'} + \ldots \right]
\end{aligned}$$

As before, we call the first term the monopole term. For the vector potential this term vanishes, because

$$\oint \mathrm{d}\mathbf{l'} = 0$$

is just the vector displacement around a closed loop. The absence of magnetic monopoles is also expressed by the Maxwell's equation $\nabla \cdot \mathbf{B} = 0$, which is what allowed us to define the vector potential in the first place.
The dominant term is usually the dipole term

$$\begin{aligned}
\mathbf{A}_{\mathrm{dip}}(\mathbf{r}) &= \frac{\mu_0 I}{4\pi r^2} \oint r' \cos\theta' \, \mathrm{d}\mathbf{l'} = \frac{\mu_0 I}{4\pi r^2} \oint (\hat{r} \cdot \mathbf{r'}) \, \mathrm{d}\mathbf{l'} \\
&= \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \hat{r}}{r^2}
\end{aligned}$$

where $\mathbf{m}$ is the magnetic dipole moment

$$\mathbf{m} \equiv I \int \mathrm{d}\mathbf{a}$$

We can construct a pure magnetic dipole by putting a current loop around the origin and shrinking it until it is infinitesimal, while keeping $\mathbf{m}$ constant. For a pure dipole we have

$$\mathbf{A}_{\mathrm{dip}}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{m \sin\theta}{r^2} \hat{\phi}$$
$$\mathbf{B}_{\mathrm{dip}}(\mathbf{r}) = \nabla \times \mathbf{A} = \frac{\mu_0 m}{4\pi r^3} (2\cos\theta \hat{r} + \sin\theta \hat{\theta})$$

4.6 Discontinuity of magnetic field at surface current

TODO

4.7 Miscellaneous results in magnetostatics

4.7.0.1 Helmholtz coil

4.7.0.2 Hall effect

Chapter 5

Magnetic fields in matter

TODO

5.1 Magnetisation

5.2 The field inside a magnetised object

5.2.1 Bound currents

5.3 The auxiliary field H

5.4 Linear and nonlinear media

Chapter 6

Electrodynamics

To make a current flow, you need to push the charges. To account for that we need to allow forces to act on charges. We define $\mathbf{f}$ as the force per unit charge it is acting on.
We can split $\mathbf{f}$ into two parts: the forces generated by a source (such as a battery, photoelectric cell or Van de Graaff generator), $\mathbf{f}_s$; and the electrostatic forces the charges exert on each other, which is just $\mathbf{E}$.

$$\mathbf{f} = \mathbf{f}_s + \mathbf{E}$$

6.1 Electromotive force

We now define the electromotive force or emf of a circuit as

$$\boxed{\mathcal{E} \equiv \oint \mathbf{f} \cdot \mathrm{d}\mathbf{l} = \oint \mathbf{f}_s \cdot \mathrm{d}\mathbf{l}.}$$

We can use $\mathbf{f}$ or $\mathbf{f}_s$ because $\oint \mathbf{E} \cdot \mathrm{d}\mathbf{l} = 0$. The name electromotive *force* is not ideal as it is actually an integral of a force per unit charge, not a force. In fact it can be interpreted as the work done per unit charge (although remember magnetic forces never do any work).
Within an ideal voltage source, the net force on the charges is zero (there is just enough force to push the electrons through the circuit), so $\mathbf{E} = -\mathbf{f}_s$. Outside the source $\mathbf{f}_s = 0$. The potential difference between the terminals (a and b) is therefore

$$V = -\int_a^b \mathbf{E} \cdot \mathrm{d}\mathbf{l} = \int_a^b \mathbf{f}_s \cdot \mathrm{d}\mathbf{l} = \oint \mathbf{f}_s \cdot \mathrm{d}\mathbf{l} = \mathcal{E}.$$

6.1.1 Motional emf

Motional emfs arise when you move a wire through a magnetic field. In order to calculate the emf we use a new quantity, the flux of $\mathbf{B}$ through a loop:

$$\Phi \equiv \int \mathbf{B} \cdot \mathrm{d}\mathbf{a}$$

Now assume that we have a loop of wire at time t; then advance to time $t + \mathrm{d}t$. The loop has now moved. TODO image.

$$\mathrm{d}\Phi = \Phi(t + \mathrm{d}t) - \Phi(t) = \Phi_{\text{ribbon}} = \int_{\text{ribbon}} \mathbf{B} \cdot \mathrm{d}\mathbf{a}.$$

Let P be a point on the wire. Let $\mathbf{v}$ be the velocity of the wire at that point, $\mathbf{u}$ the velocity of a charge along the wire at that point and $\mathbf{w} = \mathbf{v} + \mathbf{u}$ is the resultant velocity of a charge at P. The infinitesimal area of the ribbon can be written as

$$\mathrm{d}\mathbf{a} = (\mathbf{v} \times \mathrm{d}\mathbf{l})\,\mathrm{d}t.$$

Therefore

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \oint \mathbf{B} \cdot (\mathbf{v} \times \mathrm{d}\mathbf{l}).$$

Since $\mathbf{w} = (\mathbf{v} + \mathbf{u})$ and $\mathbf{u}$ is parallel to $\mathrm{d}\mathbf{l}$, we an just as well write this as

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \oint \mathbf{B} \cdot (\mathbf{w} \times \mathrm{d}\mathbf{l}).$$

Using the scalar triple-product identity we get

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = -\oint (\mathbf{B} \times \mathbf{w}) \cdot \mathrm{d}\mathbf{l}.$$

But $(\mathbf{w} \times \mathbf{B})$ is the magnetic force per unit charge, $\mathbf{f}_{\mathrm{mag}}$, so

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = -\oint \mathbf{f}_{\mathrm{mag}} \cdot \mathrm{d}\mathbf{l},$$

and the integral of $\mathbf{f}_{\mathrm{mag}}$ is the emf:

$$\boxed{\mathcal{E} = -\tfrac{\mathrm{d}\Phi}{\mathrm{d}t}}$$

6.2 Electromagnetic induction

6.2.1 Faraday's law

We have shown that moving a wire in a magnetic field can create a (motional) emf. Michael Faraday also reported on experiments where the wire was kept fixed and the *magnetic field* was moved or changed. In these cases Faraday also measured a current due to an emf.
This came as a surprise, because the force generating the emf cannot be magnetic since all the charges are stationary and sationary charges do not experience magnetic forces. But one would think there was no electric field because the wire is balanced and neutral. Faraday had an ingenious inspiration:

A changing magnetic field induces an electric field.

Empirically Faraday found the emf generated by the induced field to follow the same law as the motional emf:[1]

$$\mathcal{E} = \oint \mathbf{E} \cdot \mathrm{d}\mathbf{l} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} = -\int \frac{\partial \mathbf{B}}{\partial t} \cdot \mathrm{d}\mathbf{a}$$

This is Faraday's law, in integral form. We can convert it to differential form by applying Stokes' theorem:

$$\boxed{\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.}$$

This reduces to the old rule of $\oint \mathbf{E} \cdot \mathrm{d}\mathbf{l} = 0$ or $\nabla \times \mathbf{E} = 0$ in the static case as it should.

[1] Classically it is quite a surprising coincidence. The mechanism for Faraday's law is different from that for motional emf: one is electric, the other magnetic. That they coincide is a coincidence, but one that motivated, and is explained by, Einstein's theory of relativity.

6.2.1.1 Lenz's law

It is often difficult to keep track of signs in Faraday's law. Lenz's law gives you a way to work out which way an induced current will flow:

Nature abhors a *change* in flux.

The induced current will flow in such a direction that the flux *it* produces tends to cancel the change.

6.2.1.2 The quasistatic régime

Electromagnetic induction only occurs when there are changing magnetic fields. Often, however, we will want to use the apparatus of magnetostatics to calculate those changing fields. The is, of course, only approximately correct, but is often very close. This *quasistatic approximation* only seriously breaks down when working with electromagnetic waves and radiation.

6.2.2 Inductance

Assume we have two loops of wire at rest. If you run a steady current I_1 around loop 1, it produces a magnetic field $\mathbf{B}_1$. Some of the field lines pass through loop 2; let Φ_2 be the flux of $\mathbf{B}_1$ through the area of loop 2. From the Biot-Savart law we know that $\mathbf{B}_1$ is proportional to I_1, therefore so too is the flux though loop 2. Thus

$$\Phi_2 = M_{21} I_1$$

where M_{21} is the constant of proportionality which we call the mutual inductance of the two loops.
We can derive a formula for the mutual inductance as follows:

$$\begin{aligned} M_{21} = \frac{\Phi_2}{I_1} = \frac{1}{I_1}\int \mathbf{B}_1 \cdot \mathrm{d}\mathbf{a}_2 &= \frac{1}{I_1}\int (\nabla \times \mathbf{A}_1)\cdot \mathrm{d}\mathbf{a}_2 \\ &= \frac{1}{I_1}\oint \mathbf{A}_1 \cdot \mathrm{d}\mathbf{l}_2 \\ &= \frac{1}{I_1}\oint \left[\frac{\mu_0 I_1}{4\pi}\oint \frac{\mathrm{d}\mathbf{l}_1}{\mathscr{r}}\right]\cdot \mathrm{d}\mathbf{l}_2 \\ &= \frac{\mu_0}{4\pi}\oint_{\text{loop 1}}\oint_{\text{loop 2}} \frac{\mathrm{d}\mathbf{l}_1 \cdot \mathrm{d}\mathbf{l}_2}{\mathscr{r}}. \end{aligned}$$

using equation (4.2) for the vector potential due to a line current. This is the Neumann formula. It reveals two important facts about mutual inductance:

1. M_{21} is a purely geometric quantity, having to do with the sizes, shapes and relative positions of the two loops.

2. We can switch the roles of the two loops without changing the integral; it follows that

   $$M_{21} = M_{12}.$$

   We drop the subscripts and call them both M.

If we vary the current in loop 1, the changing flux in loop 2 will induce an emf according to Faraday's law:

$$\mathcal{E}_{\in} = -\frac{\mathrm{d}\Phi_2}{\mathrm{d}t} = -M\frac{\mathrm{d}I_1}{\mathrm{d}t}.$$

(This assumes we are in the quasistatic régime.)
A changing current in a loop induces an emf not only in nearby loops, but also in itself. We can repeat the above analysis taking loop 1 and loop 2 to be the same loop. We now denote the constant of proportionality between the flux and the current as L and we call it the self inductance (or simply inductance):

$$\Phi = LI \qquad \mathcal{E} = -L\frac{\mathrm{d}I}{\mathrm{d}t}.$$

Inductance is measured in henries (H); a henry is a volt-second per ampere.
Due to Lenz's law, inductance is an intrinsically positive quantity. The induced emf opposes the change in current. For this reason it is called a *back emf*.

6.2.2.1 Calculating the inductance of various configurations

TODO

6.3 Energy in magnetic fields

6.4 Maxwell's equations

Chapter 7

Conservation laws

7.1 Charge

7.2 Energy

7.3 Momentum

7.3.1 Angular momentum

Chapter 8

Potentials and fields in electrodynamics

Chapter 9

Electromagnetic waves

9.1 Light waves

Phenomenology, Polerization, polerizing filter
intensity square of amplitude

9.2 Electromagnetic waves in vacuum

9.3 Electromagnetic waves in matter

9.4 Absorption and dispersion

9.5 Guided waves

Chapter 10

Radiation

10.1 Dipole radiation

10.2 Radiation of point charges

Chapter 11

Gaussian system of units

Part VIII

Optics

pinhole camera specularity (diffuse / ...) geometric optical theory colours and theory rainbows X-ray crystallography (Laue diffraction) interferometer retroreflectors Lambert-beer

Chapter 1

Geometrical optics

Chapter 2

Wave optics

2.1 Interference

2.2 Diffraction

2.2.1 Single opening

2.2.2 Double slit

2.2.3 Diffraction grating

2.2.4 Fresnel diffraction

2.3 Absorption and scattering

2.4 Dispersion

2.5 Reflection

2.6 Double refraction

2.7 Polarised light

2.8 Thermal radiation

When looking at objects, we can see them because there is electromagnetic radiation that comes from the objects and travels to our eyes. At normal temperatures this radiation is light from a light source that was reflected by the object. We see the object as a certain colour because the object does not reflect light of all wavelengths equally.

Part IX

Special relativity

The physics we have described so far is what is known as classical physics and all the major conceptual ideas had been stated in some form or other by the end of the nineteenth century (Maxwell's equations were first published between 1861 and 1862). A story that is often repeated, including in introductions to quantum mechanics textbooks, is that there was a general idea shared by many eminent physicists of the time that the grand underlying principles of physics had all been discovered. Admittedly based on a very minimal amount of research, I am not at all sure this is true. The following quote is from a lecture by Lord Kelvin at the Royal Institution on April 27, 1900.

> The beauty and clearness of the dynamical theory, which asserts heat and light to be modes of motion, is at present obscured by two clouds. I. The first came into existence with the undulatory theory of light, and was dealt with by Fresnel and Dr. Thomas Young; it involved the question, how could the earth move through an elastic solid, such as essentially is the luminiferous ether? II. The second is the Maxwell–Boltzmann doctrine regarding the partition of energy. [1]

This is often paraphrased in a way that suggests that Lord Kelvin thought basically all of physics was solved except these two small clouds. In reality he almost certainly realized how big and (in his words) dense these clouds were.
The first one prompted the development of special (and later general) relativity. The second cloud required quantum mechanics to be resolved. Investigations into these new areas of physics produced lots of interesting new physics and the question of how to unite these two areas is still an open one.
We will now have a closer look at the first of the two clouds.

[1] From a 1900, April 27, Royal Institution lecture. "Lord Kelvin, Nineteenth Century Clouds over the Dynamical Theory of Heat and Light", reproduced in Notices of the Proceedings at the Meetings of the Members of the Royal Institution of Great Britain with Abstracts of the Discourses, Volume 16, p. 363–397 and Philosophical Magazine, Sixth Series, 2, 1–40 (1901).

Chapter 1

Origins and justification

TODO: review after chapter on EM
Before the theory of relativity, electric and magnetic fields were thought of as strains in an invisible medium called the (luminiferous) ether. Light waves were thought to be waves propagating through this medium, like sound waves through air.
This would imply that depending on the speed we are traveling relative to this ether, electromagnetic experiments would give different results. Looking at the laws of electrodynamics, this would seem to be correct: for example, a charge in motion produces a magnetic field, whereas a charge at rest does not; several electrodynamic laws even make explicit reference to the velocity of the charge.
Thus we might expect there to be a unique stationary reference frame, with respect to which all velocities are to be measured. In other words we expect the principle of relativity we saw in classical mechanics to not hold true in electrodynamics. Also, it is of utmost importance to find this reference frame, otherwise all our calculations become invalid.
One might suppose that finding this unique stationary reference frame would be easy. It is the only reference frame in which Maxwell's equations are valid. Therefor surely we need only conduct a simple electromagnetic experiment; any discrepancy between the experimental findings and predictions based on Maxwell's equations would indicate movement with respect to the ether frame.
Unfortunately this proved to be quite difficult. Consider for example the following simple experiment:

Example

[TODO: better example]
Suppose we put a wire loop on a train. This train then rides at constant velocity between the poles of a giant magnet.
As the loop passes through the magnetic field, an electromotive force is established according to the flux rule (TODO reference eq.),

$$\mathcal{E} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}.$$

If we now look at the situation on board the train. There would be no magnetic force, because the loop is at rest. As the magnet flies by, it would induce an electric field, according to Faraday's law (TODO reference eq). The resulting electric force would generate the

following electromotive force in the loop:

$$\mathcal{E} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}.$$

Because Faraday's law and the flux rule predict exactly the same electromotive force, people in both reference frames will get the same answer, even though the physical interpretation of the process is completely different.

The problem with the example experiment above proved quite general. When changing (inertial) reference frame, the physical interpretation changed, but "coincidences" would conspire to make the values of measurable observables the same. So it takes an uncommonly delicate experiment to do the job (TODO?).

1.1 Michelson-Morley experiment

This brings us to the famous Michelson-Morley experiment. Conducted between April and July 1887 by Albert Michelson and Edward Morley.
One of the predictions of classical electrodynamics is that electromagnetic waves travel through the vacuum at a speed

$$\frac{1}{\sqrt{\epsilon_0 \mu_0}} = 3.00 \times 10^8 \mathrm{m/s}$$

relative (presumably) to the ether.
Michelson and Morley used an interferometer to measure the speed of light in different directions. If we were moving through the ether, then we would expect light to be slower in the direction we were traveling in. Unfortunately Michelson and Morley, despite going to great lengths to achieve maximum precision, were not able to detect any difference in speed. This presented quite a problem. Over the years many unsatisfactory theories were devised to explain why our movement with respect to the ether was undetectable by the Michelson-Morley experiment, such as "ether drag" and various so-called "emission" theories.
It was not until Einstein that anybody took the result of the Michelson-Morley experiment at face value and suggested that the speed of light is a universal constant.

1.2 Einstein's postulates

Inspired by the Michelson-Morley experiment and the fact that applying the laws of electrodynamics in different inertial reference frames yields the same values for experimental observables, Einstein proposed his two famous postulates:

1. **The principle of relativity**: The laws of physics apply in all inertial reference systems.
2. **The universal speed of light**: The speed of light in a vacuum, c, is the same for all inertial observers, regardless of the motion of the source.

(TODO: second postulate redundant?) Thus the principle of relativity, which in classical mechanics is an observation, is elevated to the status of a general law. The second postulate asserts that there is no ether. This was radically new and has some strange consequences.
TODO: name lightspeed misleading, no object, no signal, nothing faster than the speed of light!!!

1.3 Gedankenexperimente and phenomena

If we accept Einstein's postulates, we are forced to accept that some strange phenomena can occur. In this section we explore thought experiments exhibiting some of the most striking new phenomena.

1.3.1 Velocity addition

In classical mechanics, if one walks at 7 km/h towards the front of a train going 100 km/h, one's speed with respect to the ground 107 km/h. This result was so obvious that until Einstein nobody has bothered to give it a name. Einstein called it Galileo's velocity addition rule. In general we can write it as follows

$$\mathbf{v}_{AB} = \mathbf{v}_{AB} + \mathbf{v}_{BC}$$

where $\mathbf{v}_{AB}$ represents the velocity of A relative to B.
If A is a light signal, this rule cannot hold, as Einstein's second postulate states that

$$v_{AC} = v_{AB} = c$$

regardless of v_{BC}. Instead we have Einstein's velocity addition rule (assuming movement along one axis):

$$v_{AC} = \frac{v_{AB} + v_{BC}}{1 + (v_{AB}v_{BC}/c^2)}.$$

We will derive this properly later as well as give a more complete treatment of relativistic kinematics, but for now we will just make a couple of remarks:

1. For speeds encountered in everyday life (i.e. $v_{AB} \ll c$ and $v_{BC} \ll c$), the denominator is so close to one that the difference between the two velocity addition formulae is negligible.

2. If $v_{AB} = c$, then we automatically have that $v_{AC} = c$:

$$v_{AC} = \frac{c + v_{BC}}{1 + (cv_{BC}/c^2)} = c.$$

1.3.2 The relativity of simultaneity

If we accept Einstein's postulates we are forced to abandon the idea of objective, universal simultaneity. Two events that happen simultaneously according to one observer, do not necessarily happen at the same time according to a different one. Even more strangely, if according to one observer one event happens before another one, there may be a different observer in a different reference frame that sees the second event happen before the first.

Example

Imagine a freight train, traveling a constant speed along a smooth, straight track (TODO image). In the centre of the train there hangs a light bulb. At a certain time the light is switched on. The light leaves the light bulb at speed c. An observer in the same reference frame as the train will see the light reach both ends of the train at the same time, because the bulb is hanging in the centre. Thus the two events in question: light reaching the front end and light reaching the back end, happen simultaneously.

To an observer on the ground these events do not happen simultaneously. As the light travels backwards (still at lightspeed), the train moves forward and thus the light moving backwards has a shorter distance to travel. The observer on the ground observes the light reach the back before it reaches the front.

Conclusion:

Two events that are simultaneous in one inertial system are not, in general, simultaneous in another.

1.3.3 Time dilation

This effect can be summed up as follows:

Moving clocks run slow.

Example

Consider the same setup as before. We now want to know how long it takes for the light from the lamp to reach the floor of the train. For an observer in the train the answer is straightforward:

$$t_{\text{train}} = \frac{h}{c}$$

where t_{train} is the time it takes the light to reach the floor and h is the height of the train. According to the ground observer however the light has to travel $\sqrt{h^2 + (v \cdot t_{\text{ground}})^2}$, see figure TODO. Thus we get

$$t_{\text{ground}} = \frac{\sqrt{h^2 + (v \cdot t_{\text{ground}})^2}}{c}$$

where v is the speed the freight train is traveling at and t_{ground} is the time it takes the light to reach the floor according to an observer on the ground.
Solving this equation for t_{ground}, we get

$$t_{\text{ground}} = \frac{h}{c}\frac{1}{\sqrt{1 - v^2/c^2}} = t_{\text{train}}\frac{1}{\sqrt{1 - v^2/c^2}} = t_{\text{train}}\gamma$$

Where we have introduced the factor

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

This result is true in general. Clocks moving at a speed v will run a factor

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

slower according to a stationary observer compared to the time measured by a comoving observer.

One might think this leads to a paradox. In the above example the freight train is moving with respect to the ground, so an observer on the ground observes time on the train as moving slower. Conversely the ground is moving with respect to the freight train and thus an observer on the freight train observes time on the ground as moving slower.
This may seem contradictory, but it actually is not. The trick is that the two observers are actually measuring different things. (TODO: explanation?)

1.3.3.1 Twin paradox.

Imagine we have twins. One of them becomes an astronaut and flies off in a rocket. Because that twin is flying fast, they will experience time dilation, meaning that from the point of view of twin who stayed at home time moves more slowly on the space ship. Thus when the astronaut returns home, they will be younger than the twin who stayed on earth.
One might be tempted to think that this reasoning can be reversed. From the point of view of the astronaut the earth is moving and so time is moving more slowly on earth, meaning that the earth bound twin would be younger. This line of reasoning is false however: the twin on the ground stays in a (roughly) inertial reference frame, whereas the astronaut has to accelerate and decelerate, thus the traveling twin is not a stationary observer.

1.3.4 Length contraction

This effect, also known as Lorentz contraction, can be summed up as follows:

Moving objects are shortened.

Example

Imagine now our lamp is at the back of the freight train and we have set up a mirror at the front. How long does it take light to travel from the back of the train to the mirror in the front and beck to where it started?
To an observer on the train the answer is

$$t_{\text{train}} = 2\frac{l_{\text{train}}}{c}$$

where l_{train} denotes the length of the train as measured by an observer on the train.
To an observer on the ground the process is more complicated.
TODO figure
We call the time, as observed by an observer on the ground, for the light signal to reach the mirror $t_{\text{ground},1}$ and the return time $t_{\text{ground},2}$. These quantities obey the following equations:

$$t_{\text{ground},1} = \frac{l_{\text{ground}} + vt_{\text{ground},1}}{c}, \qquad t_{\text{ground},2} = \frac{l_{\text{ground}} - vt_{\text{ground},2}}{c}$$

Solving for $t_{\text{ground},1}$ and $t_{\text{ground},2}$ we get

$$t_{\text{ground},1} = \frac{l_{\text{ground}}}{c - v}, \qquad t_{\text{ground},2} = \frac{l_{\text{ground}}}{c + v}$$

so the round trip time is

$$t_{\text{ground}} = t_{\text{ground},1} + t_{\text{ground},2} = 2\frac{l_{\text{ground}}}{c}\frac{1}{1 - v^2/c^2}$$

We have already derived the relationship between time intervals:

$$t_{\text{train}} = \sqrt{1 - v^2/c^2} t_{\text{ground}}$$

We conclude that

$$l_{\text{train}} = \frac{1}{\sqrt{1 - v^2/c^2}} l_{\text{ground}}$$

Lorentz contraction works very much like time dilation, except moving rulers are shorter. Many of the same remarks apply.

1.3.4.1 Dimensions perpendicular to the velocity

Objects are only shortened along the direction of its motion! We have so far taken this fact for granted. We will now justify it with another thought experiment.

Example

Imagine there is a wall next to the train tracks with a blue line 5 m above the surface, as measured by an observer on the ground. As the train goes by, a passenger paints a red line 5 m above the ground, as measured by that passenger on the train. The question now is whether the red line is above or below the blue one.
Were there contraction in the vertical direction, an observer on the ground would say that the red line is lower and the passenger would say that the blue line is lower. They cannot both be right, thus we conclude there cannot be any contraction along the vertical axis.

1.3.4.2 Ehrenfest's paradox.

TODO

Chapter 2

The mathematics of special relativity

Having discussed some phenomena of special relativity, we now want to derive a more precise mathematical description.
So far we have talked about events, such as light hitting a particular part of the train, without rigorously defining the concept. An event is something that takes place at a specific location (x, y, z) and at a specific time (t). The coordinates (x, y, z, t) of an event are only specified relative to a reference frame $\mathcal{S}$.

2.1 The Lorentz transformations

Suppose we know the coordinates (x, y, z, t) of an event E in an inertial reference frame $\mathcal{S}$ and we would like to know what the coordinates $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$ of E are in a different inertial reference frame $\bar{\mathcal{S}}$.
We can orient the axes of the frames so that the spatial axes are parallel and $\bar{\mathcal{S}}$ is moving away from $\mathcal{S}$ along the x axis at a speed v.
We start time (i.e. set $t = 0$) at the moment the origins of both frames coincide. Then, according to classical theory, we can transform coordinates from $\mathcal{S}$ to $\bar{\mathcal{S}}$ using the Galilean transformations:

$$\begin{cases} \bar{x} = x - vt \\ \bar{y} = y \\ \bar{z} = z \\ \bar{t} = t \end{cases}$$

Until Einstein these were considered self-evident. Based on the thought experiments of the previous section, we know enough to replace these transformations with the Lorentz transformations:

$$\begin{cases} \bar{x} = \gamma(x - vt) \\ \bar{y} = y \\ \bar{z} = z \\ \bar{t} = \gamma\left(t - \frac{v}{c^2}x\right) \end{cases}$$

The inverse transformations can be found by switching the sign of v (or by solving for x and t):

$$\begin{cases} x = \gamma(\bar{x} - v\bar{t}) \\ y = \bar{y} \\ z = \bar{z} \\ t = \gamma\left(\bar{t} - \frac{v}{c^2}\bar{x}\right) \end{cases}$$

The critical information to remember when deriving these equations is

1. $l_{\text{train}} > l_{\text{ground}}$;
2. $t_{\text{train}} < t_{\text{ground}}$;
3. Time and space are scaled by a factor $\gamma > 1$.

2.1.1 Einstein velocity addition

TODO

2.2 The structure of spacetime

TODO: coordinate with section on mathematics
In this section we introduce some mathematical methods to make it easier to deal with Lorentz transformations.

2.2.1 Four-vectors

In order to simplify notation we introduce the dimensionless quantity

$$\beta \equiv \frac{v}{c}.$$

We then define an object with four components (labeled with subscripts) which we call a four-vector:

$$x = (x^0, x^1, x^2, x^3) \equiv (ct, x, y, z)$$

Notice that we rescale the time dimension with c. This means we measure x^0 in units of length, like the other components, so we can easily compare and combine all components.
The Lorentz transformation considered above becomes:

$$\begin{cases} \bar{x}^0 = \gamma(x^0 - \beta x^1) \\ \bar{x}^1 = \gamma(x^1 - \beta x^0) \\ \bar{x}^2 = x^2 \\ \bar{x}^3 = x^3 \end{cases} \qquad \text{or, in matrix form} \qquad \begin{pmatrix} \bar{x}^0 \\ \bar{x}^1 \\ \bar{x}^2 \\ \bar{x}^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

Letting the Greek indices run from 0 to 3, we can now write this more compactly as

$$\bar{x}^\mu = \sum_{\nu=0}^{3} \left(\Lambda^\mu_{\ \nu}\right) x^\nu$$

2.2.1.1 Consequences of Einstein summation

The formula above for $\bar{x}^\mu$ may be considerably simplified if we use the Einstein summation convention.
TODO: guaranteed and manifestly in-/covariant?
TODO: 4-vector (spatial part in bold) transforms under Lorentz transformations. Minkowsky product invariant, covariant, contravariant, Einstein summation
TODO: Greek 0 to 3 and Latin 1 to 3.
TODO: raise / lower indices with η.
TODO: merge with mathematics.

2.2.2 The invariant interval

TODO: show that the interval stays invariant under Lorentz transformations. use that to define Lorentz transformations. -¿ like fourth type of rotation. Add translation Poincaré. Any transformation $x^\mu \to x'^\mu$ that preserves length given by metric η is Poincaré:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + \epsilon^\mu$$

2.2.3 Space-time diagrams

The invariant interval between causally related events is timelike, and their temporal ordering is the same for all inertial observers.
TODO Hartle 4.3 world line

2.2.3.1 Conformal transformations

Leave speed of light invariant

2.2.4 Light cones

spacelike, null, timelike separated

2.2.5 Differing conventions

mostly minus and mostly plus

2.3 Symmetries

What are symmetries of Minkowski space? What transformations leave $\mathrm{d}s^2$ invariant?

- Translations
- Rotations

  $$\Lambda^\intercal \eta \Lambda = \eta \qquad \eta_{\rho\sigma}$$

  Analogous to

  $$R^\intercal R = \mathbb{1}$$

  Lorentz group $\emptyset(3,1)$.

Examples of Λ:

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \Lambda = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

In general:

- 4 translations
- 3 rotations
- 3 boosts

Poincaré: all. Lorentz only rotations and boosts.

Chapter 3

Relativistic mechanics

3.1 Proper time and proper velocity

As we have seen there is no absolute notion of time. Each observer has their own clock that all run differently. There is however a clock that all observers can agree on: the clock attached to the object being observed. We call this time the proper time τ.
Due to time dilation, we know that the proper time intervals are given by

$$\mathrm{d}\tau = \sqrt{1 - v^2/c^2}\,\mathrm{d}t$$

This proper time gives us a very natural way to parametrise the world line of a particle, so we might describe the world line by the equations

$$x^\alpha = x^\alpha(\tau).$$

We can now define the proper velocity u^α as the derivatives of the position along the world line with respect to the proper time:

$$u^\alpha \equiv \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}.$$

Because x^α is a good four-vector and τ is invariant (i.e. all observers can agree on it), the four-velocity does in fact transform as a four-vector and the scalar product with itself is invariant.
If we now take the inertial frame centered around $x^\alpha(\tau)$ at a certain proper time τ moving with the same velocity as the instantaneous velocity of the particle, we see that the scalar product of u^α with itself must be

$$u^\alpha u_\alpha = \eta_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = -c^2$$

TODO: better derivation with proper velocity also if not zero rest mass
TODO:

$$\tau_{AB} = \int_A^B \mathrm{d}\tau = \int_A^B \left(-\frac{\mathrm{d}s^2}{c^2}\right)^{1/2}$$

$$\tau = \int \left(-\eta_{\mu\nu}\tfrac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\tfrac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right)^{1/2}\mathrm{d}\lambda$$

TODO define $\mathbf{v}$

$$(\Delta\tau)^2 = -(\Delta s)^2 = -\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu$$

If you do not move:

$$\Delta\tau = \Delta t$$

3.2 Relativistic energy and momentum

TODO $\mathbf{v} = \mathbf{p}/E$
Inspired by the classical definition of momentum, we can now define relativistic momentum as follows:

$$\mathbf{p} \equiv m\mathbf{u} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}}$$

We can also define the relativistic energy

$$E \equiv mcu^0 = \frac{mc^2}{\sqrt{1 - v^2/c^2}}$$

which we can identify with cp^0. Thus we can define the energy-momentum 4-vector (or just momentum 4-vector) as

$$\boxed{p^\mu \equiv mu^\mu = (E/c, \mathbf{p})}.$$

As an historical aside: Einstein originally defined the relativistic mass

$$m_{\text{rel}} \equiv \frac{m}{\sqrt{1 - v^2/c^2}}.$$

This definition differs from E by a constant factor c^2, and thus is fairly redundant.
We have not yet justified our definitions. In the classical limit ($u \ll c$), it is clear that $\mathbf{p} = m\mathbf{v}$ is just the classical momentum.
The energy is slightly more tricky. A first thing to verify is that it has the dimension of energy ($[ML^2T^{-2}]$), which it does. We can also notice that the relativistic energy is nonzero even when the object is stationary. We call this the rest energy:

$$E_{\text{rest}} \equiv mc^2$$

The remainder, which is attributable to the motion, we call the kinetic energy:

$$E_{\text{kin}} \equiv E - mc^2 = mc^2 \left(\frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right).$$

In the classical limit ($v \ll c$ or $v/c \ll 1$) we can do a Taylor expansion in powers of v^2/c^2:

$$E_{\text{kin}} = \frac{1}{2}mv^2 + \frac{3}{8}\frac{mv^4}{c^2} + \dots$$

The leading term is the classical formula for the kinetic energy.
Importantly we also have the following experimental fact (TODO???):

In every closed system, the total relativistic energy and momentum are conserved.

Finally we can take the scalar product of p^μ with itself. This quantity is not only conserved, but also scalar and thus invariant.

$$p^\mu p_\mu = -(p^0)^2 + (\mathbf{p} \cdot \mathbf{p}) = -m^2c^2$$

Multiplying this by c^2, we get the famous equation

$$\boxed{E^2 - p^2c^2 = m^2c^4}$$

3.3 Relativistic kinematics

3.3.1 Photon momenta

In classical mechanics there is no such thing as a massless particle: you could not apply a force to it ($\mathbf{F} = m\mathbf{a}$) and hence (by Newton's third law) it could not apply a force on anything else. It's kinetic energy ($\frac{1}{2}mv^2$) and momentum ($m\mathbf{v}$) would be zero as well.
In relativistic mechanics there is a loophole however. A massless particle going at lightspeed would leave the equations

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}} \quad \text{and} \quad E = \frac{mc^2}{\sqrt{1 - v^2/c^2}}$$

indeterminate (zero over zero). It might be conceivable, therefore, that there exist massless particles with energy and momentum (which means forces can act on them, because $\mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$). Incredible as it may sound, this is in fact the case: the photon is massless but carries energy and momentum. Relativity cannot predict what the energy or momentum of a photon is, but it does force light to travel at lightspeed. Later we will see that quantum mechanics allows us to calculate the energy and momentum of photons.
One does have to be slightly more careful when dealing with massless particles though. The proper time interval between any two points is zero (as is the separation), which leads to some complications. Proper time can no longer be used to parametrize the world line of a massless particle.
For massless particles we can still define the tangent vector

$$u^\alpha \equiv \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}$$

where λ is a parameter. Different parametrizations will give different tangent vectors, but they are all null vectors. Thus for massless particles we have

$$u^\alpha u_\alpha = \eta_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = 0$$

We may choose any parameter to parametrize the curve, but there is a special class of parameters, called affine parameters that we like in particular. They are the parameters the that give

$$\frac{\mathrm{d}u^\alpha}{\mathrm{d}\lambda} = 0.$$

This makes (as we shall see) the equation of motion the same as that for particles with mass.
(TODO move section down? + wave four-vector + relativistic doppler and relativistic beaming)

3.3.2 Mass-energy conversions

In a classical collision, momentum and mass are always conserved. Kinetic energy, in general, is not. During a collision kinetic energy may be transformed into other types of energy, typically thermal (a sticky collision); or other types of energy may be turned into more kinetic energy (an explosive collision). If kinetic energy is conserved, we say the collision is elastic (see also the section on classical mechanics).
In the relativistic case, momentum and total energy are always conserved, but mass and kinetic energy are not. Any change in kinetic energy means a change in rest energy, because the total energy E is conserved and

$$E_{\text{rest}} = E - E_{\text{kin}}.$$

Any change in rest energy, $E_{\text{rest}} = mc^2$, means a change in mass. So changes in kinetic energy cause changes in mass.
In the classical view, if a system loses kinetic energy, we typically say it gets transformed into thermal energy and lost. This is no different in the relativistic case, except we now must except that it has also gained mass. In general hotter objects are (ever so slightly) heavier than cold ones.
We again call a collision elastic if kinetic energy is conserved. This means rest energy and mass must also be conserved.

3.3.3 Practical strategies for relativistic kinematics

TODO

3.4 Relativistic dynamics

We have so far translated some kinematic quantities into their relativistic counterparts. Now we see what happens to Newton's laws of motion.

1. Newton's first law is built into the principle of relativity.
2. Newton's second law can be written as follows:

$$\mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$$

   This is still valid in relativistic mechanics, provided we use the relativistic momentum.
3. Newton's third law does not, in general hold true in relativistic mechanics.

3.4.1 The work-energy theorem

TODO

3.4.2 Minkowski force

TODO

3.5 Variational principle for free particle motion

Chapter 4

Relativistic electrodynamics

Writing in covariant form.
Gauge invariance because of commuting derivatives
Wave equation in terms of A^μ
D'Alembertian!!

Chapter 5

The Lorentz and Poincaré groups

TODO: Action: The transformations $T(\Lambda, \epsilon)$ induce a unitary linear transformation on vectors in the physical Hilbert space

$$\Psi \to U(\Lambda, a)\Psi$$

5.1 The Lorentz group

The Lorentz group is the group of Lorentz transformations.

5.1.1 Definition

The Lorentz group is the isometry group of the Minkowski metric η

$$\mathrm{O}(1,3) = \left\{\Lambda \in \mathbb{R}^{4\times 4} \mid \Lambda^\intercal \eta \Lambda = \eta\right\}$$

5.1.2 Connectedness

Even though SO(4) is a connected group, SO(1, 3) is not a connected group. There are two disconnections, meaning that SO(1, 3) has 4 connected subsections.

1. Taking the determinant of $\Lambda^\intercal \eta \Lambda = \eta$, we get

   $$\det(\Lambda^\intercal \eta \Lambda) = \det(\Lambda^\intercal)\det(\eta)\det(\Lambda) = \det(\eta) \qquad \Rightarrow \qquad [\det(\Lambda)]^2 = 1.$$

   As a consequence either $\det(\Lambda) = 1$ or $\det(\Lambda) = -1$. There is no path from one to the other.

2. Taking the 00-component of the equation $\eta_{\rho\sigma} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta_{\mu\nu}$, we get

   $$\begin{aligned} 1 = \eta_{00} &= \Lambda^\mu{}_0 \Lambda^\nu{}_0 \eta_{\mu\nu} \\ &= \left(\Lambda^0{}_0\right)^2 \eta_{00} + \Lambda^i{}_0 \Lambda^i{}_0 \eta_{ii} = \left(\Lambda^0{}_0\right)^2 - \Lambda^i{}_0 \Lambda^i{}_0 \end{aligned}$$

   Rearranging yields

   $$\left(\Lambda^0{}_0\right)^2 = 1 + \sum_i \left(\Lambda^i{}_0\right)^2 \geq 1 \qquad \Rightarrow \qquad \begin{cases} \Lambda^0{}_0 \geq 1 \\ \Lambda^0{}_0 \leq 1 \end{cases}$$

So we can partition the elements of $\mathrm{O}(1,3)$ into 4 subsets of transformations, only one of which is a subgroup (because it contains the identity): the proper orthochronous Lorentz group Λ^{++} with $\det(\Lambda^{++}) = 1$ and $(\Lambda^{++})^0{}_0 \geq +1$.
Any Lorentz transformation is either part of Λ^{++}, or can be made part of it by composing it with one or both of the discrete transformations P or T.
The space inversion P and time-reversal matrices are given by

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The four parts of $\mathrm{O}(1,3)$ are

1. The set of proper, orthochronous Lorentz transformations, Λ^{++} or $\mathrm{SO}(1,3)_+$, with
$$\det \Lambda = +1, \quad \Lambda^0{}_0 \geq 1.$$
This is a subgroup.
2. The set of proper, non-orthochronous Lorentz transformations, Λ^{+-}, with
$$\det \Lambda = +1, \ \Lambda^0{}_0 \leq -1.$$
This can also be written as
$$\Lambda^{+-} = PT\Lambda^{++}$$
3. The set of improper, orthochronous Lorentz transformations, Λ^{-+}, with
$$\det \Lambda = -1, \ \Lambda^0{}_0 \geq 1.$$
This can also be written as
$$\Lambda^{-+} = P\Lambda^{++}$$
4. The set of improper, non-orthochronous Lorentz transformations, Λ^{--}, with
$$\det \Lambda = -1, \ \Lambda^0{}_0 \leq -1.$$
This can also be written as
$$\Lambda^{--} = T\Lambda^{++}$$

5.1.3 Compactness

The Lorentz group is also non compact (not every open cover has a finite subcover) (TODO)

$$\mathrm{SO}(2) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \to \quad \theta \in [0, 2\pi) \quad \to \quad \text{manifold is compact}$$

$$\mathrm{SO}(1,3) \ni \begin{pmatrix} \cosh\epsilon & -\sinh\epsilon & 0 & 0 \\ -\sinh\epsilon & \cosh\epsilon & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \epsilon \in (-\infty, +\infty) \qquad \text{(boost)}$$

$$= \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \gamma \in [1, +\infty]$$

5.2 The Lorentz algebra $\mathfrak{so}(1,3)$

Consider an infinitesimal Lorentz transformation (i.e. close to identity)

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} + \omega^{\mu}{}_{\nu} \qquad (\omega \ll 1)$$

The Lorentz condition ($\Lambda^{\intercal}\eta\Lambda = \eta$) gives in this case

$$\begin{aligned}\eta_{\mu\nu} &= \Lambda^{\rho}{}_{\mu}\Lambda^{\sigma}{}_{\nu}\eta_{\rho\sigma} = (\delta^{\rho}_{\mu} + \omega^{\rho}{}_{\mu})(\delta^{\sigma}_{\nu} + \omega^{\sigma}{}_{\nu})\eta_{\rho\sigma} \\ &= \eta_{\mu\nu} + \eta_{\mu\sigma}\omega^{\sigma}{}_{\nu} + \eta_{\rho\nu}\omega^{\rho}{}_{\mu} + \mathcal{O}(\omega^2) = \eta_{\mu\nu} + \omega_{\mu\nu} + \omega_{\nu\mu} + \mathcal{O}(\omega^2)\end{aligned}$$

From this we get the condition

$$\omega_{\mu\nu} = -\omega_{\nu\mu}$$

which leaves 6 parameters.
Now we write

$$\begin{aligned}\omega_{\mu\nu} = \frac{1}{2}\left(\omega_{\mu\nu} - \omega_{\nu\mu}\right) &= \frac{1}{2}\omega_{\alpha\beta}\left(\delta^{\alpha}_{\mu}\delta^{\beta}_{\nu} - \delta^{\alpha}_{\nu}\delta^{\beta}_{\mu}\right) \\ &= -\frac{i}{2}\omega_{\alpha\beta}\left(\mathcal{J}^{\alpha\beta}\right)_{\mu\nu}\end{aligned}$$

Where we define $\left(\delta^{\alpha}_{\mu}\delta^{\beta}_{\nu} - \delta^{\alpha}_{\nu}\delta^{\beta}_{\mu}\right)$ as $-i\left(\mathcal{J}^{\alpha\beta}\right)_{\mu\nu}$. So $\mathcal{J}^{\alpha\beta}$ are the generators of the Lorentz group in the defining representation and $\frac{-\omega_{\alpha\beta}}{2}$ are the coordinates of the algebra element with respect to this basis of generators. Therefore

$$\begin{aligned}\Lambda^{\mu}_{\ \nu} &= \delta^{\mu}_{\nu} - \frac{i}{2}\omega_{\alpha\beta}\left(\mathcal{J}^{\alpha\beta}\right)_{\mu\nu} \\ \Lambda &= e^{-\frac{i}{2}\omega_{\alpha\beta}\mathcal{J}^{\alpha\beta}}\end{aligned}$$

Example

Exercise: Prove the following:

$$[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\rho\sigma}] = i\left(\eta^{\mu\rho}\mathcal{J}^{\nu\sigma} - \eta^{\nu\rho}\mathcal{J}^{\mu\sigma} + \eta^{\nu\sigma}\mathcal{J}^{\mu\rho} - \eta^{\mu\sigma}\mathcal{J}^{\nu\rho}\right)$$

Answer:

$$\begin{aligned}[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\rho\sigma}]_{ij} &= \sum_k \left(\delta^{\mu}_{i}\delta^{\nu}_{k} - \delta^{\mu}_{k}\delta^{\nu}_{i}\right)\left(\delta^{\rho}_{k}\delta^{\sigma}_{j} - \delta^{\rho}_{j}\delta^{\sigma}_{k}\right) - \sum_k \left(\delta^{\rho}_{i}\delta^{\sigma}_{k} - \delta^{\rho}_{k}\delta^{\sigma}_{i}\right)\left(\delta^{\mu}_{k}\delta^{\nu}_{j} - \delta^{\mu}_{j}\delta^{\nu}_{k}\right) \\ &= \delta^{\mu}_{i}\delta^{\rho}_{\nu}\delta^{\sigma}_{j} - \delta^{\mu}_{i}\delta^{\nu}_{\sigma}\delta^{\rho}_{j} - \delta^{\mu}_{\rho}\delta^{\nu}_{i}\delta^{\sigma}_{j} + \delta^{\mu}_{\sigma}\delta^{\nu}_{i}\delta^{\rho}_{j} - \delta^{\rho}_{i}\delta^{\sigma}_{\mu}\delta^{\nu}_{j} + \delta^{\rho}_{\mu}\delta^{\sigma}_{i}\delta^{\nu}_{j} + \delta^{\rho}_{i}\delta^{\sigma}_{\nu}\delta^{\mu}_{j} - \delta^{\rho}_{\nu}\delta^{\sigma}_{i}\delta^{\mu}_{j}\end{aligned}$$

Using

$$\left(\eta^{\mu\rho}\mathcal{J}^{\nu\sigma}\right)_{ij} = -i\left(\delta^{\rho}_{\mu}\delta^{\nu}_{i}\delta^{\sigma}_{j} - \delta^{\rho}_{\mu}\delta^{\nu}_{j}\delta^{\sigma}_{i}\right)$$

We recover the equality. But we need to take $\eta^{\mu\rho} = \delta^{\rho}_{\mu}$???

We now are going to derive the algebra $[\cdot,\cdot]$ for a generic representation D.

1. $D(\Lambda) = D(\mathbb{1}_{\omega}) \equiv \mathbb{1} - \frac{i}{2}\omega_{\alpha\beta}J^{\alpha\beta}$
   Where $J^{\alpha\beta}$ are the generators in the representation D.

2. Group properties:

$$D(\Lambda)^{-1}D(\tilde{\Lambda})D(\Lambda) = D(\Lambda^{-1}\tilde{\Lambda}\Lambda)$$

For an infinitesimal transformation $\tilde{\Lambda}$ we have

$$D(\Lambda)^{-1}(\mathbb{1} - \frac{i}{2}\tilde{\omega}_{\alpha\beta}J^{\alpha\beta})D(\Lambda) = D(\mathbb{1} - \frac{i}{2}\tilde{\omega}_{\alpha\beta}\Lambda^{\alpha}_{\ \mu}\Lambda^{\beta}_{\ \nu}J^{\mu\nu})$$

So

$$\boxed{D(\Lambda)^{-1}J^{\alpha\beta}D(\Lambda) = \Lambda^{\alpha}_{\ \mu}\Lambda^{\beta}_{\ \nu}J^{\mu\nu}}$$

3. Linearize $\Lambda \approx \mathbb{1} + \omega$ (with $\omega \ll 1$)

$$\left(\mathbb{1} + \frac{i}{2}\omega_{\rho\sigma}J^{\rho\sigma}\right)J^{\alpha\beta}\left(\mathbb{1} - \frac{i}{2}\omega_{\rho\sigma}J^{\rho\sigma}\right) + \ldots = J^{\alpha\beta} + \frac{i}{2}\omega_{\rho\sigma}\left[J^{\rho\sigma}, J^{\alpha\beta}\right] + \ldots$$

$$= \left(\delta^{\alpha}_{\ \mu} + \omega^{\alpha}_{\ \mu}\right)\left(\delta^{\beta}_{\ \nu} + \omega^{\beta}_{\ \nu}\right)J^{\mu\nu} = J^{\alpha\beta} + \omega^{\alpha}_{\ \mu}J^{\mu\beta} + \omega^{\beta}_{\ \nu}J^{\alpha\nu} = J^{\alpha\beta} - \omega_{\rho\sigma}\left(\eta^{\alpha\rho}J^{\beta\sigma} - \eta^{\beta\rho}J^{\alpha\sigma}\right) + \ldots$$

So we get

$$[J^{\mu\nu}, J^{\rho\sigma}] = i\left(\eta^{\mu\sigma}J^{\nu\sigma} - \eta^{\nu\sigma}J^{\mu\sigma} + \eta^{\nu\sigma}J^{\mu\rho} - \eta^{\mu\sigma}J^{\nu\rho}\right)$$

This is the same as for the defining representation!

Example

Recap of some important points:

$$\Lambda^{\mu}_{\ \alpha}\Lambda^{\nu}_{\ \beta}\eta_{\mu\nu} = \eta_{\alpha\beta} \qquad \rightarrow \qquad \Lambda^{\intercal}\eta\Lambda = \eta \Rightarrow \Lambda^{-1} = \eta^{-1}\Lambda^{\intercal}\eta$$

$$\begin{cases} \eta = \eta_{\alpha\beta} \\ \eta^{-1} = \eta^{\alpha\beta} \end{cases}$$

$$[J^{\mu\nu}, J^{\rho\sigma}] = i\left(\eta^{\mu\sigma}J^{\nu\sigma} - \eta^{\nu\sigma}J^{\mu\sigma} + \eta^{\nu\sigma}J^{\mu\rho} - \eta^{\mu\sigma}J^{\nu\rho}\right)$$

Defines algebra of Lorentz group.

$$D(\Lambda)^{-1}D(\tilde{\Lambda})D(\Lambda) = D(\Lambda^{-1}\tilde{\Lambda}\Lambda)$$

Let's make the following redefinitions

$$\begin{cases} K_i \equiv J^{i0} & \rightarrow \quad \text{Boost generator} \\ J_i \equiv \frac{1}{2}\epsilon_{ijk}J^{jk} & \rightarrow \quad \text{3D rotations} \end{cases}$$

The algebra commutators in K_i and J_i are given by

$$\begin{cases} [J_i, J_j] = i\epsilon_{ijk}J_k \\ [J_i, K_j] = i\epsilon_{ijk}K_k \\ [K_i, K_j] = -i\epsilon_{ijk}J_k \end{cases}$$

We introduce a new basis to decouple the commutator relations

$$N^{(+)}_i \equiv \frac{J_i + iK_i}{2}, \qquad N^{(-)}_i \equiv \frac{J_i - iK_i}{2}$$

Then the commutator relations become

$$\begin{cases} \left[N_i^{(+)}, N_j^{(+)}\right] = i\epsilon_{ijk} N_k^{(+)} & \to \quad \text{SU(2) Which is classified by an index } j_+ \\ \left[N_i^{(-)}, N_j^{(-)}\right] = i\epsilon_{ijk} N_k^{(-)} & \to \quad \text{SU(2)} \to j_- \\ \left[N_i^{(+)}, N_j^{(-)}\right] = 0 \end{cases}$$

So Lorentz group representations are classified $j = (j_+, j_-)$

$$\textbf{dim } j = (2j_+ + 1)(2j_- + 1)$$

Locally

$$\mathrm{SO}(1,3) \approx \text{“}\mathrm{SU}(2) \times \mathrm{SU}(2)''$$

5.2.1 Simplest (irreducible) Lorentz group representations

(j_+, j_-)	Dim	
Irreducible		
$(0,0)$	1	Scalar particles (spin 0)
$(1/2,0)$	2	Weyl L fermion (spin 1/2)
$(0,1/2)$	2	Weyl R fermion (spin 1/2)
$(1/2,1/2)$	4	Vector
$(1,1)$	9	Tensor particle (spin 2)
Reducible		
$(1/2,0) \oplus (0,1/2)$	4	Dirac fermion (spin 1/2)

5.2.2 Casimir of Lorentz algebra

$$C_2 = \frac{1}{2} J^{\mu\nu} J_{\mu\nu} \qquad C_2 = \frac{1}{4} \epsilon_{\mu\nu\rho\sigma} J^{\mu\nu} J^{\rho\sigma}$$

5.3 The Poincaré group

The Poincaré group is the group of Poincaré transformations

$$x^\mu \to x'^\mu = \Lambda^\mu_{\ \nu} x^\nu + \epsilon^\mu \qquad \begin{cases} \Lambda = \mathrm{O}(1,3) \\ \epsilon^\mu \in \mathbb{R}^4 \end{cases}$$

TODO: Poincaré is actually a group.

The Poincaré group is defined as

$$P = \left\{ g(\Lambda, a) \quad | \quad \Lambda \in \mathbb{R}^{4\times 4}, \quad a \in \mathbb{R}^4, \quad \Lambda^\intercal \eta \Lambda = \eta \right\}$$

where $\eta = \mathrm{diag}(1,-1,-1,-1)$. With the following composition law:

$$\begin{aligned} g_2 \cdot g_1 &= (\Lambda_2, a_2)(\Lambda_1, a_1) = (\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2) \\ g^{-1} &= (\Lambda, a)^{-1} = (\Lambda^{-1}, -\Lambda^{-1} a) \end{aligned}$$

The Poincaré group looks like a direct product, but isn't quite one. We call it a semidirect product

$$P = \mathrm{O}(1,3) \rtimes \mathbb{R}^4$$

5.4 Poincaré algebra

TODO: Hamiltonian P^μ (or $-P^\mu$)
First we look at translations. They form a 4 dim invariant subgroup of $\mathbb{R}^4$

$$\begin{cases} x^\mu \to x'^\mu = x^\mu + \epsilon^\mu = (1 + i\epsilon^\nu \rho_\nu)\, x^\mu \\ \rho_\nu = -\partial_\nu, \qquad g_\epsilon = e^{i\epsilon^\mu \rho_\mu} \end{cases}$$

$$[\rho_\nu, \rho_\mu] = 0 \qquad \to \qquad \text{Translations commute}$$

Translations are also an abelian subgroup of P:

$$(\mathbb{1}, \epsilon_1) \cdot (\mathbb{1}, \epsilon_2) = (\mathbb{1}, \epsilon_1 + \epsilon_2) = (\mathbb{1}, \epsilon_2) \cdot (\mathbb{1}, \epsilon_1)$$

$$D^{-1}(g_\epsilon) D(g_{\tilde{\epsilon}}) D(g_\epsilon) = D(g_\epsilon^{-1} g_{\tilde{\epsilon}} g_\epsilon) = D(g_{\tilde{\epsilon}})$$

We introduce the generic representation $D(g_\epsilon) = e^{i\epsilon^\mu P_\mu}$ where P_μ are generators of D and we try to find properties of P_μ

$$e^{-i\epsilon P}\,(1 + i\tilde{\epsilon} P)\, e^{i\epsilon P} = 1 + i\tilde{\epsilon} P$$

$$e^{-i\epsilon P} i\tilde{\epsilon} P e^{i\epsilon P} = i\tilde{\epsilon} P$$

$$(1 - i\epsilon^\mu P_\mu)\, i\tilde{\epsilon}^\nu P_\nu\, (1 + i\epsilon^\mu P_\mu) = i\tilde{\epsilon}^\nu P_\nu$$

$$\cancel{i\tilde{\epsilon}^\nu P_\nu} - \epsilon^\mu \tilde{\epsilon}^\nu\, [P_\mu, P_\nu] + \mathcal{O}(\epsilon^2) = \cancel{i\tilde{\epsilon}^\nu P_\nu}$$

So

$$[P_\mu, P_\nu] = 0$$

Translations + rotations do, however, not commute. We write generic elements of rotations, translations and the Poincaré group in the following way

$$g_\Lambda(\Lambda, 0), \qquad g_\epsilon(\mathbb{1}, \epsilon), \qquad g(\Lambda, \epsilon)$$

And we see that

$$g_\Lambda g_\epsilon = (\Lambda, 0)(\mathbb{1}, \epsilon) = (\Lambda, \Lambda\epsilon)$$
$$g_\epsilon g_\Lambda = (\mathbb{1}, \epsilon)(\Lambda, 0) = (\Lambda, \epsilon)$$

These objects are not the same, so translations and rotation do not comute.
We now calculate the commutators

$$D(g_\Lambda)^{-1} D(g_\epsilon) D(g_\Lambda) = D(g_\Lambda^{-1} g_\epsilon g_\Lambda) = D(g_{\Lambda^{-1}\epsilon^{-1}} g_{\tilde{\epsilon}})$$

Which we expand for small $\epsilon \ll 1$:

$$D(g_\Lambda)^{-1}\,(\mathbb{1} + i\epsilon^\mu P_\mu)\, D(g_\Lambda) = 1 + i(\Lambda^{-1})^\mu{}_\nu \epsilon^\nu P_\mu$$

$$\begin{cases} D(g_\Lambda)^{-1} P_\mu D(g_\Lambda) = \Lambda_\mu{}^\nu P_\nu \\ \boxed{D(g_\Lambda)^{-1} P^\mu D(g_\Lambda) = \Lambda^\mu{}_\nu P^\nu} \end{cases}$$

Expanding for small $\omega_{\alpha\beta} \ll 1$

$$(\mathbb{1} + \frac{1}{2}\omega_{\alpha\beta} J^{\alpha\beta}) P^\mu (\mathbb{1} - \frac{1}{2}\omega_{\alpha\beta} J^{\alpha\beta}) = (\delta^\mu_\nu + \omega^\mu{}_\nu) P^\nu$$

So

$$\frac{i}{2}\omega_{\alpha\beta}\left[J^{\alpha\beta}, P^{\mu}\right] = \omega^{\mu}_{\ \nu} P^{\nu}$$

$$\boxed{\left[J^{\alpha\beta}, P^{\mu}\right] = i\left(\eta^{\mu\alpha}P^{\beta} - \eta^{\mu\beta} - \eta^{\mu\beta}P^{\alpha}\right)}$$

The full set of commutators for the Poincaré algebra, where we define

$$K_i \equiv J^{i0}, \qquad J_i \equiv \frac{1}{2}\epsilon_{ijk}J^{jk}, \qquad H \equiv P_0, \qquad P_i$$

is as follows

$$\begin{array}{ll}
[J_i, J_k] = i\epsilon_{ijk}J_k & [J_i, K_j] = i\epsilon_{ijk}K_k \\
[J_i, P_j] = i\epsilon_{ijk}P_k & [J_i, H] = 0 \rightarrow \text{good quantum number} \\
[K_i, K_j] = -i\epsilon_{ijk}J_k & [K_i, P_j] = i\delta_{ij}H \\
[K_i, H] = iP_i & [P_j, P_j] = 0 \\
[P_i, H] = 0 & [H, H] = 0
\end{array}$$

5.4.1 Casimirs of the Poincaré algebra

P^{μ} form a subalgebra. Because $[P^{\mu}, P^{\nu}] = 0$ they can be diagonalised simultaneously so that

$$P^{\mu}\left|P\right\rangle = p^{\mu}\left|P\right\rangle$$

Where p^{μ} is an eigenvalue of the operator P^{μ}.
P^2 is a scalar and

$$\left[P^2, P^{\mu}\right] = 0, \qquad \left[P^2, J^{\mu\nu}\right] = 0$$

$$P^2\left|P\right\rangle = M^2\left|P\right\rangle$$

With M the mass.

- P^2 is a Casimir for the Poincaré group
- Any relativistic particle is identified by M
- M identifies the irreducible representation of Poincaré
- Spin is not a good relativistic quantum number (Spin ¿ ¡ Boost). So we use helicity = projection of spin along the direction of motion.

We introduce the Pauli-Lubanski (pseudo-)vector

$$\boxed{W^{\mu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}J_{\nu\rho}P_{\sigma}}$$

$W^2 \equiv W^{\mu}W_{\mu}$ is a Casimir of the Poincaré group

$$\left[W^2, P^{\mu}\right] = 0, \qquad \left[W^2, J^{\alpha\beta}\right] = 0$$

$$W^2\left|P, \sigma\right\rangle \propto \frac{\sigma(\sigma+1)}{2}\left|P, \sigma\right\rangle$$

To classify any irreducible representation of the Poincaré group: P^2 and W^2 give mass and spin.
See exercises!!

Part X

General relativity

TODO: watch out for $c = 1$.

Chapter 1

Gravity and covariance

In the previous section we have discussed relativistic mechanics and rewritten the laws of electrodynamics in a covariant way. Now we turn to gravity; the goal of general relativity is to find a relativistic theory of gravity.
Electrodynamics was already intrinsically relativistic. Unfortunately we are not so lucky with Newton's theory of gravity. It has some decidedly non-relativistic properties that mean that it must be an approximation of a more fundamental theory.
We can express Newton's theory of gravity as follows (see also the section on Newtonian mechanics):

$$\mathbf{F}_{12} = -G_N \frac{m_1 m_2}{|\mathbf{r}_1 - \mathbf{r}_2|^2} \hat{n}_{12} \qquad \left(\hat{n}_{12} \text{ is the unit vector } \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right)$$

Problems for making it covariant include:

- Gravity depends on a spatial distance measured at a given time, both of which are not invariant. According to which observer should we measure these quantities?
- There is an action at a distance: Newtonian gravity acts instantly. However no signal can move faster than the speed of light. This is a problem.

Of course it is not always necessary to use relativistic gravity. Newtonian gravity is adequate in many situations. In particular the Newtonian approximation is a good one if

$$\frac{G_N M}{R c^2}$$

is much less than one. Otherwise we must use the formulae for relativistic gravity.

Chapter 2

Equivalence principles

2.1 Weak equivalence principle

$$\mathbf{F}_g = -m_g \boldsymbol{\nabla}\Phi \qquad \mathbf{a} = \frac{\mathbf{F}_g}{m_I} = -\left(\frac{m_g}{m_I}\right)\boldsymbol{\nabla}\Phi$$

The weak equivalence principle states that

$$m_g = m_I$$

Which means that the acceleration of an object due to gravity is independent of the nature of the object. In other words gravity is purely geometric.
Equivalently we can state the WEP as follows

> The motion of freely-falling particles are the same in a gravitational field and a uniformly accelerated frame, in small enough regions of space.

2.2 Einstein equivalence principle

The basic idea is that locally there is no way to distinguish between uniform acceleration and an external gravitational field, no matter what the experiment (so it applies to all of physics). Formally

> In small enough regions of spacetime, the laws of physics reduce to those of spacial relativity; it is impossible to detect the existence of a gravitational field by means of local experiments.

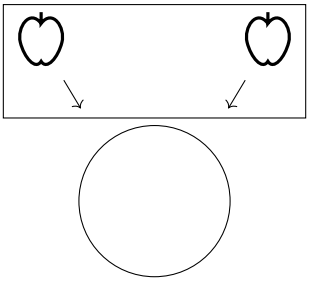
1. The weak equivalence principle holds
2. There is local Lorentz invariance (the space is locally Minkowski and special relativity holds)

It is impossible to detect gravity through local, non-gravitational experiments

3. There is local position invariance.

$$\partial_\mu g_{\nu\rho}(x_p) = 0$$

Local Lorentz invariance has been verified to 10^{-24} and local position invariance to 10^{-4}.
We can tell acceleration from gravity if time / space is long enough:



In general iff no nonuniformities of Φ can be detected.

2.3 Strong equivalence principle

Object only feels result of curvature from other objects.

self-interaction strongly limited

> General relativity is uniquely compatible with the strong equivalence principle.

2.4 Experimental evidence

2.4.1 Eötvös experiment

[Insert picture here] Eötvös parameter:

$$\eta = 2\frac{a_2 - a_1}{a_2 + a_1} = \frac{\left(\frac{m_g}{m_I}\right)_2 - \left(\frac{m_g}{m_I}\right)_1}{\left(\frac{m_g}{m_I}\right)_2 + \left(\frac{m_g}{m_I}\right)_1}$$

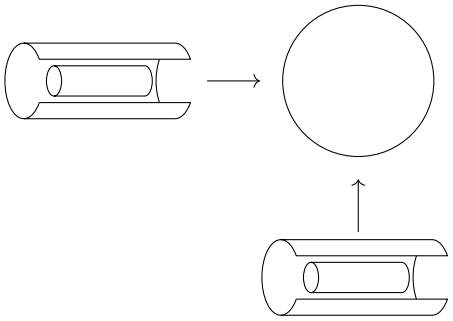
With the following results:

$$\begin{aligned} \eta_{\text{Be,T}} &= (0,3 \pm 1,8) \cdot 10^{-13} \\ \eta_{\text{Be,Al}} &= (-0,7 \pm 1,3) \cdot 10^{-13} \\ \eta_{\text{Earth,Moon}} &= (\quad \pm \quad) \cdot 10 \end{aligned}$$

2.4.2 MICROSCOPE

MICRO-Satellite à traînée Compensée pour l'Observation du Principe d'Equivalence (MICROSCOPE). Tests WEP up to 10^{-15}.

$$\eta < 2,1 \times 10^{-14}$$



2.5 Towards general covariance

2.5.1 Newtonian gravity in spacetime terms

6.6 Hartle

$$\mathrm{d}s^2 = -\left(1 + \frac{2\Phi(x^i)}{c^2}\right)(c\,\mathrm{d}t)^2 + \left(1 - \frac{2\Phi(x^i)}{c^2}\right)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)$$

Static weak field approximation

Chapter 3

Consequences of the equivalence principles and overview of general relativity

In this section we explore some of the consequences of the equivalence principles. Then we give a very quick overview of way we will be building up the theory of general relativity in this section. This should hopefully serve to justify the somewhat involved mathematics of the next section.

3.1 Gravitational redshift

$$\omega_\infty = \left(1 - \frac{G_N M}{Rc^2}\right) \omega_*$$

3.2 Geometric interpretation of gravity

spacetime: differentiable manifold Chapt 6.5 Hartle
Gravity as a property of space, so we need to review how we see space. We need it to be approximately flat in small areas. This idea is captured in the rigourous, mathematical notion of a manifold.
We take spacetime to be a differentiable manifold, meaning we can still do calculus to points in the manifold.
We also still want to be able to use vectors, and in general tensors, on the manifold. Here we run into a slight problem: vectors in different points are incompatible with each other. Concepts such as the displacement and position vector only work in flat space. Instead we construct a vector space (and more generally tensors) in each point. That vector space is called the tangent space in a point. (TODO Why + link to curvature)
We can however link the tangent spaces in neighbouring points using connections. These are not inherent in a manifold and specify the curvature. There is another (more intuitive) way to specify the curvature: using a metric. Here our concept of metric will be slightly different (!) than when we discussed metric spaces above. A major difference will be that we will not require it to be positive definite.

3.3 Causality

3.4 Rindler spacetime

3.4.1 Uniformly accelerated observers

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2$$

We study uniformly accelerating object (indestiguishable from gravity in local experiments):

$$\alpha^2 = \alpha^\mu \alpha_\mu = a^2 \qquad x(0) = \frac{1}{a}$$

With α the acceleration. Following is basic SR kinematics

$$x(t) = x(0) + \frac{1}{a}\sqrt{1 + a^2 t^2} - \frac{1}{a} = \frac{1}{a}\sqrt{1 + a^2 t^2}$$

$$\begin{cases} t(\tau) = \frac{1}{a}\sinh(a\tau) \\ x(\tau) = \frac{1}{a}\sqrt{1 + \sinh^2(a\tau)} \end{cases} \qquad x^\mu(\tau) = \frac{1}{a}\begin{pmatrix}\sinh(a\tau)\\ \cosh(a\tau)\end{pmatrix}$$

[Note to self: look up Minkowski diagram]
Still compatible with assumptions:

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \begin{pmatrix}\cosh(a\tau)\\ \sinh(a\tau)\end{pmatrix} \qquad \alpha^\mu \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = a\begin{pmatrix}\sinh(a\tau)\\ \cosh(a\tau)\end{pmatrix}$$

$$\alpha^\mu \alpha_\mu = a^2(-\sinh^2(a\tau) + \cosh^2(a\tau)) = a^2$$

3.4.2 Adapted coordinates

Change variables so that $\xi = 0$ gives trajectory of object.

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto \begin{pmatrix} \tau \\ \xi \end{pmatrix}$$

A t = 0 both observers have same spacetime
Change of coordinates is singular: can only describe part of spacetime.

$$\begin{cases} t(\tau, \xi = 0) = \frac{1}{a}\sinh(a\tau) \\ x(\tau, \xi = 0) = \frac{1}{a}\cosh(a\tau) \end{cases}$$

$$x(\tau = 0, \xi) = \frac{f(\xi)}{a} ??$$

We want $x = 0$ is $\xi = -\infty$, so log, because $x = 0$ will be singular:

$$x \in]0, +\infty[, t \in \mathbb{R} \to \xi \in \mathbb{R}, \tau \in \mathbb{R}$$

So we choose:

$$\begin{cases} t(\tau, \xi) = \frac{e^{a\xi}}{a}\sinh(a\tau) \\ x(\tau, \xi) = \frac{e^{a\xi}}{a}\cosh(a\tau) \end{cases}$$

$$\begin{cases} \tau(t,x) = \frac{1}{a}\operatorname{arctanh}(t/x) \\ \xi(t,x) = \frac{1}{2a}\log[a^2(x^2-t^2)] \end{cases}$$

Argument of log has to be positive, so only describes wedge, not even light cone.

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2$$

$$\begin{aligned} \mathrm{d}s^2 &= -\left(\frac{\partial t}{\partial \xi}\,\mathrm{d}\xi + \frac{\partial t}{\partial \tau}\,\mathrm{d}\tau\right)^2 + \left(\frac{\partial x}{\partial \xi}\,\mathrm{d}\xi + \frac{\partial x}{\partial \tau}\,\mathrm{d}\tau\right)^2 \\ &= -\left(e^{a\xi}\sinh(a\tau)\,\mathrm{d}\xi + e^{a\xi}\cosh(a\tau)\,\mathrm{d}\tau\right)^2 + \left(e^{a\xi}\cosh(a\tau)\,\mathrm{d}\xi + e^{a\xi}\sinh(a\tau)\,\mathrm{d}\tau\right)^2 = \quad e^{2a\xi} \end{aligned}$$

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2 = e^{2a\xi}(-\,\mathrm{d}\tau^2 + \mathrm{d}\xi^2)$$

3.4.3 Rindler spacetime (time dilation, event horizons)

Now take observer at $\xi = 0$, so $t = \frac{1}{a}\sinh(a\tau)$ and $x = \frac{1}{a}\cosh(a\tau)$. Then observe $\xi = 1$:

$$\begin{cases} t = \frac{e^a}{a}\sinh(a\tau) = \frac{1}{a'}\sinh(a'\tau') & a' = ae^{-a} < a \\ x = \frac{e^a}{a}\cosh(a\tau) = \frac{1}{a'}\cosh(a'\tau') & \tau' = \frac{a}{a'}\tau = e^a\tau > \tau \end{cases}$$

Less acceleration means time flows faster, more acceleration means time flows slower.
Now for $\xi = -1$:

$$\begin{cases} t = \frac{e^{-a}}{a}\sinh(a\tau) = \frac{1}{a''}\sinh(a''\tau'') & a'' = ae^{a} > a \\ x = \frac{e^{-a}}{a}\cosh(a\tau) = \frac{1}{a''}\cosh(a''\tau'') & \tau'' = \frac{a}{a''}\tau = e^{-a}\tau < \tau \end{cases}$$

Clock in satellite: two competing effects.
Redshift: $\sqrt{-g_{\tau\tau}} = \sqrt{e^{2a\xi}} = e^{a\xi}$
Even if metric is singular, the space is not singular! (Extend I to II, and then to IV)

$$g_{\mu\nu} = \begin{pmatrix} -e^{2a\xi} & 0 \\ 0 & e^{2a\xi} \end{pmatrix} \qquad \partial_\tau g_{\mu\nu} = 0$$

$$u^\mu = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad u = \partial_\tau$$

$$\begin{aligned} u^\mu u_\mu = g_{\tau\tau} &= -e^{2a\xi} < 0 \\ &= a^2(t^2 - x^2) \end{aligned}$$

u is timelike in I as it should be, but can also be lightlike or spacelike in different sections.

Chapter 4

Explaining the geometry of gravity: Einstein's equations

4.1 The Einstein-Hilbert action

4.2 Einstein's equations

4.3 The stress-energy tensor

(stress tensor for a gas of particles)

4.4 Couplings to other fields

$$x^\mu : \mathbb{R} \to \mathbb{R}^{1,3} : \tau \mapsto$$

Now we describe free motion of massive particle. The action is:

$$S[x^\mu(\tau)] = -m \int \mathrm{d}s$$

Describe for arbitrary metric

$$\mathrm{d}s^2 = \mathrm{d}x^\mu \, \mathrm{d}x^\nu g_{\mu\nu}(x)$$

Extremise S! Trajectory of massive particle given by maximum. (Massless more difficult).

$$\delta S = -m \int \delta \left(\sqrt{-\dot{x}^\mu \dot{x}^\nu g_{\mu\nu}(x)} \right) \mathrm{d}\tau = m \int \frac{1}{2\sqrt{-\dot{x}^2}} \left(2\dot{x}^\mu \tfrac{\mathrm{d}\delta x^\nu}{\mathrm{d}\tau} g_{\mu\nu} + \dot{x}^\mu \dot{x}^\nu \frac{\partial g_{\mu\nu}}{\partial x^\rho} \delta x^\rho \right) \mathrm{d}\tau$$

Using that $\dot{x}^2$ is constant, this is equal to

$$m \int \frac{1}{2\sqrt{-\dot{x}^2}} \left(\ddot{x}^\mu g_{\mu\nu} \delta x^\nu - \dot{x}^\mu \delta x^\nu \frac{\partial g_{\mu\nu}}{\partial x^\rho} \dot{x}^\rho + \frac{1}{2} \dot{x}^\mu \dot{x}^\nu \frac{\partial g_{\mu\nu}}{\partial x^\rho} \delta x^\rho \right) \mathrm{d}\tau = 0$$

$$\begin{aligned} \delta S &= -m \int \frac{1}{2\sqrt{-\dot{x}^2}} \left(\ddot{x}^\mu g_{\mu\rho} + \dot{x}^\mu \dot{x}^\nu \partial_\nu g_{\mu\rho} - \frac{1}{2} \dot{x}^\mu \dot{x}^\nu \partial_\rho g_{\mu\rho} \right) \delta x^\rho \, \mathrm{d}\tau \\ &= -m \int \frac{g_{\mu\rho}}{2\sqrt{-\dot{x}^2}} \left(\ddot{x}^\mu + g^{\mu\sigma} \partial_\tau g_{\mu\rho} \dot{x}^\nu \dot{x}^\tau - \frac{1}{2} g^{\mu\sigma} \partial_\sigma g_{\mu\tau} \dot{x}^\nu \dot{x}^\tau \right) \delta x^\rho \, \mathrm{d}\tau \end{aligned}$$

$$\delta S = -m \int \frac{g_{\mu\rho}}{2\sqrt{-\dot{x}^2}} \frac{D\dot{x}^\mu}{D\tau} \delta x^\rho \, \mathrm{d}\tau$$

$$\frac{D\dot{x}^\mu}{D\tau} = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \frac{1}{2} g^{\mu\sigma} \left(\partial_\nu g_{\rho\sigma} + \partial_\rho g_{\nu\sigma} - \partial_\sigma g_{\nu\rho}\right) \dot{x}^\nu \dot{x}^\rho = 0$$

Christoffel symbols

$$\Gamma^\mu_{\nu\rho} \equiv \frac{1}{2} g^{\mu\sigma} \left(\partial_\nu g_{\rho\sigma} + \partial_\rho g_{\nu\sigma} - \partial_\sigma g_{\nu\rho}\right)$$

What about partical with other parameter than τ? Or massless particles? We want a procedure that is reparametrisation invariant. Change of coordinates that is never singular:

$$\tau \to \tau' = \tau'(\tau) \qquad \frac{\mathrm{d}\tau'}{\mathrm{d}\tau} \neq 0 \qquad \mathrm{d}s^2 = \mathrm{d}x^\mu \, \mathrm{d}x^\nu g_{\mu\nu}(x)$$

$$\mathrm{d}s^2 = \mathrm{d}\tau^2 \gamma_{\tau\tau}(\tau) = \mathrm{d}\tau' \, \mathrm{d}\tau' \gamma_{\tau'\tau'} = \mathrm{d}\tau' \, \mathrm{d}\tau' \frac{\mathrm{d}\tau}{\mathrm{d}\tau'} \frac{\mathrm{d}\tau}{\mathrm{d}\tau'} \gamma_{\tau\tau}$$

$$\gamma_{\tau'\tau'} = \left(\tfrac{\mathrm{d}\tau}{\mathrm{d}\tau'}\right)^2 \gamma_{\tau\tau}$$

$$\gamma_{\tau\tau} = \left(\tfrac{\mathrm{d}\tau}{\mathrm{d}\tau'}\right)^2 \gamma_{\tau'\tau'}$$

Proper time: $\gamma_{\tau\tau} = -1$ Polyakov's action:

$$S = -\frac{1}{2} \mathrm{d}\tau \sqrt{\gamma} \left[\gamma^{\tau\tau} \tfrac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \tfrac{\mathrm{d}x^\nu}{\mathrm{d}\tau} g_{\mu\nu}(x) + m^2\right]$$

(τ generic parameter)
We can also describe motion of particle, string or brane:

- 0-dim object $\to$ 1-dim worldline
- 1-dim object $\to$ 2-dim worldsheet

Write action in terms of e ($\gamma_{\tau\tau}(\tau) = -e^2(\tau)$)

$$S[x^\mu, e] = \frac{1}{2} \int \mathrm{d}\tau [e^{-1} \dot{x}^\mu \dot{x}^\nu g_{\mu\nu} - e m^2]$$

$$\begin{aligned} \frac{\delta S}{\partial x^\mu} &\sim e^{-1} [2\dot{x}^\mu \delta\dot{x}^\nu g_{\mu\nu} + \dot{x}^\mu \dot{x}^\nu \partial_\rho g_{\mu\nu} \delta x^\rho] \\ &= \int \mathrm{d}\tau e^{-1} \left[-\frac{D\dot{x}^\mu}{D\tau} g_{\mu\nu} \delta x^\nu\right] = 0 \end{aligned}$$

For massless particle: simply put $m = 0$

$$\frac{\partial S}{\partial e} \sim -e^{-2}\dot{x}^2 - m^2 = 0$$

$$m = 0 \textbf{ particle: } \dot{x}^2 = 0$$

$$\textbf{for } m \neq 0\textbf{: } e^2 = \tfrac{m^2}{-\dot{x}^2} \qquad e = \tfrac{m}{\sqrt{-\dot{x}^2}}$$

Equation of motion (?):

$$\frac{D\dot{x}^\mu}{\tau} = 0 = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0$$

Using the following

$$\sigma = f(\tau): \qquad \frac{\mathrm{d}}{\mathrm{d}\tau} = \frac{\mathrm{d}f}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}\sigma} \qquad \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} = \ddot{f}\frac{\mathrm{d}}{\mathrm{d}\sigma} + (\dot{f}^2)\frac{\mathrm{d}^0}{\mathrm{d}\sigma^0}$$

we get

$$\ddot{f}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma} + \dot{f}^2\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\sigma^2} + \Gamma^\mu_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\rho}{\mathrm{d}\sigma}(\dot{f})^2 = 0$$

Lecture 18/10
Polyakov's action ($\mathrm{d}s^2|_{\text{on woldline}} = -e(\tau)^2\,\mathrm{d}\tau$)

$$S = \frac{1}{2}\int \mathrm{d}\tau(e^{-1}\dot{x}^2 - em^2)$$

$$\delta_e S = -\frac{1}{2}\int \mathrm{d}\tau(+e^{-2}\dot{x}^2 + m^2)\delta e$$

$$\begin{aligned}\delta_x S &= \frac{1}{2}\int \mathrm{d}\tau(2e^{-1}\dot{x}^\mu g_{\mu\nu}\tfrac{\mathrm{d}\delta x}{\mathrm{d}\tau} + e^{-1}\dot{x}^\mu\dot{x}^\rho\partial_\nu g_{\mu\rho}\delta x^\nu)\\ &= \int \mathrm{d}\tau e^{-1}(\frac{D\dot{x}^\mu}{\tau}g_{\mu\nu}\delta x + e^{-2}\tfrac{\mathrm{d}e}{\mathrm{d}\tau}\dot{x}^\mu g_{\mu\rho}\delta x^\nu)\end{aligned}$$

Where

$$\frac{D\dot{x}^\mu}{\tau} \equiv= \frac{\mathrm{d}\dot{x}}{\mathrm{d}\tau} + \Gamma^\mu_{\nu\rho}\dot{x}^\nu\dot{x}^\rho$$

with $\Gamma^\mu_{\nu\rho} = \frac{1}{2}$. We use the following notation:

$$\Big\{(\mu\nu) = \tfrac{1}{2}(\mu\nu + F_\nu\mu)[\mu\nu] = \tfrac{1}{2}(\mu\nu - F_\nu\mu)$$

For example with EM tensor F: $F_{[\mu\nu]} = \frac{1}{2}(F_{\mu\nu - F_\nu\mu}) = 0$ (F antisymm).
Now this:

$$g_{00} = -(1 + 2\frac{V(\mathbf{x})}{m}) \qquad g_{ij} = \delta_{ij}$$

Using $\frac{V}{m}$

$$\begin{aligned}S = -m\int\sqrt{(1+\frac{2V}{m}) - |\tfrac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}|}\tfrac{\mathrm{d}t}{\mathrm{d}\tau}\,\mathrm{d}\tau \\ &\approx -m\int(1 + \frac{V}{m} - \frac{1}{2}|\mathbf{x}|^2)\,\mathrm{d}t\\ &= -m\int\mathrm{d}t + \int\left[\frac{1}{2}m|\mathbf{x}|^2 - V(\mathbf{x})\right]\mathrm{d}t\end{aligned}$$

In worldline (1D?) ($\gamma : \tau \mapsto x^\mu(\tau)$)

$$S_1 = q\int_\gamma \mathrm{d}x^\mu A_\mu(x)$$

$$\delta A_\mu = \partial_\mu = \partial_\mu \Lambda \overset{\text{gauge invariance}}{\to} \delta_\Lambda S_1 = q \int_\gamma \mathrm{d}x^\mu \partial_\mu \Lambda$$

$$\begin{aligned}
\delta_x S_1 &= q \int_\gamma \mathrm{d}\delta x^\mu A_\mu + q \int_\gamma \mathrm{d}x^\mu \partial_\nu A_\mu \delta x^\nu \\
&= q \int_\gamma \left[\tfrac{\mathrm{d}\delta x^\mu}{\mathrm{d}\tau} A_\mu + \tfrac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \partial_\nu A_\mu \delta x^\nu \right] \mathrm{d}t \\
&= q \left[-\delta x^\mu \partial_\nu A_\mu \dot{x}^\nu + \dot{x}^\nu \partial_\nu A_\mu \delta x^\nu \right] \mathrm{d}t \\
&= -q \int F_{\mu\nu} \dot{x}^\mu \, \mathrm{d}x^\nu \, \mathrm{d}\tau
\end{aligned}$$

$$\frac{D\dot{x}^\mu}{\tau} = \frac{q}{m} F^{\mu\nu}$$

Chapter 5

Symmetries and simplifications

Chapter 6

Some physics of gravitation

6.1 Weak field approximation

6.2 Gravitational redshift (GPS)

Weak gravity

$$\mathrm{d}s^2 = -(1+2\phi)\,\mathrm{d}t^2 + \delta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j$$

Earth: $\phi = -G_N\frac{M_\oplus}{r}$ [picture with satellites] Two competing effects: satellites going faster but also in weaker gravity.

$$\frac{\nu_{s\to e}}{\nu_s} = \frac{\nu_{s\to e}}{\nu_e} = \sqrt{\frac{g_{\mu\nu}(r_s)\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}}{g_{00}(r_e)}} = \sqrt{\frac{1+2\phi(r_s)-v_s^2}{1+2\phi(r_e)}}$$

Using $v_s^2 = G\frac{M_\oplus}{R_\oplus+h}$ and $\phi(r_s = R_\oplus + h) = -\frac{GM_\oplus}{R_\oplus+h}$, we get

$$= \sqrt{\frac{1-3G\frac{M_\oplus}{R_\oplus+h}}{1-2G\frac{M_\oplus}{R_\oplus}}} \approx 1 - \frac{3}{2}G\frac{M_\oplus}{R_\oplus+h} + G\frac{M_\oplus}{R_\oplus}$$

$$= ????missed$$

Sometimes redshift, sometimes blueshift. (Missed calculation of when)

6.3 Gravity waves

Small, time dependent perturbations

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

Where $|h| \sim 10^{-21}$?? Not produced in big quantities (we do not see effects in cosmic microwave background radiation) Indirect detection: period of pulsars (like PSR 1913+16). They collapse because they lose (radiate) energy as gravitational waves. Also direct observation.
CFR. EM

$$\begin{cases}\mathrm{d}F = 0\\ \mathrm{d}{*}F = 0\end{cases} \qquad \to \qquad \Box A^\mu - \partial^\mu\partial_\nu A^\nu = 0$$

$$\begin{cases} \Box A^\mu = 0 \\ \partial_\mu A^\mu = 0 \qquad \text{gauge} \end{cases}$$

$$A^\mu = c^\mu e^{-ik_\mu x^\mu}$$

$$k^2 = 0 \qquad k_\mu c^\mu = 0 \qquad A^\mu = ak^\mu + \tilde{c}^\mu$$

$$k^\mu = |k| \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad \tilde{c}^\mu = \begin{pmatrix} 0 \\ \tilde{c}^1 \\ \tilde{c}^2 \\ 0 \end{pmatrix}$$

We want to do the same with $R_{\mu\nu} = 0$. In order to do that we need to linearize

$$R_{\mu\nu}(h) = 0$$

We raise and lower indices with η because rest is of order h^2.

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}$$

$$h^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}$$

$$\Gamma^\mu_{\nu\rho}(h) = \frac{1}{2}\eta^{\mu\sigma}\left(\partial_\nu h_{\rho\sigma} + \partial_\rho h_{\nu\sigma} - \partial_\sigma h_{\rho\nu}\right)$$

$$\begin{aligned} R_{\mu\nu}(h) &= R_{\alpha\mu\ \nu}^{\ \ \alpha}(h) = \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\mu \Gamma^\alpha_{\alpha\nu} \\ &= \frac{1}{2}\left[\eta^{\alpha\sigma}\partial_\alpha\left(\partial_\mu h_{\nu\sigma} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}\right) - \eta^{\alpha\sigma}\partial_\mu\left(\partial_\alpha h_{\nu\sigma} + \partial_\nu h_{\alpha\sigma} - \partial_\sigma h_{\alpha\nu}\right)\right] \end{aligned}$$

$$R_{\mu\nu}(h) = \frac{1}{2}\left[\partial^\mu\partial^\sigma h_{\nu\sigma} + \partial_\nu\partial^\sigma h_{\mu\sigma} - \Box_\eta h_{\mu\nu} - \partial_\mu\partial^\sigma h_{\nu\sigma} - \partial_\mu\partial_\nu h + \partial_\mu\partial^\sigma h_{\nu\sigma}\right] = 0$$

$$\boxed{\Box h_{\mu\nu} + \partial_\mu\partial\nu h - \partial_\nu\partial^\sigma h_{\nu\sigma} - \partial_\nu\partial^\sigma h_{\mu\sigma} = 0}$$

Now we need to fix invariance under diffeomorphism. How many functions an we use? In EM we can fix one.
Missed bunch (h is trace)
deDonder gauge

$$\boxed{\Gamma^\rho_{\mu\nu} g^{\mu\nu} = 0}$$

Again missed a bunch

$$\boxed{\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h = 0} \qquad \text{gauge}$$

So we have: From Einstein ($R_{\mu\nu}(h) = 0$) and the deDonder gauge ($\Gamma^\rho_{\mu\nu} g^{\mu\nu} = 0$) we get

$$\begin{cases} \Box h_{\mu\nu} = 0 \\ \partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h = 0 \end{cases}$$

$\Box h_{\mu\nu} = 0$ implies $\Box h = 0$. To put it in a better form we call

$$\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$$

for the trace we have

$$\bar{h} = h - \frac{1}{2}4h = -h$$

We then get

$$\begin{cases} \Box \bar{h}_{\mu\nu} = 0 \\ \partial^\mu \bar{h}_{\mu\nu} = 0 \end{cases}$$

$$c_{\mu\nu} = \begin{pmatrix} c_{00} & c_{01} & c_{02} & -c_{00} \\ c_{01} & c_{11} & c_{12} & -c_{10} \\ c_{02} & c_{12} & c_{22} & -c_{20} \\ -c_{10} & -c_{10} & -c_{20} & +c_{00} \end{pmatrix}$$

$$\bar{h}_{\mu\nu} = c_{\mu\nu} e^{-ik_\rho x^\rho}, \qquad k^2 = 0, \qquad k^\mu c_{\mu\nu}$$

$$k^\mu = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad c_{0\nu} + c_{3\nu} = 0$$

$$\begin{aligned} \delta h_{\mu\nu} &= \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu \\ \delta h_{\mu\nu} &= \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu - \partial^\rho \epsilon_\rho \eta_{\mu\nu} \\ \delta c_{\mu\nu} &= k_\mu \epsilon_\nu + k_\nu \epsilon_\mu - k^\rho \epsilon_\rho \eta_{\mu\nu} \end{aligned}$$

so

$$\begin{aligned} \delta c_{00} &= -2\epsilon_0 + (\epsilon_0 + \epsilon_3) = \epsilon_3 - \epsilon_0 \\ \delta c_{01} &= -\epsilon_1 \\ \delta c_{02} &= -\epsilon_2 \\ \delta c_{12} &= 0 \\ \delta c_{11} &= -(\epsilon_0 + \epsilon_3) \\ \delta c_{22} &= -(\epsilon_0 + \epsilon_3) \end{aligned}$$

We finally get the transverse traceless polarization vector

$$c_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & c_{11} & c_{12} & 0 \\ 0 & c_{12} & -c_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}z^2 + \mathrm{d}x^2(1 + c_{11}\cos(t+z)) + \mathrm{d}y^2(1 - c_{11}\cos(t-z))$$

$$c_{11}, c_{12} \sim 10^{-21}$$

Missed several lectures??
Lecture 22/11 (Approx size of c_{11}??)

6.3.1 Gravity waves: detection

$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}z^2 + \mathrm{d}x^2(1 + c_{11}\cos\alpha or \kappa(t-z)) + \mathrm{d}y^2(1 + c_{11}\cos\alpha or\kappa(t-z)) + 2\,\mathrm{d}x\,\mathrm{d}y c_{12}\cos\alpha or\kappa(t-z)$

Geodesic detection equation

$$\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} = -\delta^i \partial_i \phi(x(t)) \qquad x^i(t) = x^i_*(t) + \chi^i(t)$$

$$\frac{\mathrm{d}^2 x^i_*(t)}{\mathrm{d}t^2} + \frac{\mathrm{d}^2 \chi^i(t)}{\mathrm{d}t^2} = -\delta^{ij}\partial_j \phi(x^*(t)) - \delta^{ij}\partial_\kappa \partial_j \phi(x^*(t))\chi^*(t) + \mathcal{O}(\chi^2)$$

So

$$\frac{\mathrm{d}^2 \chi^i(t)}{\mathrm{d}t^2} \approx -\delta^{ij} \underbrace{\partial_\kappa \partial_j \phi(x^*(t))}_{\text{tidal tensor}} \chi^*(t)$$

How does difference between $x^\mu(\tau_x)$ and $y^\mu(\tau_y)$ change with time? (We set $\tau = \tau_x = \tau_y$). $\epsilon^\mu(\tau_x, \tau_y) = x^\mu(\tau_x) - y^\mu(\tau_y)$
Simplified calculation with special reference frame, but with covariant result. (Caroll does it fully covariantly, much longer) Choose coordinates (Fermi normal coordinates) such that

$$\begin{aligned}
&g_{\mu\nu}(x_*) = \eta_{\mu\nu} \\
&\partial_\rho g_{\mu\nu}(x_*) = 0 = \Gamma^\mu_{\nu\rho}(x^*) \\
&g_{\mu\nu}(x_*) = \eta_{\mu\nu} + \mathcal{O}(|x - x^*|^2) \\
&\partial_\rho g_{\mu\nu}(x(t)) = 0 = \Gamma^\mu_{\nu\rho}(x(t))
\end{aligned}$$

is satisfied along geodesic! (Is possible, proved by Fermi)

$$\ddot{y}^\mu + \Gamma^\mu_{\nu\rho}(y(\tau))\dot{y}^\nu \dot{y}^\rho = 0$$

We compute ϵ

$$\begin{aligned}
\ddot{\epsilon}^\mu &= \ddot{y}^\mu - \ddot{x}^\mu \\
&= -\Gamma^\mu_{\nu\rho}(y)\dot{y}^\nu \dot{y}^\rho \\
&= ???????
\end{aligned}$$

$$y^\mu(\tau) = x^\mu(\tau) + \epsilon^\mu(\tau)$$

$$\ddot{\epsilon}^\mu = -\epsilon^\sigma \partial_\sigma \Gamma^\mu_{\nu\rho}(x)\dot{x}^\nu \dot{x}^\rho \qquad \text{first order in } \epsilon$$

$$\begin{aligned}
\frac{D^2\epsilon^\mu}{D\tau} = \frac{D}{D\tau}\left[\frac{D\epsilon^\mu}{D\tau}\right] &= \frac{D}{D\tau}\left(\dot{\epsilon}^\mu + \Gamma^\mu_{\nu\rho}(x)\epsilon^\nu \dot{x}^\rho\right) \\
&= \ddot{\epsilon}^\mu + \tfrac{\mathrm{d}}{\mathrm{d}\tau}\left(\Gamma^\mu_{\nu\rho}(x)\epsilon^\nu \dot{x}^\rho\right) + \cancel{\Gamma^\mu_{\nu\rho}(x)\frac{D\epsilon^\nu}{D\tau}\dot{x}^\rho} \\
&= \ddot{\epsilon}^\mu + \dot{x}^\sigma \partial_\sigma \Gamma^\mu_{\nu\rho}(x)\epsilon^\nu \dot{x}^\rho \\
&= -\epsilon^\sigma \partial_\sigma \Gamma^\mu_{\nu\rho}(x)^\cdot{}^\nu \dot{x}^\rho + \dot{x}^\sigma \partial_\sigma \Gamma^\mu_{\nu\rho}\epsilon^\nu \dot{x}^\rho \\
&= \epsilon^\sigma \dot{x}^\nu \dot{x}^\rho \left(-\partial_\sigma \Gamma^\mu_{\nu\rho}(x) + \partial_\nu \Gamma^\mu_{\sigma\rho}(x)\right) = -R_{\sigma\nu\ \rho}^{\ \mu}\epsilon^\sigma \dot{x}^\nu \dot{x}^\rho
\end{aligned}$$

$$\boxed{\frac{D^2\epsilon^\mu}{D\tau^2} = -R_{\sigma\nu\ \rho}^{\ \mu}\epsilon^\sigma \dot{x}^\nu \dot{x}^\rho}$$

Now we use this metric

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}z^2 + \mathrm{d}x^2(1 + c_{11}\cos\kappa(t-z)) + \mathrm{d}y^2(1 + c_{11}\cos\kappa(t-z)) + 2\,\mathrm{d}x\,\mathrm{d}y c_{12}\cos\kappa(t-z)$$

geodesic doesn't change ?!? But proper distance changes.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \qquad h_{00} = h_{0i} = h_{zz} = h_{z0} = h_{zi} = 0$$
$$h_{ij} \qquad i,j = x,y$$
$$\begin{cases} \dot{x}^0 = 1 = \dot{y}^0 \\ \dot{x}^{i,z} = 0 = \dot{y}^{i,z} \end{cases} \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
$$\frac{\mathrm{d}^2\epsilon^\mu}{\mathrm{d}t^2} = -R_{\sigma 0\ 0}^{\ \mu}\epsilon^\sigma$$

$$\begin{aligned} \tfrac{\mathrm{d}^2\epsilon^i}{\mathrm{d}t^2} &= -R_{j0\ 0}^{\ i}\epsilon^j \\ &= -\left(\partial_j\Gamma^i_{00} - \partial_0\Gamma^i_{j0}\right)\epsilon^j \end{aligned}$$

$$\frac{\mathrm{d}^2\epsilon^i}{\mathrm{d}t^2} = +\partial_0\Gamma^i_{j0}\epsilon^j = \frac{1}{2}\delta^{ik}\left(\tfrac{\mathrm{d}^2}{\mathrm{d}t^2}h_{kj}\right)\epsilon^j$$
$$\begin{cases} \ddot{\epsilon}^x = -\frac{1}{2}c_{11}\cos(\kappa t)\epsilon^x \\ \ddot{\epsilon}^y = +\frac{1}{2}c_{11}\cos(\kappa t)\epsilon^y \end{cases}$$
$$\begin{cases} \epsilon^x(t) = \left(1 + \frac{1}{2}c_{11}\cos(\kappa t)\right)\epsilon^x(0) + \mathcal{O}(c_1 1^2) \\ \epsilon^y(t) = \left(1 - \frac{1}{2}c_{11}\cos(\kappa t)\right)\epsilon^y(0) + \mathcal{O}(c_1 1^2) \end{cases}$$

If quantizable, gravity has spin 2. A theorem states that particles of spin more than 2 are completely decoupled. What about 3/2? Squishing and stretching of circle [picture]

6.4 Physical effects of gravity waves (geodesic deviation equation)

Chapter 7

Gravity outside a spherical mass: Schwarzschild's solution

7.1 Solving Einstein's equations using symmetries

Solve Einstein equtions. Solve for $R_{\mu\nu} = 0$.
We assume spherical symmetry, so we have 3 Killing vectors.

$$[K_I, K_S] = \epsilon_{IJK} K_K \qquad I, J = 1, 2, 3$$

$$K_I = K_I^\mu \partial_\mu \qquad D_{[\mu} K_{I\nu]} = 0$$

We also assume static spacetime ($\neq$ stationary). This assumption is not necessary to reach Schwarzschild, but helps.

$$\mathrm{d}s^2 = g_{00}\,\mathrm{d}t^2 + g_{0i}\,\mathrm{d}t\,\mathrm{d}x^i + g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j$$

Stationary means you can foliate space layers [fig].

$$\exists \xi_t = \xi_t^\mu \partial_\mu, \qquad \xi_t^2 < 0 \qquad D_{[\mu}\xi_{t\nu]} = 0$$

$$\Leftrightarrow \xi_t^\mu \partial_\mu g_{\nu\rho} + g_{\mu\nu}\partial_\rho \xi_t^\mu + g_{\mu\rho}\partial_\nu \xi_t^\mu = 0$$

Introduce t: $\xi_t = \partial_t$
Static means something more: there is no mixing of time and space coordinates(?)

$$K_I^\mu \partial_\mu g_{\nu\rho} + g_{\mu\nu}\partial_\rho K_I^\mu + g_{\mu\rho}\partial_\nu K_I^\nu = 0$$

$$K_I^i \partial_i g_{00} + \cancel{2 g_{i0} \partial_0 K_I^i} = 0 \qquad \Rightarrow \qquad g_{00}(x) = g_{00}(r)$$

In spherical coordinates:

$$f(r)\,\mathrm{d}r^2 + g(r)\left[\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right]$$

Ansatz for the metric:

$$\mathrm{d}s^2 = -A^2(r)\,\mathrm{d}t^2 + B^2(r)\,\mathrm{d}r^2 + C(r)^2\left[\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right]$$

10 equations. "Parity invariances":

$$\begin{cases} t \to -t & R_{tr} = 0 = R_{\theta\phi} \\ \phi \to -\phi & R_{t\theta} = 0 = R_{r\theta} \\ \theta \to -\theta & R_{t\phi} = 0 = R_{r\phi} \end{cases}$$

So only 4 equations left.
We will use fieldbines (easier if we choose them intelligently). (Fieldbines are only defined up to a Lorentz transformation)

$$g_{00} = -A^2 \quad g_{rr} = B^2 \quad g_{\theta\theta} = r^2 \quad g_{\phi\phi} = r^2 \sin^2\theta$$

$e^a \to \omega \to \mathbb{R}$

$$\mathrm{d}s^2 = e^a \otimes e^b \eta_{ab} = -e^{0\,2} + (e^1)^2 + (e^2)^2 + (e^3)^2$$

$$\begin{cases} e^0 = A(r)\,\mathrm{d}t \\ e^1 = B(r)\,\mathrm{d}r \\ e^2 = r\,\mathrm{d}\theta \\ e^3 = r\sin\theta\,\mathrm{d}\phi \end{cases}$$

$$\mathrm{d}e^a + \omega^a_{\ b} e^b = 0 \qquad R^a_{\ b} = \mathrm{d}\omega^a_{\ b} + \omega^a_{\ c}\omega^c_{\ d}$$

Missed stuff (in lecture notes)
Lecture 06/12 Ward identity for Compton scattering

$$\mathcal{M} = -ie^2 \bar{u}' \epsilon'^*_\nu \epsilon_\mu \left(\gamma^\nu \frac{\not{p}}{} \right)$$

[Fig see notes] Dirac equation.
Rewrite

$$\epsilon_\mu \not{p} \gamma^\mu u = \not{p} eps\not{i} lon u = (-)???$$

Which gives us

$$\mathcal{M} = -ie^2 \bar{u}' \left(\frac{\not{\epsilon}' \not{k} \not{\epsilon} + 2(p\cdot\epsilon)\not{\epsilon}'^*}{2(p\cdot k)} + \frac{-\not{\epsilon}\not{k}\not{\epsilon}'^* + 2(p'\cdot\epsilon)\not{\epsilon}'^*}{-2(p'\cdot k)} \right) u(p)$$

On-shell condition $\not{k}\not{k} = k^2 = 0$

7.2 Scattering form a ("semi" classical) external field

No solutions satisfying simultaneously on-shellness and momentum conservation (there exist no real Minkowski moments)
Coulomb field Semi-classicla external field.

$$\mathcal{M} = ie\bar{u}' A_{\mathrm{ext}}(\bar{q}) u = ie\bar{u}'\gamma^\mu u A_{\mathrm{ext}\mu}$$

Flux

$$\phi = \frac{v_{\mathrm{rel}}}{V} = \frac{|\bar{p}|}{VE}$$

Unpolarised x-section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega'} = \frac{(2m\alpha Z)^2}{|\bar{q}|^4}\frac{1}{2}\sum_{\text{spin}}|\bar{u}\gamma^0 u|^2$$

$$|\bar{\mathcal{M}}|^2 = \frac{1}{2}\operatorname{Tr}\left((\not{p}' + m)\gamma^0(\not{p} + m)\gamma_0\right)$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{(\alpha Z)^2}{4E^2v^4\sin^4\frac{\theta}{2}}\left(1 - v^2\sin^2\frac{\theta}{2}\right)$$

If $Z = 1$ then hydrogen so proton charge.
Diagrams $M \to \infty, \bar{p}_{\text{proton}} = \bar{0}$
J_ν is proton 4-current.

$$\bar{u}_r(\bar{0})\gamma^0 u_s(\bar{0}) = \delta_{rs}$$

$$\delta\mathcal{M} = 0 \Leftrightarrow \sum_{i\in\text{in}} Q_i - \sum_{j\in\text{out}} Q_j = 0$$

Lecture 07/12 Missed first half: Calculation of Schwarzschild

$$\begin{aligned} BR_{tt} + AR_{rr} &= \frac{2}{r}\frac{A'}{B^2}B + \frac{2}{r}\frac{B'}{B^2}A \\ &= \frac{2}{rB^2}\left(A'B + B'A\right) = 0 \end{aligned}$$

$$B = A^{-1} \qquad B' = (A^{-1})' = -A^{-2}A'$$

$$-\frac{A'}{AB^2} + \frac{B'}{B^3} + \left(1 - \frac{1}{B^2}\right)\frac{1}{r} = -AA' - AA' + (1 - A^2)\frac{1}{r} = 0$$

$$\mathrm{d}s^2 = -\left(1 - \frac{2m}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2m}{r}\right)^{-1}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)$$

$\mathbb{R} \times \mathrm{SO}(3)$ isometry group. 4 Killing vectors.

7.3 Finding Schwarzschild

Chapter 8

Physics of the solar system

8.1 Effective potential for geodesics

8.2 Geodesics for massive particles and Mercury's procession

8.3 Geodesics for massless particles

8.3.1 Bending of light rays

8.3.2 Light emitted from compact stars

8.3.3 Shapiro time delay

Schwarschild radius

$$r^* = 2m$$

m is mass of object (compare to Kumar energy $\mathcal{E}$??)

8.4 Geodesics

$$\mathcal{L} = g_{\mu\nu}(x)\dot{x}^\mu\dot{x}^\nu = \epsilon$$

Where $\epsilon = \begin{cases} 0 & (m=0) \\ -1 & (m\neq 0) \end{cases}$ We fill in the metric to get:

$$\mathcal{L} = -\left(1-\frac{2M}{r}\right)\dot{t}^2 + \left(1-\frac{2M}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2$$

$$Q = K_\mu\dot{x}^\mu$$

$$\mathcal{E} = -\xi_\mu\dot{x}^\mu = -\xi^\mu g_{\mu\nu}\dot{x}^\nu = -g_{t\nu}\dot{x}^\nu = -g_{tt}\dot{t} = \left(1-\frac{2M}{r}\right)\dot{t}$$

$$\begin{cases} L_1 = \sin\phi\partial_\theta + \frac{\cos\phi}{\tan\theta}\partial_\phi \\ L_2 = \cos\phi\partial_\theta - \frac{\sin\phi}{\tan\theta}\partial_\phi \\ L_3 = \partial_\phi \end{cases}$$

$$\mathbf{l} = \mathbf{L}_\mu \dot{x}^\mu \qquad \mathbf{l} = \begin{pmatrix} 0 \\ 0 \\ l \end{pmatrix}$$

$$\dot{l}(??) = \frac{\mathcal{E}}{1 - \frac{2M}{r(t)}} \qquad \overset{r\to\infty}{\to} \qquad \mathcal{E} = \gamma$$

Lecture 13/12 Missed first half ?? Other stuff(?) ???

$$m \equiv G_N \frac{M}{c^2}$$

$$\mathrm{d}s^2 = -\left(1 - \frac{2m}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2m}{r}\right)^{-1} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)$$

$$\dot{t} = \frac{E}{1 - \frac{2m}{r}} \qquad \theta = \frac{\pi}{2} \qquad \dot{\phi} = \frac{l}{r^2}$$

Extreme $\epsilon = 0 \qquad r = 3m$

$$\epsilon = -1 \qquad r_\pm = \frac{l^2}{2m}\left(1 \pm \sqrt{1 - r(\frac{m}{l})^2??}\right)$$

$$\mathcal{L}_{\text{eff}} = \frac{\dot{r}^2}{2} + V_{\text{eff}}(r) = E_{\text{eff}} = \frac{E^2 + \epsilon}{2}$$

$$V_{\text{eff}} = \epsilon\frac{m}{r} + \frac{l^2}{2r^2} - \underbrace{m\frac{l^2}{r^3}}_{\text{Relativistic correction??}}$$

Massive geodesics $\epsilon = -1$

$$\frac{\mathrm{d}V_{\text{eff}}}{\mathrm{d}r} = -\epsilon\frac{m}{r^2} - \frac{l}{r^3} + 3m\frac{l^2}{r^4} = 0$$

second half
Escape velocity

$$l = 0 \qquad E = 1 \qquad \frac{\dot{r}}{2} - \frac{m}{r} = 0$$

$$\dot{r} = \sqrt{\frac{2m}{r}} \qquad \overset{r\to r_s = 2m}{\to} \qquad 1$$

With r_s the Schwarzschild radius. (1 means speed of light -¿ no escape)
ISCO (Inner most stable circular orbit?)
Circular orbits $l \geq \sqrt{12}m$

$$r_+ = \frac{12m^2}{2m} = 6m$$

$$\dot{r} = ????$$

(Energy gain massive (6%): several times fusion process)??
Kepler's law

$$\omega = \frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\dot{\phi}}{\dot{t}} = \frac{l}{r_+^2 E}\left(1 - \frac{2m}{r_+}\right)$$

$r = r_+$

$$\frac{m}{r^2} - \frac{l^2}{r^2}\left(1 - 3\frac{??}{??}??\right)???$$

??????
Perihelion procession

$$\Delta\phi = 2\int_{r_-}^{r_+} \frac{\mathrm{d}\phi}{\mathrm{d}r}\,\mathrm{d}r$$

$$\left(\tfrac{\mathrm{d}\phi}{\mathrm{d}r}{}^2\right) = \frac{\dot{\phi}^2}{\dot{r}^2} = \frac{l^2}{r^4}\frac{1}{2(E_{\text{eff}} - V_{\text{eff}})}$$

$$u = \frac{1}{r} \qquad \mathrm{d}u = -\frac{1}{r^2}\,\mathrm{d}r$$

$$(\left(\tfrac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2)^2 = r'^2 = \frac{2r^4}{l^2}\,(E_{\text{eff}} - V_{\text{eff}})$$

$$u'^2 = \frac{r'^2}{r^4} = \frac{2}{l^2}\,(E_{\text{eff}} - V_{\text{eff}}(u)) = \frac{2}{l^2}E_{\text{eff}} + \frac{2m}{l^2}u - u^2 + 2mu^3(?)$$

$$2u'u" = \left(\frac{2m}{l^2} - 2u + 6mu^2\right)u'$$

$$u" = \frac{m}{l^2} - u + \underbrace{3mu^2}_{\text{GR correction}}$$

$$u = u_N + u_E \qquad u_N" = \frac{m}{l^2} - u_N \qquad u_E" = -u_E + 3mu_N^2$$

Chapter 9

Gravity in the solar system

Chapter 10

Schwarzschild black hole

10.1 Accelerated and freely falling observers

10.2 Collapse to black hole

10.3 Near-horizon metric

10.4 Eddington-Finkelstein and Kruskal-Szekeres coordinates

10.5 Global extensions

$$\mathrm{d}s^2 = -\left(1-\frac{2m}{r}\right)\mathrm{d}t^2 + \left(1-\frac{2m}{r}\right)^{-1}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2+\sin^2\theta\,\mathrm{d}\phi^2\right)$$

$r = 2m$ coordinate singularity.
Black holes useful to understand gravity.
Take observer:

$$u^0 = \left(1-\frac{2m}{r}\right)^{-1/2} \qquad \mathbf{u} = 0$$

This is not an inertial observer:

$$\begin{aligned}\alpha^\mu &= \frac{Du^\mu}{D\tau} = u^\nu D_\nu u^\mu \\ &= \cancel{\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}} + \Gamma^\mu_{\nu\rho}u^\nu u^\rho = \Gamma^\mu_{tt}\left(1-\frac{2m}{r}\right)^{-1}\end{aligned}$$

$$\begin{aligned}\Gamma^\mu_{tt} &= \frac{1}{2}g^{\mu\rho}\left(\cancel{2\partial_t g_{t\rho}} - \partial_\rho g_{tt}\right) = -\frac{1}{2}g^{\mu\rho}\partial_\rho g_{tt} \\ &= \frac{1}{2}g^{\mu\rho}\partial_\rho\left(1-\frac{2m}{r}\right)\end{aligned}$$

$$\alpha^r = \Gamma^r_{tt}\left(1-\frac{2m}{r}\right)^{-1} = \cancel{\left(1-\frac{2m}{r}\right)^{-1}}\frac{1}{2}\cancel{\left(1-\frac{2m}{r}\right)}\partial_r\left(-2\frac{m}{r}\right) = \frac{m}{r^2}$$

$$\alpha^\mu = \begin{pmatrix} 0 \\ \frac{m}{r^2} \\ 0 \\ 0 \end{pmatrix}$$

The acceleration you need stand still is $\sqrt{\alpha^2}$ (coordinate invariant)!

$$\alpha^2 = \alpha^\mu\alpha^\nu g_{\mu\nu} = (\alpha^r)^2 g_{rr} = \frac{m^2}{r^4}(1-\frac{2m}{r})^{-1}(??????)$$

At $r = m$ this is infinite!! So impossible to stand still
Time to get $r = r_s$ for obs at ∞

$$T = \int_{r_0}^{r_s}\frac{\mathrm{d}t}{\mathrm{d}r}\,\mathrm{d}r \overset{\rho=r-r_s}{=}\sim \int_{r_0+r_s}^{0}\frac{\mathrm{d}\rho}{\rho} \to \infty$$

$$\frac{\mathrm{d}t}{\mathrm{d}r} = \frac{\dot{t}}{\dot{r}} = \frac{E}{\left(1-\frac{2m}{r}\right)}\left(E^2-1+\frac{2m}{r}\right)^{-1/2}$$

[Fig 1]
Proper time

$$\mathcal{T} = \int_{r_0}^{r_s=2m}\frac{\mathrm{d}\tau}{\mathrm{d}r}\,\mathrm{d}r$$

$$\tfrac{\mathrm{d}\tau}{\mathrm{d}r}\big|_{??} = ??$$

???

$$\mathcal{T} = -\frac{1}{\sqrt{2m}}\int_{r_0}^{r_s}\sqrt{\frac{rr_0}{r_0-r}}\,\mathrm{d}r$$

We change coordinates $r = \frac{r_0}{2}(1+\cos\eta)$ and $\mathrm{d}r = -\frac{r_0}{2}\sin\eta\,\mathrm{d}\eta$

$$\begin{aligned}\mathcal{T} &= +\frac{1}{\sqrt{2m}}\int_0^{\eta_f}\sqrt{\frac{r_0^2(1+\cos\eta)}{r_0(1-\cos\eta)}}\frac{r_0}{2}\sin\eta\,\mathrm{d}\eta \\ &= \frac{r_0^{3/2}}{2\sqrt{2m}}\int_0^{eta_f}(1+\cos\eta)\,\mathrm{d}\eta = \sqrt{\frac{r_0^3}{8m}}(\eta_f+\sin\eta_f)\end{aligned}$$

$$\frac{\mathrm{d}t}{\mathrm{d}r} = \frac{E}{\left(1-\frac{2m}{r}\right)}\left(E^2-1+\frac{2m}{r}\right)^{-1/2}$$

[Fig 2]

10.6 Near-horizon metric

Change coordinates:

$$\mathrm{d}s^2 = -\left(\frac{\rho}{2m+\rho}\right)\mathrm{d}t^2 + \left(\frac{2m+\rho}{\rho}\right)\mathrm{d}\rho^2 + (2m+\rho)^2\underbrace{\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)}_{\mathrm{d}\Omega^2_{S^2}}$$

Take $\rho \ll 2m$:

$$\mathrm{d}s^2 = -\frac{\rho}{2m}\,\mathrm{d}t^2 + 2m\frac{\mathrm{d}\rho^2}{\rho} + 4m^2\,\mathrm{d}\Omega^2_{S^2}$$

Using $\rho = \frac{e}{denom}$??? We get Rindler!
Extending Rindler [Fig 3]

$$\mathrm{d}s^2 = -r^2\,\mathrm{d}t^2 + \mathrm{d}r^2$$

Light ray $\mathrm{d}r = \pm r\,\mathrm{d}t$ [Fig 4]

$$r = e^{\rho}$$
$$\mathrm{d}s^2 = e^{2\rho}\left(-\mathrm{d}t^2 + \mathrm{d}\rho^2\right)$$

Light ray $\mathrm{d}\rho = \pm\,\mathrm{d}t$

$$\begin{cases}\text{In falling} & \mathrm{d}\rho = -\,\mathrm{d}t \quad \to \quad v = t + \rho = t + \log r \\ \text{Outgoing} & \mathrm{d}\rho = \mathrm{d}t \quad \to \quad u = t - \rho = t - \log r\end{cases}$$

Coordinate change

$$\mathrm{d}s^2 = e^{v-u}(-\,\mathrm{d}u\,\mathrm{d}v)$$

Explanation of how to extend spacetime coordinates.
Radial light rays

$$\left(1 - \frac{2m}{r}\right)\mathrm{d}t = \pm\,\mathrm{d}r$$
$$\pm\,\mathrm{d}t = \frac{\mathrm{d}r}{1-\frac{2m}{r}} = \mathrm{d}\tilde{r}$$
$$\tilde{r} = r + 2m\log\frac{r-2m}{2m}$$

Regge-Wheeler coordinates

$$\mathrm{d}s^2 = \left(1 - \frac{2m}{r(\tilde{r})}\right)\left(-\,\mathrm{d}t^2 + \mathrm{d}\tilde{r}^2\right) + r(\tilde{r})\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)$$

$$\begin{cases} u = t - \tilde{r} \\ v = t + \tilde{r}\end{cases}$$

$$\begin{cases}\text{Ingoing} & \mathrm{d}s^2 = -\left(1-\frac{2m}{r}\right)\mathrm{d}v^2 + 2\,\mathrm{d}v\,\mathrm{d}r + r^2\,\mathrm{d}{\Omega_{S^2}}^2 \\ \text{Outgoing} & \mathrm{d}s^2 = -\left(1-\frac{2m}{r}\right)\mathrm{d}u^2 - 2\,\mathrm{d}u\,\mathrm{d}r + r^2\,\mathrm{d}{\Omega_{S^2}}^2\end{cases}$$

Eddington- Finkelstein

$$g|_{r=2m} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

Chapter 11

Black hole physics

11.1 Time translation in Kruskal coordinates

11.2 Null hypersurfaces

11.3 Surface gravity

11.4 Penrose diagrams

Nima Arkam-Hamed (???)
Don't modify gravity - understand it !

$$\mathrm{d}s^2 = \underbrace{1 - \frac{2m}{r}}_{A}\,\mathrm{d}u + 2\,\mathrm{d}u\,\mathrm{d}r r^2\,\mathrm{d}{\Omega_{S^2}}^2??$$

Horizon

$$S = r - 2m = 0 \qquad \mathcal{N} = \{x | S(x) = 0\}$$

$$lperp\mathcal{N}, \quad lparallel\mathcal{N} \qquad l^\mu = f(x) g^{\mu\nu}\partial_\nu S$$

$$g = \begin{pmatrix} -A & 1 \\ 1 & 0 \end{pmatrix}$$

missed blackboard

$$\xi^\nu D_\nu \xi^\mu = \kappa \xi^\mu$$

Where κ is surface gravity (constant!!!, not function)

$$\kappa = -\frac{1}{2} D_\mu \xi_\nu D^\mu \xi^\nu|_N$$

is constant on N
$\xi^2 = -1$ at $r = \infty$

$$\xi^u D_u \xi^u = \cancel{\xi^u \partial_u \xi^u} + \xi^u \Gamma^u_{u\rho} \xi^\rho = \Gamma^u_{uu} = \frac{1}{2} g^{u\rho} \left(\cancel{2\partial_u g_{u\rho}} - \partial_\rho g_{(?)} \right)$$

$$= -\frac{1}{2} g^{ur} \partial_r \frac{2m}{r} = \frac{m}{r^2}|_{r=2m} = \boxed{\frac{1}{4m} = \kappa}$$

Surface gravity: gravity on surface of black hole as seen by an observer far away

11.5 Relation between kappa and T (TODO: math mode)

Temperature related to periodicity in Euclidean time

$$Z = \prod_i \int \mathrm{d}p_i \, \mathrm{d}q_i e^{-\beta E(p,q)} \qquad \beta = \frac{1}{T}$$

$$\mathcal{L} = \partial_\mu \varphi \partial^\mu \varphi + V(\varphi)$$

$$Z = tze^{-\beta H} = \int D\varphi e^{-\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3 x \mathcal{L}(\varphi)}$$

$$\mathrm{d}s^2_{\mathrm{Schw}} = \left(1 - \frac{2m}{r}\right) \mathrm{d}\eta^2 + \left(1 - \frac{2m}{r}\right) \mathrm{d}r^2 + r^2 \, \mathrm{d}{\Omega_{S^2}}^2$$

$r - 2m = \frac{\rho^2}{8m}$ $t = n\eta$(???)
$\rho \ll 1$ $r \sim 2m$

$$\approx \underbrace{\frac{\rho^2}{16m^2} \mathrm{d}\eta^2}_{\rho^2 \kappa^2 \, \mathrm{d}\eta^2 + \mathrm{d}\rho^2} + \mathrm{d}\rho^2 + 4m^2 \, \mathrm{d}{\Omega_{S^2}}^2 + corrections$$

$$T_{BH} = ... = ...$$

11.6 Penrose diagrams

Interested only in causal structure.
Conformal transformation

$$(M, g) \mapsto (\tilde{M}, \tilde{g}) \qquad \tilde{g}_{\mu\nu} = \omega^2(x) g_{\mu\nu}$$

preserves inner product (angles)

$$\frac{A^\mu g_{\mu\nu} B^\nu}{\sqrt{A_g^2 B_g^2}} = \frac{A^\mu \tilde{g}_{\mu\nu} B^\nu}{\sqrt{A_{\tilde{g}}^2 B_{\tilde{g}}^2}}$$

2d spacetimes are conformally flat

$${\mathrm{d}s_{2d}}^2 = e^{2\sigma(x)(?)} \left(-\mathrm{d}H^2 + \mathrm{d}x^2\right)$$

$$\mathcal{R}_{0101} \neq 0 \qquad \omega = e^\sigma$$

$$\tilde{R} = R(\omega^2 g) = e^{-2\sigma} \left(R - 2g^{\mu\nu} D_\mu D_\nu \sigma\right) = 0$$

[picture of lightcone: future/past lightcone $\mathcal{I}^\pm$, future past timelike past $i^\pm$, spacelike seperation]

1. use 4d coords that allow reduction to 2d

2. choose new 2d coords with finite range
3. conf transf to rid of factors due to point 2
4. coord change to minkowski
5. represent the result in a diagram

11.6.1 Minkowski spacetime $\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$

1. $\mathrm{d}s^2 = \underbrace{-\mathrm{d}t^2 + \mathrm{d}r^2}_{\mathcal{I} \qquad -\mathrm{d}u\,\mathrm{d}v} + r^2 \underbrace{\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right)}_{\mathrm{d}{\Omega_{S^2}}^2}$ [fig 1]
2. $u = \tan\tilde{u}$
   $v = \tan\tilde{v}$
   $$\tilde{u}, \tilde{v} \in]-\pi/2, \pi/2[\qquad \tilde{u} \leq \tilde{v}$$
   $\mathrm{d}s^2 = -\frac{1}{\cos^2\tilde{u}\cos^2\tilde{v}}\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v} + \left(\frac{\tan\tilde{v}\tan\tilde{u}}{2}\right)^2 \mathrm{d}{\Omega_{S^2}}^2$
3. $g \to \omega^2 g$
4.

pic 2

Kruskal

1.
2. $\mathrm{d}s^2 = -32\frac{m^3}{r}e^{-r/2m(?)}\,\mathrm{d}U\,\mathrm{d}V + r^2(U,V)\,\mathrm{d}\Omega^2_{S^2}$
3. $U = \tan\tilde{U}$
   $V = \tan\tilde{V}$
   $$\tilde{U}, \tilde{V} \in]\pi/2, \pi/2[$$
4. $\omega = 2\cos\tilde{U}\cos\tilde{V}$
   $$\mathrm{d}s^2 = -\frac{128}{r}m^3 e^{-r/2m}\,\mathrm{d}\tilde{U}\,\mathrm{d}\tilde{V} + r^2\cos^2\tilde{U}\cos^2\tilde{V}\,\mathrm{d}\Omega^2$$
5. $\tilde{U} = \tilde{T} - \tilde{X} \qquad \tilde{V} = \tilde{T} + \tilde{X}$

Chapter 12

Other black holes

12.1 Charged black holes

$$M_p = 1$$

$$\frac{\mathcal{L}}{\sqrt{-g}} = \frac{\mathcal{R}}{2} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

$$\mathcal{R}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = T_{\mu\nu}$$

Dynamic BHs

$$F = \frac{1}{2} \, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu F_{\mu\nu}$$

$$\mathrm{d}F = *J_m$$

$$\mathrm{d}*F = *J_e$$

$$\mathrm{d}s^2 = -f^2(r) \, \mathrm{d}t^2 + f^{-2}(r) \, \mathrm{d}r^2 + r^2 \, \mathrm{d}{\Omega_{S^2}}^2$$

$$q = \frac{1}{4\pi} \int_{S^2} *F \qquad p = \frac{1}{4\pi} \int_{S^2} F$$

$$E^r = F^{0r} = \frac{q}{r^2} \qquad F = \frac{p}{r^2} \left(r^2 \cos\theta (\wedge ?) \, \mathrm{d}\theta \, \mathrm{d}\phi \right) + \frac{q}{r^2} \, \mathrm{d}r (\wedge ?) \, \mathrm{d}t$$

$$B^r = \frac{P}{r^2}$$

$$f(r) = \frac{\Delta}{r^2} \qquad \Delta = r^2 - 2mr + p^2 + q^2$$

$$\begin{array}{c|c|l} \Delta = 0 & \text{horizons} & r_\pm = m \pm \sqrt{m^2 - (p^2 + q^2)} \\ r = 0 & \text{singularity} & \end{array}$$

[picture numerline: $r = 0 | \Delta > 0 | r_- | \Delta < 0 | r_+ | \Delta > 0$]

1. no horizons $m < p^2 + q^2$, naked singularity (problematic!?)
2. regular BHs $m > \sqrt{p^2 + q^2}$, 2 horizons
3. external BH $m = \sqrt{p^2 + q^2}$ one double horizon $r_+ = r_- = m$

$$\mathrm{d}s^2 = -\frac{\Delta}{r^2}\,\mathrm{d}v^2 + 2\,\mathrm{d}v\,\mathrm{d}r + r^2\,\mathrm{d}\Omega_{S^2}{}^2$$

For $r \to r_+, r \to r_-$: $\mathrm{d}s^2 \to \text{Rindler} \times S^2$
[pic 1]

$$A_{\text{hor}} = 4\pi r_\pm^2 \overset{\text{extremal}}{=} 4\pi(q^2 + p^2)$$

$$\Delta = (r-m)^2$$

$$\Delta l_{r_0 r=m} = \int_m^{r_0} \frac{r}{\Delta}\,\mathrm{d}r = \int_m^{r_0} \left(1 + \frac{m}{r-m}\right)\mathrm{d}r = \infty$$

$r = m + \rho$ $\Delta = \rho^2$

$$\mathrm{d}s^2 \qquad \overset{\rho \ll 1}{\to} \qquad \underbrace{-\frac{\rho^2}{m^2}\,\mathrm{d}t^2 + m^2\frac{\mathrm{d}\rho^2}{\rho^2}}_{A\,\mathrm{d}S_2} \underbrace{+}_{\times} \underbrace{m^2\,\mathrm{d}\Omega_{S^2}{}^2}_{S^2}$$

(Bertotti - Robinson) (long neck)
Now possible to combine solutions ??

$$\begin{aligned}
\mathrm{d}a^2 &= -\left(1-\frac{m}{r}\right)^2\mathrm{d}t^2 + \left(1-\frac{m}{r}\right)^{-2}\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega_{S^2}{}^2 \\
&= -\left(1-\frac{m}{r}\right)^2\mathrm{d}t^2 + \left(1-\frac{m}{r}\right)^{-2}\left[\mathrm{d}r^2 + (r-m)^2\,\mathrm{d}\Omega_{S^2}{}^2\right] \\
&= -\left(1+\frac{m}{\tilde{r}}\right)^2\mathrm{d}t^2 + \left(1+\frac{m}{\tilde{r}}\right)^{-2}\underbrace{\left[\mathrm{d}\tilde{r}^2 + \tilde{r}^2\,\mathrm{d}\Omega_{S^2}{}^2\right]}_{\mathrm{d}\mathbf{x}\cdot\mathrm{d}\mathbf{x}} \\
&= -\left(1+\frac{m}{\tilde{r}}\right)^2\mathrm{d}t^2 + \left(1+\frac{m}{\tilde{r}}\right)^{-2}\mathrm{d}\mathbf{x}^2
\end{aligned}$$

Where $\tilde{r} = r - m$

$$\Delta_3 H = \delta^3(r)$$

$$H = 1 + \Sigma_i \frac{m_i}{\sqrt{(\mathbf{x} - \mathbf{x}_i ??)}}$$

We can put extremal black holes wherever we want.
Schwinger process

$$M \geq Q \qquad m \leq q$$

$$M - m \geq Q - q$$

You need $m \leq q$ for elementary particles if you don't want naked singularities, so gravity has to be the weakest force.

12.2 Rotating black holes

Kerr-Newman metric (rotating and charged)

$$\mathrm{d}s^2 = -\frac{\Delta - a^2\sin^2\theta}{\Sigma}\,\mathrm{d}t^2 + \frac{(r^2+a^2)^2 - \Delta a^2\sin^2\theta}{\Sigma}\sin^2\theta\,\mathrm{d}\phi^2 - 2a\sin^2\theta\,\mathrm{d}t\,\mathrm{d}\phi\frac{(r^2+a^2)-\Delta}{\Sigma} + \frac{\Sigma}{\Delta}\,\mathrm{d}r^2 + \Sigma\,\mathrm{d}\theta^2$$

With

$$\Delta = r^2 - 2mr + a^2 + e^2 \qquad \Sigma = r^2 + a^2\cos^2\theta \qquad a = \frac{J}{m} \qquad e = \sqrt{p^2+q^2}$$

J really is angular momentum (Kummar charge ...)
We want asymptotically Minkowski.
Spherical symmetry too strong We have timelike killing vector

$$K_t = \partial_t \qquad K_t^2|_\infty = -1$$

Axial symmetry $K_\phi = \partial_\phi$
We will consider Kerr metric (charge e=0)
No hair theorem: one you fix charges black hole is uniquely defined.

- $\Sigma = 0$ is a real singularity; $r = 0, \theta = \frac{\pi}{2}$ is a ring
- horizons $\Delta = 0$ $r_\pm = m \pm \sqrt{m^2 - a^2}$
  extremal $J = m^2 \Leftrightarrow a = m$

(Impossible to reach extremality)
Horizons are not the horizons where K_t vanishes, but where K goes to zero:

$$K = \partial_t + \Omega_H\partial_\phi \qquad \Omega_H = \frac{a}{r_\pm^2 + a^2}$$

$$\mathcal{K}_\pm = \frac{r_\pm - r_\mp}{2(r_\pm^2 + a^2)}$$

$$\xi_\phi K_\phi^\mu g_{\mu\nu}\,\mathrm{d}x^\nu = \ldots\,\mathrm{d}\phi + \ldots\,\mathrm{d}t$$

(Ker-Newman not metric outside rotating star: disregards stress-energy tensor)

$$l = K_\phi^\mu g_{\mu\nu}\dot{x}^\nu = g_{\phi\phi}\dot{\phi} + g_{\phi t}\dot{t}$$

$$l = 0 \qquad \frac{\mathrm{d}\phi}{\mathrm{d}t} = -\frac{g_{\phi t}}{g_{\phi\phi}}$$

12.3 Frame dragging

Framedragging! Spacetime is rotating and drags everything along with it.
Now we will view metric with r very large. Obviously Minkowski, but first corrections?

$$\mathrm{d}s^2_{r\to\infty} = -\underbrace{\frac{r^2-2mr}{r^2}}_{(1-\frac{2m}{r})}\mathrm{d}t^2 + \left(1-\frac{2m}{r}\right)^{-1}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right) - 2a\sin^2\theta\frac{2mr}{r^2}\,\mathrm{d}t\,\mathrm{d}\phi$$

So

$$\mathrm{d}s^2_{r\to\infty} = \mathrm{d}{s_{schw}}^2 - 4\frac{J}{r}\sin^2\theta\,\mathrm{d}t\,\mathrm{d}\phi$$

If we transform

$$\begin{cases} z = r\cos\theta \\ x = r\sin\theta\cos\phi \\ y = r\sin\theta\cos\sin\phi \end{cases}$$

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + 4\frac{J}{r^3}\left(x\,\mathrm{d}y - y\,\mathrm{d}x\right)\mathrm{d}t$$

$$k^\mu = (u^0, 0, 0, u^2) \qquad \text{motion at} \quad x = y = 0$$
$$S^\mu = (0, S^x, S^y, 0) \qquad S^\mu U_\mu = 0$$
$$U^\nu D)\nu S^\mu = 0$$
$$\dot{S}^\mu + \Gamma^\mu_{\nu\rho} u^\nu S^\rho = 0$$
$$\Gamma^x_{xt}|_{x=y=0} = 0$$

$$\begin{aligned} \Gamma^x_{yt}\big|_{x=y=0} &= \frac{1}{2}g^{xx}\left(\partial_y g_{tx} + \partial_t g_{yx} - \partial_x g_{yt}\right) \\ &= \frac{1}{2}(-4J)\left[\partial_y\left(\frac{y}{r^3} + \partial_x\left(\frac{x}{r^3}\right)\right)\right] \\ &= -4\frac{J}{z^3} \end{aligned}$$

$$\begin{cases} \dot{S}^x - 4\frac{J}{z^3}u^0 S^y = 0 \\ \dot{S}^y + 4\frac{J}{z^3}u^0 S^x = 0 \end{cases}$$

$$(S^x + iS^y) + i\left(4\frac{J}{z^3}u^0\right)(S^x + iS^y) = 0$$

This is Lense-Thiving effect.

12.4 Ergosphere & Penrose process

$$K_t^2 = 0 \qquad r^2 - 2mr^2 + a^2(1 - \sin^2\theta) = 0$$
$$r_e(\theta) = m + \sqrt{m^2 - a^2\cos^2\theta}$$
$$\Delta = 0 \qquad r_+ = m + \sqrt{m^2 - a^2}$$

[Fig 1]

$$r = \text{const} \qquad \theta = \frac{\pi}{2} \qquad \text{lightlike motion}$$

$$\mathrm{d}s^2 = -\frac{\Delta - a^2}{r^2}\,\mathrm{d}t^2 + \frac{(r^2 + a^2)^2 - \Delta a^2}{r^2}\,\mathrm{d}\phi^2$$
$$- 2a\frac{r^2 + a^2 - \Delta}{r^2}\,\mathrm{d}t\,\mathrm{d}\phi = 0$$

- $r \to \infty$

$$-\mathrm{d}t^2 + r^2\,\mathrm{d}\phi^2 = 0 \Leftrightarrow \frac{\mathrm{d}\phi}{\mathrm{d}t} = \pm\frac{1}{r}$$

- $r = r_e = 2m \quad \Delta = a^2$

$$\frac{\left(4m^2 + a^2\right)^2 - a^4}{4m^2}\,\mathrm{d}\phi^2 - 2a\,\mathrm{d}\phi\,\mathrm{d}t = 0$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = 0 \qquad \frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{a}{2m^2 + a^2}$$

- Farther inside: not even able to stand still

$$\epsilon = K_t^\mu g_{\mu\nu}\dot{x}^\nu = \epsilon_1\epsilon_2$$

$\epsilon - \epsilon_1 = \epsilon_2 > \epsilon$ if $\epsilon_0 < 0$

$$\epsilon - \Omega_H l = -K_t^\mu P_\mu - \Omega_H K_\phi^\mu P_\mu \geq 0$$

$$l \leq \frac{\epsilon}{\Omega_H}, \qquad \delta l \leq \frac{\delta\epsilon}{\Omega_H}$$

$$\delta J \leq \frac{\delta m}{\Omega_H} = \frac{r_t^2 + a^2}{a}\delta m = \frac{\left(m + \sqrt{m^2 - \frac{J^2}{m^2}}\right)^2 + \frac{J^2}{m^2}}{\frac{J}{m}}\delta m$$

$$\begin{aligned} J\delta J &\leq m\delta m\left(\left(m^2 + \sqrt{m^2 - \frac{J^2}{m^2}}\right)^2 + \frac{J}{m^2}\right)\delta m \\ &\leq 2m\delta m\left(m^2 + \sqrt{m^4 - J^2}\right) \end{aligned}$$

$$\boxed{\delta\left(m^2 + \sqrt{m^4 - J^2}\right) \geq 0}$$

Chapter 13

Black hole mechanics

13.1 Black hole thermodynamics

Zeroth: T constant in a system in thermal eq.
—
κ constant at the event horizon (if $T_{\mu\nu}$ obeys the dominant e.c.)

First: $\mathrm{d}U = T\,\mathrm{d}S + \phi\,\mathrm{d}Q + \Omega\,\mathrm{d}J$
—

$$\mathrm{d}M = \frac{\kappa}{8\pi}\,\mathrm{d}A + \phi_h\,\mathrm{d}Q + \Omega_H\,\mathrm{d}J$$

(for a stationary BH)

Second: $\Delta S \geq 0$ in a closed system.
—
$\Delta A_{\mathrm{hor}}(t) \geq 0$ for asympt hor BHs. (If $T_{\mu\nu}$ satisfies the weak e.c. and cosmic sensorship(??))

$$\mathrm{d}s^2 = -\frac{\Delta}{r}\,\mathrm{d}t^2 + \frac{r}{\Delta}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega_{S^2}$$

$\Delta = (r - 2m)$ $r_s = 2m$
Field bines (not unique!!)

$$e^0 = \sqrt{\frac{\Delta}{r}}\,\mathrm{d}t \qquad e^1 = \sqrt{\frac{r}{\Delta}}\,\mathrm{d}r \qquad e^2 = r\,\mathrm{d}\theta \qquad e^3 = r\sin\theta\,\mathrm{d}\phi$$

$$A_{\mathrm{hor}} = \int_{S^2} e^2 \wedge e^3 = r_s^2 \int_{S^2} \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = 4\pi r_s^2 = 16\pi m^2$$

$$\boxed{m^2 = \frac{A_{\mathrm{hor}}}{16\pi}}$$

$$2m\delta m = \frac{\delta A_{\text{hor}}}{16\pi} \qquad \kappa \frac{1}{4m}$$

$$\delta m = \frac{1}{32\pi m}\delta A_{\text{hor}} = \frac{1}{8\pi}\kappa\delta A_{\text{hor}}$$

$$T = \frac{\kappa}{2\pi} \qquad \delta m = T\delta\left(\frac{A_{\text{hor}}}{4}\right)$$

By analogy we get

$$S_{BH} = \frac{A_{\text{hor}}}{4}$$

$$A_{\text{hor}} = 8\pi\left(m^2 + \sqrt{m^2 - J^2}\right)$$

$$\delta\left(m^2 + \sqrt{m^4 - J^2}\right) \geq 0$$

$$\delta A_{\text{hor}} \geq 0$$

Is this actually entropy?? I.e.

$$S = \log N$$

number of microstates $(e^1 0)^8 0$. But no hair theorem in GR states that BH unique if defined mass and charge, so only 1 state.
(Extremal BH: $\sim$ infinitely far away for everybody)
Second law means a black hole cannot split.
Consider two black holes (of mass M_1, M_1) merge to form a black hole of mass M. The energy emitted as gravitational waves is $E = M_1 + M_2 - M$

$$\eta = \frac{E}{M_1 + M_2} = 1 - \frac{M}{M_1 + M_2}$$

is the efficiency (by definition) Now the areas are $A_{1,2} = 16\pi M_{1,2}^2$. Because $A \geq A_1 + A_2$,

$$M \geq \sqrt{M_1^2 + M_2^2}$$

This puts a constraint on the efficiency:

$$\eta = 1 - \sqrt{\frac{m_1^2 + M_2^2}{(M_1 + M_2)^2}} \leq 1 - \frac{1}{\sqrt{2}}$$

Which is about 23%, which is ver very good.
In the opposite process (splitting)

$$M \leq \sqrt{M_1^2 + M_2^2} \leq M_1 + M_2$$

But also (from energy conservation)

$$M \geq M_1 + M_2$$

13.1.1 Energy conditions

$$T_{\mu\nu} = (p + \rho)u_\mu u_\nu + p g_{\mu\nu}$$

- Dominant: $\rho \geq |p|$
  We want energy flow to go slower than the speed of light
- Weak: $T_{\mu\nu} t^\mu t^\nu \geq 0 \qquad t^2 < 0$
- Null: $T_{\mu\nu} l^\mu l^\nu \geq 0 \qquad l^2 = 0$
- Strong:

$$P + \rho \geq 0 \qquad \rho + 3p \geq 0 \qquad \underbrace{(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T)}_{R_{\mu\nu}} t^\mu t^\nu \geq 0 \qquad t^2 \leq 0$$

Useful to prove theorems, but there are forms of matter that violate them.

Chapter 14

Beyond GR

No experimental need to go beyond GR. Dark matter looks like it really is matter. Theoretically we have some puzzles that need to be solved. One example is the entropy of black holes. Also GR and QM are incompatible. GR clearly incomplete: there should be singularities. Currently gravity tested down to micrometer.
Metric emergent phenomenon?
Locality, equivalence principle and unitarity can't all be true

Chapter 15

Gravity as geometry

How to describe gravity?

- Universality: every entity interacts with all others (not dependent on an intrinsic property like charge)
  - Newton: gravitational charge = mass
  - SR: mass = energy
- Acceleration of an object does not depend on its mass, only on where you put it (geometrical nature of gravity):

$$\mathbf{F} = m\mathbf{a} = G_N \frac{Mm}{r^2} \Rightarrow \mathbf{a} = G_N \frac{M}{r^2}$$

- Energy sources generate gravity $\rightarrow$ modify spacetime.
  Spacetime geometry affects the motion of energy sources.

$$\mathrm{d}s^2 = \mathrm{d}x^\mu g_{\mu\nu} \, \mathrm{d}x^\nu$$

  Where g is a spacetime metric.

$$\text{gravity} \Rightarrow g_{\mu\nu} \neq \eta_{\mu\nu}$$

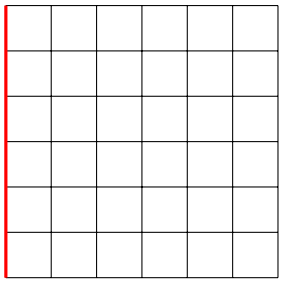
The opposite implication is not true! It could simply be due to a non inertial reference frame.

Example

Example of a different metric due to rotating reference frame.
We start in an inertial frame (while setting $c = 1$)

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$$

(a) $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$



(b) $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2$

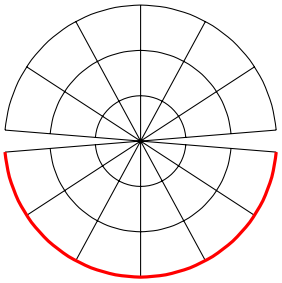


Figure 15.1: Transformation $x = r\cos\theta, y = r\sin\theta$

We now transform to the following rotating reference frame:

$$\begin{cases} x' = \cos(\omega t)x + \sin(\omega t)y \\ y' = -\sin(\omega t)x + \cos(\omega t)y \\ z' = z \\ t' = t \end{cases}$$

We now get the following squared line element:

$$\mathrm{d}s^2 = -(1 - \omega^2(x^2 + y^2))\,\mathrm{d}t^2 + 2\omega(x\,\mathrm{d}y - y\,\mathrm{d}x)\,\mathrm{d}t + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$$

Gravity is given by a change in metric so you cannot transform and get Minkowski.

15.1 Some more geometry

gravity $\leftrightarrow$ geometry of spacetime
Different (non-Euclidean) geometries possible. Visualization in extra dimension. Also intrinsic definition?

- Axiomatic
- From the distance between nearby points (then larger distances by integration)

Any free observer sees the same spacetime, measured with the line element

$$\mathrm{d}s^2 = \mathrm{d}x^\mu \eta_{\mu\nu}\,\mathrm{d}x^\nu$$

using the Minkowsky metric. Here we use the mostly plus convention:

$$\eta = \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1}_3 \end{pmatrix}$$

See also intrinsic and extrinsic curvature.

Figure 15.2: Positive curvature

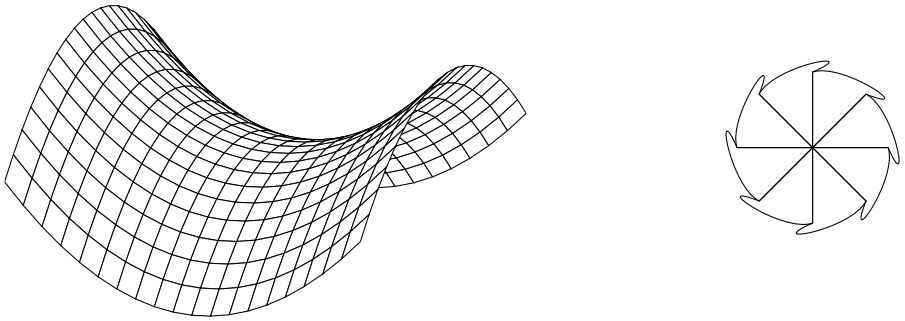


Figure 15.3: Negative curvature

(a) Intrinsic curvature

(b) Extrinsic curvature

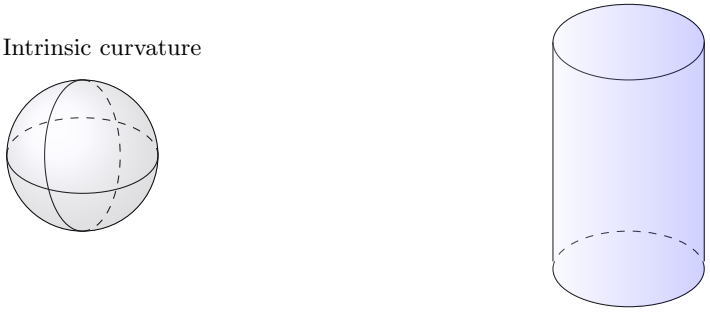


Figure 15.4: Intrinsic and extrinsic curvature

Chapter 16

The equivalence principle

Given an arbitrary spacetime

$$\mathrm{d}s^2 = \mathrm{d}x^\mu g_{\mu\nu}(x)\,\mathrm{d}x^\nu$$

we want the following properties to hold:

1. Special relativity should be a special case:

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \mathcal{O}(x - x_0)^2$$

2. The same physics should hold for arbitrary transformations

Coordinates have no meaning

3. there should be gravity

We have

$$x^\mu \mapsto x'^\mu = f^\mu(x)$$

with

$$\mathrm{d}x'^\mu = \Lambda^\mu{}_\nu(x)\,\mathrm{d}x^\nu \qquad \Lambda \in \mathrm{GL}(1,3)$$

So locally we have a Lorentz group.

Part XI

Quantum mechanics

Chapter 1

Measurement theory

Chapter 2

Postulates

2.1 States, observables and measurement

2.1.1 States

A state on a Hilbert space is a positive trace class operator with trace 1.

Proposition XI.1. *Let ρ be a state. The following are equivalent:*

1. *ρ is extremal;*
2. *ρ is a projector;*
3. *$\rho = \rho^2$;*
4. *$\mathrm{Tr}[\rho^2] = 1$;*
5. *$\|\rho\| = 1$;*
6. *$S(\rho) := -\mathrm{Tr}[\rho \ln \rho] = 0$.*

The quantity $S(\rho) := -\mathrm{Tr}[\rho \ln \rho]$ is called the von Neumann entropy.

2.1.2 Effects

An effect on a Hilbert space is a bounded operator in $[0, \mathrm{id}]$.

Chapter 3

Quantum theory

3.1 Quantum statistics

Let $\mathcal{H}$ be a Hilbert space, A a symmetric operator on $\mathcal{H}$ and ψ a unit vector in $\mathrm{dom}(A)$. Then we define

- the expectation value of A at ψ is defined as $\langle A \rangle_\psi := Q_A(\psi) = \langle \psi | A | \psi \rangle$;
- the uncertainty or standard deviation of A at ψ is defined as

$$\Delta_\psi A := \| \big(A - \langle A \rangle_\psi \, \mathrm{id} \big) \psi \|.$$

The expectation value $\langle A \rangle_\psi$ is always real (by **??**).

Lemma XI.2. $\langle A \rangle_\psi = \langle \psi | A | \psi \rangle = \mathrm{Tr}\left(A | \psi \rangle \langle \psi | \right)$.

Lemma XI.3. *Let $\mathcal{H}$ be a Hilbert space, A a symmetric operator on $\mathcal{H}$ and ψ a unit vector in* $\mathrm{dom}(A)$. *Then $\Delta_\psi A = 0$ if and only if ψ is an eigenvalue of A.*
In this case the eigenvalue of ψ is $\langle A \rangle_\psi$.

Proof. Assume $\Delta_\psi A = \| \big(A - \langle A \rangle_\psi \, \mathrm{id} \big) \psi \| = 0$. Then $\big(A - \langle A \rangle_\psi \, \mathrm{id} \big) \psi = 0$, or $A\psi = \langle A \rangle_\psi \, \psi$, which means that ψ is an eigenvector with eigenvalue $\langle A \rangle_\psi$.
Conversely, suppose ψ is an eigenvector of A with eigenvalue λ, i.e. $A\psi = \lambda\psi$. Then

$$\langle A \rangle_\psi = \langle \psi, A\psi \rangle = \langle \psi, \lambda\psi \rangle = \lambda \langle \psi, \psi \rangle = \lambda.$$

Thus $A\psi = \langle A \rangle_\psi \, \psi$ and $\Delta_\psi A = 0$. □

Lemma XI.4. *Let $\mathcal{H}$ be a Hilbert space, A a symmetric operator on $\mathcal{H}$ and ψ a unit vector in* $\mathrm{dom}(A)$. *Then*

$$\big(\Delta_\psi A \big)^2 = \Big\langle \big(A - \langle A \rangle_\psi \, \mathrm{id} \big) \psi, \big(A - \langle A \rangle_\psi \, \mathrm{id} \big) \psi \Big\rangle = \langle A\psi, A\psi \rangle - \langle A \rangle_\psi^2 \, .$$

If ψ is also in $\mathrm{dom}(A^2)$, *then*

$$\big(\Delta_\psi A \big)^2 = \Big\langle \big(A - \langle A \rangle_\psi \, \mathrm{id} \big)^2 \Big\rangle_\psi = \big\langle A^2 \big\rangle_\psi - \langle A \rangle_\psi^2 \, .$$

Proof. We calculate (using symmetry, the fact that $\langle A\rangle_\psi$ is real and the fact that ψ is a unit vector)

$$\begin{aligned}\big(\Delta_\psi A\big)^2 &= \Big\langle \big(A - \langle A\rangle_\psi \,\mathrm{id}\,\big)\psi, \big(A - \langle A\rangle_\psi \,\mathrm{id}\,\big)\psi \Big\rangle \\ &= \langle A\psi, A\psi\rangle - \Big\langle A\psi, \langle A\rangle_\psi \,\psi\Big\rangle - \Big\langle \langle A\rangle_\psi \,\psi, A\psi\Big\rangle + \Big\langle \langle A\rangle_\psi \,\psi, \langle A\rangle_\psi \,\psi\Big\rangle \\ &= \langle A\psi, A\psi\rangle - \langle A\rangle_\psi \langle \psi, A\psi\rangle - \langle A\rangle_\psi \langle \psi, A\psi\rangle + \langle A\rangle_\psi^2 \cancel{\langle \psi, \psi\rangle} \\ &= \langle A\psi, A\psi\rangle - \langle A\rangle_\psi^2 - \langle A\rangle_\psi^2 + \langle A\rangle_\psi^2 \\ &= \langle A\psi, A\psi\rangle - \langle A\rangle_\psi^2 .\end{aligned}$$

Now assume $\psi \in \mathrm{dom}(A^2)$. Then $\psi \in \mathrm{dom}\big((A - \langle A\rangle_\psi \,\mathrm{id})^2\big)$ and thus $(A - \langle A\rangle_\psi \,\mathrm{id})\psi \in \mathrm{dom}\big(A - \langle A\rangle_\psi \,\mathrm{id}\,\big)$. The first equality follows by symmetry. For the second equality we have a similar reasoning starting from the equation above. □

3.1.1 Uncertainty relations

Theorem XI.5 (Heisenberg's uncertainty principle)**.** *Let $\mathcal{H}$ be a Hilbert space and A, B symmetric operators on $\mathcal{H}$. Let ψ be a unit vector in* $\mathrm{dom}(AB) \cap \mathrm{dom}(BA)$*. Then*

$$(\Delta_\psi A)(\Delta_\psi B) \geq \frac{1}{2}\big|\,\langle [A,B]\rangle_\psi\,\big|.$$

Proof. By the Cauchy-Schwarz inequality, XVI.8, we have

$$\begin{aligned}(\Delta_\psi A)(\Delta_\psi B) &= \|\big(A - \langle A\rangle_\psi \,\mathrm{id}\,\big)\psi\| \;\|\big(B - \langle B\rangle_\psi \,\mathrm{id}\,\big)\psi\| \\ &\geq |\Big\langle \big(A - \langle A\rangle_\psi \,\mathrm{id}\,\big)\psi, \big(B - \langle B\rangle_\psi \,\mathrm{id}\,\big)\psi\Big\rangle| \\ &\geq |\,\mathfrak{Im}\,\Big\langle \big(A - \langle A\rangle_\psi \,\mathrm{id}\,\big)\psi, \big(B - \langle B\rangle_\psi \,\mathrm{id}\,\big)\psi\Big\rangle| \\ &= |\,\mathfrak{Im}\,\Big(\langle A\psi, B\psi\rangle - \Big\langle A\psi, \langle B\rangle_\psi \,\psi\Big\rangle - \Big\langle \langle A\rangle_\psi \,\psi, B\psi\Big\rangle + \Big\langle \langle A\rangle_\psi \,\psi, \langle B\rangle_\psi \,\psi\Big\rangle\Big)| \\ &= |\,\mathfrak{Im}\,\Big(\langle A\psi, B\psi\rangle - \cancel{\langle A\rangle_\psi \langle B\rangle_\psi}\Big)| \\ &= \frac{1}{2}|\langle A\psi, B\psi\rangle - \overline{\langle A\psi, B\psi\rangle}| = \frac{1}{2}|\langle A\psi, B\psi\rangle - \langle B\psi, A\psi\rangle|.\end{aligned}$$

Since $\psi \in \mathrm{dom}(AB)$, we have $B\psi \in \mathrm{dom}(A)$. Similarly $A\psi \in \mathrm{dom}(B)$. Thus, by symmetry, we have

$$\frac{1}{2}|\langle A\psi, B\psi\rangle - \langle B\psi, A\psi\rangle| = \frac{1}{2}|\langle \psi, AB\psi\rangle - \langle \psi, BA\psi\rangle| = \frac{1}{2}|\langle \psi, [A,B]\psi\rangle| = \frac{1}{2}\big|\,\langle [A,B]\rangle_\psi\,\big|.$$

□

3.2 Time evolution

3.2.1 The Schrödinger equation

$$H = i\hbar\frac{\mathrm{d}}{\mathrm{d}t}$$

Lemma XI.6. $U = e^{Ht/i\hbar} = e^{-iHt/\hbar}$.

3.2.2 Schrödinger and Heisenberg pictures

Lemma XI.7. *Let $\mathcal{H}$ be a Hilbert space and $\{e_i\}$ an orthonormal basis. Let U be a unitary operator on $\mathcal{H}$. Then $\{Ue_i\}$ is also an orthonormal basis for $\mathcal{H}$.*

$\{e_i\}$ gives Schrödinger and $\{Ue_i\}$ gives Heisenberg.

3.2.3 Adiabatic theorem

Theorem XI.8. *Let $\mathcal{H}$ be a Hilbert space and consider a path $H : [0,1] \to \mathcal{SA}(\mathcal{H}) : s \mapsto H(s)$ of self-adjoint operators on the Hilbert space.*
Take $\epsilon > 0$. If the Hamiltonian of a system is given by $H(\epsilon t)$ for times $t \in [0, \epsilon^{-1}]$. Then the state ρ of the system satisfies

$$i\epsilon \frac{\mathrm{d}\rho(s)}{\mathrm{d}s} = \big[H(s), \rho(s)\big],$$

where $s = \epsilon t \in [0,1]$.
Assume $\lambda(s) \in \sigma\big(H(s)\big)$ is an isolated point of the spectrum for all $s \in [0,1]$. It is an eigenvalue and let $P(s)$ be the orthogonal projector onto the eigenspace. Set $\Delta(s) \coloneqq \inf d\big(\lambda, \sigma(H(s)) \setminus \{\lambda(s)\}\big)$.
Assume $\rho(0) \leq P(0)$. Consider the fidelity $F(s) \coloneqq \operatorname{Tr}\big(P(s)\rho(s)\big)$. Then

$$F(1) \geq 1 - \epsilon O\left(\frac{\|H'\|^2}{\Delta^3} + \frac{\|H''\|}{\Delta^2}\right).$$

Proof. We first write down a differential equation for the fidelity (where $' = \frac{\mathrm{d}}{\mathrm{d}s}$)

$$\frac{\mathrm{d}F(s)}{\mathrm{d}s} = \operatorname{Tr}(P'\rho) + \operatorname{Tr}(P\rho').$$

The second term is zero because P commutes with H:

$$\begin{aligned}
\operatorname{Tr}(P\rho') &= -\epsilon^{-1} i \operatorname{Tr}\Big(P\big[H,\rho\big]\Big) \\
&= -\epsilon^{-1} i \operatorname{Tr}\Big(P(H\rho - \rho H)\Big) \\
&= -\epsilon^{-1} i \operatorname{Tr}\Big(PH\rho\Big) + \epsilon^{-1} i \operatorname{Tr}\Big(\rho HP\Big) \\
&= -\epsilon^{-1} i \operatorname{Tr}\Big(PH\rho\Big) + \epsilon^{-1} i \operatorname{Tr}\Big(\rho PH\Big) \\
&= -\epsilon^{-1} i \operatorname{Tr}\Big(\big[PH,\rho\big]\Big) = 0,
\end{aligned}$$

by XVI.175.
Now set $Q(s) = \mathrm{id} - P(s)$ and consider the pseudoinverse $(H - \lambda\,\mathrm{id})^+$ (TODO by continuous functional calculus?). Then $(H - \lambda\,\mathrm{id})(H - \lambda\,\mathrm{id})^+ = Q = (H - \lambda\,\mathrm{id})^+(H - \lambda\,\mathrm{id})$ by continuous

functional calculus (TODO ref). Now we can expand $P' = PP'Q + QP'P$ (by XIV.121.2), so

$$\begin{aligned}
\tfrac{\mathrm{d}F(s)}{\mathrm{d}s} &= \operatorname{Tr}(P'\rho) \\
&= \operatorname{Tr}(PP'Q\rho + QP'P\rho) \\
&= \operatorname{Tr}\Big(PP'(H-\lambda\,\mathrm{id})^+(H-\lambda\,\mathrm{id})\rho + (H-\lambda\,\mathrm{id})(H-\lambda\,\mathrm{id})^+P'P\rho\Big) \\
&= \operatorname{Tr}\Big(P'(H-\lambda\,\mathrm{id})^+(H-\lambda\,\mathrm{id})\rho P + (H-\lambda\,\mathrm{id})^+P'P\rho(H-\lambda\,\mathrm{id})\Big) \\
&= \operatorname{Tr}\Big(P'(H-\lambda\,\mathrm{id})^+\big(H\rho P - \rho\lambda P\big) + (H-\lambda\,\mathrm{id})^+P'\big(P\rho H - \lambda P\rho\big)\Big) \\
&= \operatorname{Tr}\Big(P'(H-\lambda\,\mathrm{id})^+\big(H\rho - \rho H\big)P + (H-\lambda\,\mathrm{id})^+P'P\big(\rho H - H\rho\big)\Big) \\
&= \operatorname{Tr}\Big(P'(H-\lambda\,\mathrm{id})^+\big[H,\rho\big]P - (H-\lambda\,\mathrm{id})^+P'P\big[H,\rho\big]\Big) \\
&= \operatorname{Tr}\Big(PP'(H-\lambda\,\mathrm{id})^+\big[H,\rho\big] - (H-\lambda\,\mathrm{id})^+P'P\big[H,\rho\big]\Big) \\
&= \operatorname{Tr}\Big(P'Q(H-\lambda\,\mathrm{id})^+\big[H,\rho\big] - (H-\lambda\,\mathrm{id})^+QP'\big[H,\rho\big]\Big) \\
&= \operatorname{Tr}\Big(P'(H-\lambda\,\mathrm{id})^+\big[H,\rho\big] - (H-\lambda\,\mathrm{id})^+P'\big[H,\rho\big]\Big) \\
&= \operatorname{Tr}\Big(\big[P',(H-\lambda\,\mathrm{id})^+\big]\cdot\big[H,\rho\big]\Big) \\
&= i\epsilon\operatorname{Tr}\Big(\big[P',(H-\lambda\,\mathrm{id})^+\big]\rho'\Big)
\end{aligned}$$

Integrating this w.r.t. s gives

$$F(1) - F(0) = i\epsilon\int_0^1 \operatorname{Tr}\Big(\big[P',(H-\lambda\,\mathrm{id})^+\big]\rho'\Big)\,\mathrm{d}s.$$

We fill in that $F(0) = 1$ and perform integration by parts to obtain

$$\begin{aligned}
F(1) &= 1 + i\epsilon\int_0^1 \operatorname{Tr}\Big(\big[P',(H-\lambda\,\mathrm{id})^+\big]\rho'\Big)\,\mathrm{d}s \\
&= 1 + i\epsilon\operatorname{Tr}\Big(\big[(H-\lambda\,\mathrm{id})^+,P'\big]\rho\Big)_0^1 - i\epsilon\int_0^1 \operatorname{Tr}\Big(\big[(H-\lambda\,\mathrm{id})^+,P'\big]'\rho\Big)\,\mathrm{d}s
\end{aligned}$$

□

Proposition XI.9. *Let $\mathcal{H}$ be a Hilbert space, $H : [0,1] \to \mathcal{SA}(\mathcal{H})$ and $P(s)$ the projector on an eigenspace associated to some subset of the spectrum σ' that is separated from the rest by a gap $g(s)$. Let $\rho(s)$ be the solution the the Heisenberg equation*

$$i\epsilon\frac{\mathrm{d}\rho(s)}{\mathrm{d}s} = [H(s),\rho(s)] \qquad \rho(0) = P(0).$$

Then we can write

$$\rho(s) = \sum_{n=0}^{N}\epsilon^n B_n(s) - \epsilon^N\int_0^s U_\epsilon(s,r)\dot{B}_N(r)U_\epsilon(r,s)\,\mathrm{d}r,$$

where the B_n satisfy the recursion relation

$$\begin{cases} B_0(s) = P(s) \\ B_n(s) = \frac{1}{2\pi}\int_\Gamma R_H(z)\big[P,\dot{B}_{n-1}\big]R_H(z)\,\mathrm{d}z + S_n - 2PS_nP \end{cases},$$

where Γ is some Jordan curve that contains σ', but not the rest of the spectrum,

$$S_n \coloneqq \sum_{m=1}^{n-1} B_m B_{n-m}$$

and $U_\epsilon(s,t)$ is the propagator that satisfies

$$\begin{cases} i\epsilon \frac{\partial}{\partial s} U_\epsilon(s,r) = H U_\epsilon(s,r) \\ U_\epsilon(r,r) = \mathrm{id}\,. \end{cases}$$

Lemma XI.10. *Let $\mathcal{H}$ be a Hilbert space and $H : [0,1] \to \mathcal{SA}(\mathcal{H})$ a function that belongs to the Gevrey class $G^\alpha(R)$. Then*

$$\| \frac{\mathrm{d}^k B_n(s)}{\mathrm{d}s^k} \| \leq L(n,k) \coloneqq \frac{1}{(10n+0.3)^2} g^{-2n-k} \big(2CR(k+3n)\big)^{k+3n}.$$

Proof. □

3.3 Quantum thermodynamics

3.3.1 KMS states

`http://www.ueltschi.org/teaching/2022-TCC.html`.

Chapter 4

Other formulations of quantum mechanics

4.1 Path-integral formulation

4.2 Phase-space formulation

4.2.1 The Wigner transform

4.2.1.1 The Wigner function

4.2.1.2 The Wigner function of a pure state

Let $\psi \in L^2(\mathbb{R}, \mathbb{C})$ be a state, meaning that $\|\psi\|_2 = 1$. Then the corresponding Wigner function W_ψ is defined by

$$W_\psi : \mathbb{R}^2 \to \mathbb{C} : (x, p) \mapsto \frac{1}{\pi\hbar}$$

We have defined the codomain of W_ψ as $\mathbb{C}$, but in fact the image is real.

Lemma XI.11. *Let $\psi \in L^2(\mathbb{R}, \mathbb{C})$ be such that $\|\psi\|_2 = 1$. Then* $\mathrm{im}(W_\psi) \subseteq \mathbb{R}$.

4.2.1.3 The Wigner function of a mixed state

Chapter 5

Approximations

5.1 Approximating eigenvectors

5.1.1 Power series expansion of a non-degenerate level

5.2 Approximating evolutions

Chapter 6

Investigations of systems

6.1 Stepped potentials

(Use density to solve general potentials?)

6.2 Coulomb interaction

6.3 Harmonic oscillator

6.3.1 Creation and annihilation operators

Lemma XI.12.

6.3.2 Gibbs state

Proposition XI.13. *Let $H = \hbar\omega\big(a^*a + \frac{1}{2}\big)$ be the Hamiltonian of a harmonic oscillator, with $\omega > 0$. Consider the Gibbs state*

$$\rho = \frac{e^{-\beta H}}{\mathcal{Z}},$$

with inverse temperature $\beta > 0$. Then

1. *the Gibbs state exists;*
2. *the the partition function is given by $\mathcal{Z} = \frac{1}{2\sinh\left(\frac{\beta\hbar\omega}{2}\right)}$;*
3. *the Gibbs state can be expanded as*

$$\rho = \big(1 - e^{-\beta\hbar\omega}\big)\sum_{n\in\mathbb{N}} e^{-n\beta\hbar\omega}|n\rangle\langle n|,$$

where the sum converges (TODO: weakly or strongly??).

Proof. (1) Since βH is a positive self-adjoint operator, $-\beta H$ is dissipative by XVI.124, and thus the existence of $e^{-\beta H}$ is given by XVII.38 and the fact that $\lambda \in \rho(-\beta H)$ for all $\lambda > 0$, so $\lambda\,\mathrm{id} + \beta H$ is surjective.

The existence of the Gibbs state will then follow if we can show that $e^{-\beta H}$ is trace-class. This is given by (2).
(2) We use XVII.15.2 to calculate

$$\begin{aligned}
\operatorname{Tr}\left(e^{-\beta H}\right) &= \sum_{n\in\mathbb{N}} \langle n|e^{-\beta H}|n\rangle \\
&= \sum_{n\in\mathbb{N}} \langle n|e^{-\beta\hbar\omega\left(n+\frac{1}{2}\right)}|n\rangle \\
&= e^{-\frac{\beta\hbar\omega}{2}} \sum_{n\in\mathbb{N}} \left(e^{-\beta\hbar\omega}\right)^n \\
&= e^{-\frac{\beta\hbar\omega}{2}} \frac{1}{1-e^{-\beta\hbar\omega}} \\
&= \frac{1}{e^{\frac{\beta\hbar\omega}{2}} - e^{-\frac{\beta\hbar\omega}{2}}} = \frac{1}{2\sinh\left(\frac{\beta\hbar\omega}{2}\right)}.
\end{aligned}$$

(3) TODO. □

Chapter 7

Circuit model

7.1 Qubits

A qubit is an element of $\mathbb{C}P^2$.

7.1.1 $\mathfrak{su}(2)$ and the Pauli matrices

The qubit Hamiltonians are elements of $\mathfrak{u}(2)$. Getting rid of the arbitrary global phase gives us elements of $\mathfrak{su}(2)$.

Lemma XI.14. *For all $\sigma \in \mathfrak{su}(2)$, the eigenvalues λ_1, λ_2 are real and $\lambda_1 = -\lambda_2$.*

Proof. For all $\sigma \in \mathfrak{su}(2)$, σ is self-adjoint and $\mathrm{Tr}[\sigma] = \lambda_1 + \lambda_2 = 0$. □

Corollary XI.14.1. *For all $\sigma \in \mathfrak{su}(2)$, we have $\|\sigma\|_2 = \sqrt{2}\|\sigma\|$, where $\|\cdot\|_2$ is the Hilbert-Schmidt norm.*

Proof. We have $\|\sigma\|_2 = \sqrt{\lambda_1^2 + \lambda_2^2} = \sqrt{2}|\lambda_1| = \sqrt{2}\|\sigma\|$. □

Corollary XI.14.2. *The inner product*

$$\langle \cdot, \cdot \rangle : (\sigma_1, \sigma_2) \mapsto \frac{1}{2}\,\mathrm{Tr}[\sigma_1 \sigma_2]$$

on $\mathfrak{su}(2)$ yields the operator norm as the norm associated with this inner product.

Corollary XI.14.3. *Let σ be a unit vector in $\mathfrak{su}(2)$. Then σ has eigenvalues ± 1. This means σ is unitary.*

Corollary XI.14.4. *There is a bijection between the unit vectors in $\mathfrak{su}(2)$ and the rank-1 projections on $\mathbb{C}^2$ given by $\mathfrak{su}(2)/\mathbb{R} \to \sigma \mapsto E_1^\sigma$, where E_1^σ is the eigenspace of σ associated with eigenvalue $+1$.*

Proof. We construct an inverse of the map. Let P_1 be the orthogonal projector on E_1^σ. Then $\sigma = 2P_1 - \mathrm{id}$. □

Proposition XI.15. *Let $\sigma \in \mathfrak{su}(2)$. Then $\sigma^2 = \|\sigma\|^2\, \mathbf{1}$.*

Proof. We have

$$\sigma^2 = \|\sigma\|^2 \left(\frac{\sigma}{\|\sigma\|}\right)^2 = \|\sigma\|^2 \frac{\sigma}{\|\sigma\|} \left(\frac{\sigma}{\|\sigma\|}\right)^* = \|\sigma\|^2 \mathbf{1},$$

where we have used that $\frac{\sigma}{\|\sigma\|}$ is unitary by XI.14.3. □

Corollary XI.15.1. *The real algebra generated by* $\mathfrak{su}(2)$ *is the Clifford algebra* $\mathrm{Cl}_{3,0}$*. The unit pseudoscalar is* $i\mathbf{1}$*.*

Note that $\mathfrak{su}(2)$ is a Lie algebra and thus closed under the Lie bracket, but not an algebra under operator composition. Thus the algebra generated by $\mathfrak{su}(2)$ is larger than $\mathfrak{su}(2)$ (indeed $\sigma^2 = \|\sigma\|^2 \mathbf{1} \notin \mathfrak{su}(2)$).

7.1.1.1 Pauli matrices

Proposition XI.16 (Pauli matrices)**.** *The Clifford algebra* $\mathfrak{su}(2)$ *has an orthonormal basis*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Proposition XI.17. *The Pauli matrices obey*

1. $[\sigma_i, \sigma_j] = 2i\varepsilon_{i,j,k}\sigma_k$*;*
2. $\{\sigma_i, \sigma_j\} = 2\delta_{i,j}\mathbb{1}_2$*;*
3. $\sigma_i\sigma_j = \delta_{i,j}\mathbb{1}_2 + i\varepsilon_{i,j,k}\sigma_k$

Corollary XI.17.1. *TODO* $\sigma_i\sigma_j\sigma_i =$*.*

eg $\sigma_x\sigma_y\sigma_x = -\sigma_y$.

7.1.1.2 Qubit bases

We call

- the eigenbasis of σ_z the Z-basis or computational basis and write the elements

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix};$$

- the eigenbasis of σ_x the X-basis or coherence basis and write the elements

$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \qquad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

7.1.2 The Bloch sphere

Because $\mathfrak{su}(2)$ is the Clifford algebra $\mathrm{Cl}_{3,0}$, we can use the unit sphere in $\mathbb{R}^3$ to represent the unit vectors in $\mathfrak{su}(2)$. By XI.14.4 we can also use it to represent the rank-1 projections on $\mathbb{C}^2$. The unit sphere with as x, y, z axes the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ is called the Bloch sphere.
TODO image of Bloch sphere.
TODO spherical coordinates: $\cos\theta/2\,|0\rangle + e^{i\phi}\sin\theta/2\,|1\rangle$

Lemma XI.18.

$$e^{\phi/2i\sigma}e^{\theta/2i\sigma^\perp}\,|1\rangle = \cos\theta/2\,|0\rangle + e^{i\phi}\sin\theta/2\,|1\rangle$$

7.2 Gates

7.2.1 One-qubit gates

7.2.1.1 Hadamard gate

The Hadamard gate or Walsh-Hadamard gate W is the linear operation determined by the following matrix:

$$W = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

We have the following mappings

$$W|0\rangle = |+\rangle \qquad W|+\rangle = |0\rangle$$
$$W|1\rangle = |-\rangle \qquad W|-\rangle = |1\rangle$$

Lemma XI.19. *The Hadamard gate is its own inverse:* $W^2 = \mathbb{1}$.

7.2.1.2 Pauli gates

Each of the Pauli matrices determines an operation on $\mathbb{C}^2$. In this context we often write X, Y and Z for σ_x, σ_y and σ_z.

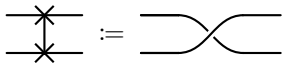
Lemma XI.20. *The Pauli gates are their own inverses:* $X^2 = Y^2 = Z^2 = \mathbb{1}$.

7.2.2 Controlled gates

Let H_1, H_2 be two Hilbert spaces. Let P be a projector on H_1 and L an operator on H_2. The operation of L controlled by P is the operation $P \otimes L + (\mathrm{id}_{H_1} - P) \otimes \mathrm{id}_{H_2}$ in $H_1 \otimes H_2 \to H_1 \otimes H_2$.

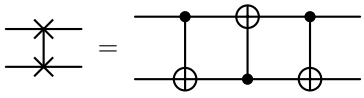
7.2.3 Two-qubit gates

7.2.3.1 Swap



$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Proposition XI.21.



Chapter 8

Eigenpath traversal

8.1 Quantum Zeno effect

Proposition XI.22. *Consider a path of states $|\psi(s)\rangle$ where $s \in [0,1]$. Assume that, for fixed d and all δ,*

$$|\langle\psi(s)|\psi(s+\delta\rangle|^2 \geq 1 - d^2\delta^2.$$

Then

8.2 Adiabatic quantum computation

8.3 Evolution through measurement

8.3.1 Phase randomisation

$|\psi_0\rangle = |E_0(0)\rangle = \sum_i \alpha_i |E_i(s_1)\rangle$

$$\begin{aligned}\rho(s_1) &= \frac{1}{T}\int_0^T e^{-iH(s_1)t}|\psi_0\rangle\langle\psi_0|e^{iH(s_1)t}\,\mathrm{d}t \\ &= \sum_{i,j}\alpha_i\alpha_j^*\frac{1}{T}\left(\int_0^T e^{-i(E_i-E_j)t}\,\mathrm{d}t\right)|E_i\rangle\langle E_j| \\ &= \sum_{i,j}\alpha_i\alpha_j^*\left(\frac{i(e^{-i(E_i-E_j)T}-1)}{T(E_i-E_j)}\right)|E_i\rangle\langle E_j|\end{aligned}$$

Theorem XI.23 (Randomised dephasing)**.** *Let $|\psi(s)\rangle$ be a nondegenerate eigenstate of $H(s)$ and $\{\omega_j\}$ the energy differences to the other eigenstates $|\psi_j(s)\rangle$. Let T be a random variable. Then, for all states ρ, we have*

$$\|(M_l - e^{-iH(s)T}\rho e^{-iH(s)T}\|_{tr} \leq \epsilon = \sup_{\omega_j}|\Phi(\omega_j)|$$

Chapter 9

Variational quantum algorithms

Chapter 10

Algorithms

10.1 Some building blocks

10.1.1 Preparing states

10.1.1.1 Uniform superposition

Lemma XI.24. *Consider the Hilbert space $\mathcal{H}_{\mathbb{F}_2^k}$. We can generate a uniform superposition of computational basis states by applying a Hadamard gate to each register, initialised to zero:*

$$\frac{1}{2^k}\sum_{b\in\mathbb{F}_2^k}|b\rangle = W^{\otimes k}|0\rangle^{\otimes k}.$$

Projection on the n-dimensional uniform superposition is given by $\frac{1}{n}\mathbb{J}_{n\times n}$.

10.1.2 Quantum Fourier transform

Set $R_k := \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{pmatrix}$.
TODO: produces swapped result!
Requires n Hadamard gates and $\frac{n(n-1)}{2}$ controlled R-gates. At most $n/2$ swaps. Each with three CNOTs (XI.21).

10.1.3 Quantum phase estimation

Needs large number of ancillas.

Proposition XI.25. *Let U be a unitary operator and $|\psi\rangle$ an eigenvector of U. acting on an m-qubit register and .*

10.1.4 Quantum singular value transformation

10.2 Quantum oracle querying

10.2.1 Oracle set-up

For any finite set A, we define a mapping

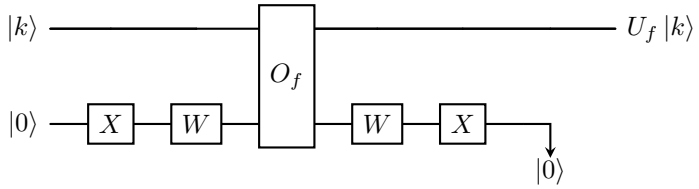
$$O : (A \to \mathbb{F}_2) \to \mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_{\mathbb{F}_2}) : f \mapsto O_f$$

where

$$O_f(|k\rangle \otimes |q\rangle) \coloneqq |k\rangle \otimes |f(k) + q\rangle .$$

We call O_f the quantum oracle of f.

Lemma XI.26. *We can convert an oracle O_f to a unitary operator $U_f : \mathcal{H}_A \to \mathcal{H}_A : |k\rangle \mapsto (-1)^{f(k)} |k\rangle$ using only a constant number of Hadamard and NOT gates:*



If A is finite, then $U_f = \mathrm{id} - 2\sum_{k \in f^{-1}(1)} |k\rangle\langle k|$, which is a Householder matrix.

Proof. We show that the circuit transforms $|k\rangle$ into $(-1)^{f(k)} |k\rangle$.
Given input $|k\rangle$, the oracle acts on $|k\rangle \otimes |-\rangle = |k\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = \frac{|k\rangle\otimes|0\rangle}{\sqrt{2}} - \frac{|k\rangle\otimes|1\rangle}{\sqrt{2}}$ and produces

$$\frac{|k\rangle \otimes |f(k)\rangle}{\sqrt{2}} - \frac{|k\rangle \otimes |f(k)+1\rangle}{\sqrt{2}} = \begin{cases} |k\rangle \otimes |-\rangle & (f(k) = 0) \\ -|k\rangle \otimes |-\rangle & (f(k) = 1) \end{cases}$$
$$= (-1)^{f(k)} |k\rangle \otimes |-\rangle .$$

Given that XW is the inverse of WX, the output $U_f |k\rangle$ is equal to $(-1)^{f(k)} |k\rangle$. The second register is left in the state $|0\rangle$, thus we can either measure of just discard it.
Finally, assume A finite. Then

$$\begin{aligned} U_f &= U_f \,\mathrm{id} = U_f \sum_{k \in A} |k\rangle\langle k| = \sum_{k \in f^{-1}(0)} U_f |k\rangle\langle k| + \sum_{k \in f^{-1}(1)} U_f |k\rangle\langle k| \\ &= \sum_{k \in f^{-1}(0)} |k\rangle\langle k| - \sum_{k \in f^{-1}(1)} |k\rangle\langle k| \\ &= \sum_{k \in f^{-1}(0)} |k\rangle\langle k| - \sum_{k \in f^{-1}(1)} |k\rangle\langle k| + \sum_{k \in f^{-1}(1)} |k\rangle\langle k| - \sum_{k \in f^{-1}(1)} |k\rangle\langle k| \\ &= \left(\sum_{k \in f^{-1}(0)} |k\rangle\langle k| + \sum_{k \in f^{-1}(1)} |k\rangle\langle k| \right) - 2 \sum_{k \in f^{-1}(1)} |k\rangle\langle k| \\ &= \mathrm{id} - 2 \sum_{k \in f^{-1}(1)} |k\rangle\langle k| \end{aligned}$$

□

10.2.2 Deutsch's problem

Problem statement

Suppose we have a 1-bit to 1-bit function $f : \{0,1\} \to \{0,1\} : x \mapsto f(x)$.
Determine whether f is

- constant, i.e. $f(0) = f(1)$; or
- balanced, i.e. $f(0) \neq f(1)$.

Proposition XI.27. *The classical query complexity of Deutch's problem is 2. The quantum query complexity is 1.*

Proof. The classical case is clear.
In the quantum case, we have access to the unitary $U_f : \mathcal{H}_{\mathbb{F}_2} \to \mathcal{H}_{\mathbb{F}_2}$ from XI.26. The state $WU_fW|0\rangle$ is as follows:

$$\begin{aligned} WU_fW|0\rangle &= WU_f|+\rangle = WU_f\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \\ &= W\left(\frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}}\right) \\ &= W\begin{cases} (-1)^{f(0)}\frac{|0\rangle+|1\rangle}{\sqrt{2}} = (-1)^{f(0)}|+\rangle & (f(0)=f(1)) \\ (-1)^{f(0)}\frac{|0\rangle-|1\rangle}{\sqrt{2}} = (-1)^{f(0)}|-\rangle & (f(0)\neq f(1)) \end{cases} \\ &= \begin{cases} (-1)^{f(0)}|0\rangle & (f(0)=f(1)) \\ (-1)^{f(0)}|1\rangle & (f(0)\neq f(1)) \end{cases} \end{aligned}$$

If we measure this state, then we will certainly measure 0 is f is constant and 1 if f is balanced.
□

10.2.3 Grover's search algorithm

Problem statement

Suppose we have a set $\mathcal{N}$ of N items, some of which are marked. WLOG we can take $\mathcal{N}$ to be the set of bit strings of length ν. The marked items form a subset $\mathcal{M}$ of size M. Suppose we have an oracle that tells us whether an object is marked or not:

$$f : \mathcal{N} \to \{0,1\} : x \mapsto \begin{cases} 0 & (x \in \mathcal{M}) \\ 1 & (x \notin \mathcal{M}) \end{cases}$$

The problem is to find any marked item.

Classically we need to check $\Theta(N/M)$ items on average. Grover's algorithm has a query complexity of $O(\sqrt{N/M})$.

10.2.3.1 The algorithm

Define the states

$$|\mathcal{N}\rangle \coloneqq \frac{1}{\sqrt{N}} \sum_{n\in\mathcal{N}} |n\rangle\,, \qquad \text{and} \qquad |\mathcal{M}\rangle \coloneqq \frac{1}{\sqrt{M}} \sum_{m\in\mathcal{M}} |m\rangle\,.$$

Using XI.26, construct the unitary $U_f = \mathrm{id}_{\mathcal{H}_N} - 2\sum_{m\in\mathcal{M}} |m\rangle\langle m|$. Also construct the Householder matrix

$$U = W^{\otimes\nu}(\mathrm{id} - 2|0\rangle\langle 0|)W^{\otimes\nu} = \mathrm{id} - \mathbb{J}_{2\times 2}^{\otimes\nu} = \mathrm{id} - \mathbb{J}_{2\nu\times 2\nu} = \mathrm{id} - 2|\mathcal{N}\rangle\langle\mathcal{N}|.$$

(TODO Householder circuit).
Now consider the subspace $S \coloneqq \mathrm{span}\{|\mathcal{M}\rangle\,, |\mathcal{N}\rangle\}$. We take an orthonormal basis $\{|\mathcal{M}\rangle\,, |\mathcal{M}'\rangle\}$ of S where

$$|\mathcal{M}'\rangle \coloneqq \sqrt{\frac{N}{N-M}}\,|\mathcal{N}\rangle - \sqrt{\frac{M}{N-M}}\,|\mathcal{M}\rangle = \frac{1}{\sqrt{N-M}} \sum_{k\in\mathcal{N}\setminus\mathcal{M}} |k\rangle\,.$$

Then we can write

$$|\mathcal{N}\rangle = \sqrt{\frac{N-M}{N}}\,|\mathcal{M}'\rangle + \sqrt{\frac{M}{N}}\,|\mathcal{M}\rangle$$

and (using $\mathrm{id} = |\mathcal{M}\rangle\langle\mathcal{M}| + |\mathcal{M}'\rangle\langle\mathcal{M}'|$)

$$\begin{aligned} U_f &= \mathrm{id} - 2|\mathcal{M}\rangle\langle\mathcal{M}| = -|\mathcal{M}\rangle\langle\mathcal{M}| + |\mathcal{M}'\rangle\langle\mathcal{M}'| \\ U &= \mathrm{id} - 2\left(\sqrt{\frac{N-M}{N}}\,|\mathcal{M}'\rangle + \sqrt{\frac{M}{N}}\,|\mathcal{M}\rangle\right)\left(\sqrt{\frac{N-M}{N}}\,\langle\mathcal{M}'| + \sqrt{\frac{M}{N}}\,\langle\mathcal{M}|\right) \\ &= \frac{N-2M}{N}|\mathcal{M}\rangle\langle\mathcal{M}| + \frac{2M-N}{N}|\mathcal{M}'\rangle\langle\mathcal{M}'| - 2\left(\frac{\sqrt{M}\sqrt{N-M}}{N}\right)\left(|\mathcal{M}\rangle\langle\mathcal{M}'| + |\mathcal{M}'\rangle\langle\mathcal{M}|\right) \end{aligned}$$

So S is invariant under both U and U_f and we can write the matrix G of $-UU_f$ in this basis:

$$G = \begin{pmatrix} 1 - \frac{2M}{N} & -2\frac{\sqrt{M}\sqrt{N-M}}{N} \\ 2\frac{\sqrt{M}\sqrt{N-M}}{N} & 1 - \frac{2M}{N} \end{pmatrix}.$$

We note that $1 - \frac{2M}{N}$ is some number between -1 and 1, so we can write it as $\cos\alpha$. Then

$$\sin\alpha = \sqrt{1-\cos^2\alpha} = \sqrt{1 - \left(1 - 2\frac{M}{N}\right)^2} = \sqrt{4\frac{M}{N} - 4\frac{M^2}{N^2}} = 2\sqrt{\frac{M}{N}}\sqrt{1-\frac{M}{N}} = 2\frac{\sqrt{M}\sqrt{N-M}}{N}.$$

Thus

$$G = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = R_\alpha.$$

We now want to apply G a certain number of times, say k, such that

$$G^k\,|\mathcal{N}\rangle = G^k \begin{pmatrix} \sqrt{M/N} \\ \sqrt{1-M/N} \end{pmatrix} = R_{k\alpha} R_{\tan^{-1}(\sqrt{N/M-1})} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |\mathcal{M}\rangle\,.$$

TODO:

$$\sqrt{\frac{N}{M}} \approx T_k(\cos\alpha) - \sqrt{\frac{N}{M} - 1}\,U_{k-1}(\cos\alpha)$$

or

$$0 \approx \sqrt{\frac{N}{M} - 1} T_k(\cos\alpha) + U_{k-1}(\cos\alpha)\sin\alpha$$

and

$$(\cos\alpha + \sin\alpha)^k = T_k(\cos\alpha) + \sin\alpha U_{k-1}(\cos\alpha).$$

We also have $\cos\alpha + \sin\alpha = \frac{(\sqrt{N-M}-\sqrt{M})^2+2M}{N}$.

10.2.3.2 Analogue Grover

Consider the Hamiltonian $H = |\mathcal{M}\rangle\langle\mathcal{M}| + |\mathcal{N}\rangle\langle\mathcal{N}|$, which we can rewrite in terms of $|\mathcal{M}\rangle$ and $|\mathcal{M}'\rangle$:

$$H = \left(1 + \frac{M}{N}\right)|\mathcal{M}\rangle\langle\mathcal{M}| + \left(1 - \frac{M}{N}\right)|\mathcal{M}'\rangle\langle\mathcal{M}'| + \frac{\sqrt{M}\sqrt{N-M}}{M}\Big(|\mathcal{M}\rangle\langle\mathcal{M}'| + |\mathcal{M}'\rangle\langle\mathcal{M}|\Big),$$

This operator is self-adjoint and leaves the subspace S invariant, so S has a basis of eigenvectors for it.
By XVII.15.2, all elements of $e^{-iHt}|\mathcal{N}\rangle$ lie in S. W.r.t. the basis $\{|\mathcal{M}\rangle, |\mathcal{M}'\rangle\}$, we have

$$e^{-iHt}|\mathcal{N}\rangle = \begin{pmatrix} \sqrt{M/N}\cos(\sqrt{M/N}t) - i\sin(\sqrt{M/N}t) \\ \sqrt{1-M/N}\cos(\sqrt{M/N}t) \end{pmatrix}.$$

Thus $\langle\mathcal{M}|e^{-iHt}|\mathcal{N}\rangle = \frac{M}{N}\cos^2(\sqrt{M/N}t) + \sin^2(\sqrt{M/N}t)$. This is equal to 1 when $t = \frac{\pi}{2}\sqrt{\frac{N}{M}}$.
Thus we can produce $|\mathcal{M}\rangle$ in $O(\sqrt{N/M})$ time.
TODO: optimality.

10.2.3.3 Adiabatic Grover

$$H_0 = \mathbf{1} - |\mathcal{N}\rangle\langle\mathcal{N}| = \mathbb{1} - \mathbb{J}/N$$
$$H_f = \mathbf{1} - \sum_{m\in\mathcal{M}} |m\rangle\langle m| = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_{N-M} \end{pmatrix}.$$

Then

$$H(s) = (1-s)H_0 + sH_f = \begin{pmatrix} (1-s)\mathbb{1} + \frac{s-1}{N}\mathbb{J} & \frac{s-1}{N}\mathbb{J} \\ \frac{s-1}{N}\mathbb{J} & \mathbb{1} + \frac{s-1}{N}\mathbb{J} \end{pmatrix} = \begin{pmatrix} (1-s)\mathbb{1}_M & 0 \\ 0 & \mathbb{1}_{N-M} \end{pmatrix} + \frac{s-1}{N}\mathbb{J}.$$

Then, setting $A = \begin{pmatrix} (1-s-\lambda)\mathbb{1}_M & 0 \\ 0 & (1-\lambda)\mathbb{1}_{N-M} \end{pmatrix}$

$$\begin{aligned}
\det(H(s)-\lambda) &= \det\left(A + \frac{s-1}{N}\mathbb{J}^{n\times 1}\mathbb{J}^{1\times n}\right) \\
&= \det(A)\det\left(\mathbb{1}_N + \frac{s-1}{N}A^{-1}\mathbb{J}^{n\times 1}\mathbb{J}^{1\times n}\right) \\
&= \det(A)\det\left(1 + \frac{s-1}{N}\mathbb{J}^{1\times n}A^{-1}\mathbb{J}^{n\times 1}\right) \\
&= (1-s-\lambda)^M(1-\lambda)^{N-M}\left(1 + \frac{M(s-1)}{(1-s-\lambda)N} + \frac{(N-M)(s-1)}{(1-\lambda)N}\right) \\
&= (1-s-\lambda)^{M-1}(1-\lambda)^{N-M-1}\left(\lambda^2 - \lambda + s(1-s)\frac{N-M}{N}\right),
\end{aligned}$$

where we have used the matrix determinant lemma to simplify the calculation.
We see that there are four distinct eigenvalues:

$$\begin{aligned}
\lambda_{0,1} &= \frac{1}{2}\left(1 \pm \sqrt{1 - 4(1-\frac{M}{N})s(1-s)}\right) && \text{with multiplicity 1}\\
&= \frac{1}{2}\left(1 \pm \sqrt{\frac{M}{N} + 4(1-\frac{M}{N})(s-\frac{1}{2})^2}\right)\\
\lambda_2 &= 1 - s && \text{with multiplicity } M-1\\
\lambda_3 &= 1 && \text{with multiplicity } N-M-1.
\end{aligned}$$

Define $\Delta \coloneqq \lambda_1 - \lambda_0 = \sqrt{1 - 4(1-\frac{M}{N})s(1-s)}$. TODO: we can ignore $\lambda_{2,3}$! (Why?)

Lemma XI.28.

$$\int_0^1 \frac{1}{\Delta}\,\mathrm{d}s = \sqrt{\frac{\frac{N}{M}}{4\left(\frac{N}{M}-4\right)}}\log\left(\frac{\sqrt{\frac{N}{M}\left(\frac{N}{M}-1\right)} + \left(\frac{N}{M}-1\right)}{\sqrt{\frac{N}{M}\left(\frac{N}{M}-1\right)} - \left(\frac{N}{M}-1\right)}\right) \sim \log\left(\frac{N}{M}\right).$$

Next we are interested in the eigen vectors associated to $\lambda_{0,1}$. Let Q_1 be a unitary transformation that maps $\mathbb{J}^{M\times 1}$ to $\begin{pmatrix}\sqrt{M}\\0\\0\\\vdots\end{pmatrix} \in \mathbb{C}^M$. Similarly let Q_2 be a unitary transformation that maps $\mathbb{J}^{(N-M)\times 1}$ to $\begin{pmatrix}\sqrt{N-M}\\0\\0\\\vdots\end{pmatrix} \in \mathbb{C}^{(N-M)}$ and define $Q = \begin{pmatrix}Q_1 & 0\\0 & Q_2\end{pmatrix}$, which is also unitary.
Then we have

$$\begin{aligned}
QH(s)Q^{-1} &= \begin{pmatrix}(1-s)\mathbb{1}_M & 0\\0 & \mathbb{1}_{N-M}\end{pmatrix} + \frac{s-1}{N}Q\mathbb{J}^{N\times 1}(\mathbb{J}^{N\times 1})^*Q^*\\
&= \begin{pmatrix}(1-s)\mathbb{1}_M & 0\\0 & \mathbb{1}_{N-M}\end{pmatrix} + \frac{s-1}{N}\begin{pmatrix}\sqrt{M}\\0\\\vdots\\0\\\sqrt{N-M}\\0\\\vdots\end{pmatrix}\begin{pmatrix}\sqrt{M}\\0\\\vdots\\0\\\sqrt{N-M}\\0\\\vdots\end{pmatrix}^*.
\end{aligned}$$

All but two of the eigenvectors are just elements of $\mathcal{N}$. To study the other two we can simplify by "removing the zeros". Now the eigenvalue problem becomes

$$\left(\begin{pmatrix}1-s-\lambda_{0,1} & 0\\0 & 1-\lambda_{0,1}\end{pmatrix} + \frac{s-1}{N}\begin{pmatrix}\sqrt{M}\\\sqrt{N-M}\end{pmatrix}\begin{pmatrix}\sqrt{M}\\\sqrt{N-M}\end{pmatrix}^*\right)\mathbf{v} = 0.$$

Setting $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ gives us the equations

$$\begin{aligned} 0 &= (1 - s - \lambda_{0,1})x_1 + \frac{s-1}{N}(Mx_1 + \sqrt{M}\sqrt{N-M}x_2) \\ 0 &= (1 - \lambda_{0,1})x_2 + \frac{s-1}{N}(\sqrt{M}\sqrt{N-M}x_1 + (N-M)x_2) \end{aligned}$$

reshuffling and dividing these equations gives

$$\frac{(1 - s - \lambda_{0,1})x_1}{(1 - s - \lambda_{0,1})x_2} = \frac{-\frac{s-1}{N}(Mx_1 + \sqrt{M}\sqrt{N-M}x_2)}{-\frac{s-1}{N}(\sqrt{M}\sqrt{N-M}x_1 + (N-M)x_2)} = \frac{\sqrt{M}(\sqrt{M}x_1 + \sqrt{N-M}x_2)}{\sqrt{N-M}(\sqrt{M}x_1 + \sqrt{N-M}x_2)} = \frac{\sqrt{M}}{\sqrt{N-M}}.$$

So we have eigenvectors $\mathbf{v}_{0,1} = (\sqrt{M}(1 - \lambda_{0,1}), \sqrt{N-M}(1 - s - \lambda_{0,1}))^{\mathrm{T}}$ of $QH(s)Q^{-1}$. The corresponding eigenvectors of $H(s)$ are then given by

$$Q^* \mathbf{v}_{0,1} = \begin{pmatrix} (1 - \lambda_{0,1})\mathbb{J}^{M \times 1} \\ (1 - s - \lambda_{0,1})\mathbb{J}^{(N-M) \times 1} \end{pmatrix}.$$

For computational ease we will keep on working with the eigenvectors $\mathbf{v}_{0,1}$ of $QH(s)Q^{-1}$. We denote by $|0\rangle$ and $|1\rangle$ the normalisation of $\mathbf{v}_{0,1}$.

10.2.3.4 Poisson projective measurement

The dynamics of the system is governed by the differential equation

$$\frac{\mathrm{d}\rho}{\mathrm{d}s} = \Lambda(|0\rangle\langle 0|\rho|0\rangle\langle 0| + |1\rangle\langle 1|\rho|1\rangle\langle 1| - \rho).$$

Now we can rewrite $\frac{\mathrm{d}\rho}{\mathrm{d}t}$ in the basis $|0\rangle, |1\rangle$: let $i, j, k, l = 0, 1 \mod 2$ with implicit summation

$$\begin{aligned} \tfrac{\mathrm{d}\rho}{\mathrm{d}s} &= \tfrac{\mathrm{d}}{\mathrm{d}s}\left(|i\rangle\langle i|\rho|j\rangle\langle j|\right) \\ &= \tfrac{\mathrm{d}}{\mathrm{d}s}(\langle i|\rho|j\rangle)|i\rangle\langle j| + \langle i|\rho|j\rangle \tfrac{\mathrm{d}}{\mathrm{d}s}(|i\rangle\langle j|) \\ &= \tfrac{\mathrm{d}}{\mathrm{d}s}(\langle i|\rho|j\rangle)|i\rangle\langle j| + \langle i|\rho|j\rangle|k\rangle\langle k|\tfrac{\mathrm{d}}{\mathrm{d}s}|i\rangle\langle j| - \langle i|\rho|j\rangle|i\rangle\langle j|\tfrac{\mathrm{d}}{\mathrm{d}s}|l\rangle\langle l| \\ &= \tfrac{\mathrm{d}}{\mathrm{d}s}(\langle i|\rho|j\rangle)|i\rangle\langle j| + \langle i|\rho|j\rangle|i+1\rangle\langle i+1|\tfrac{\mathrm{d}}{\mathrm{d}s}|i\rangle\langle j| - \langle i|\rho|j\rangle|i\rangle\langle j|\tfrac{\mathrm{d}}{\mathrm{d}s}|j+1\rangle\langle j+1| \\ &= \left(\tfrac{\mathrm{d}}{\mathrm{d}s}(\langle i|\rho|j\rangle) + \langle i+1|\rho|j\rangle\langle i|\tfrac{\mathrm{d}}{\mathrm{d}s}|i+1\rangle - \langle i|\rho|j+1\rangle\langle j+1|\tfrac{\mathrm{d}}{\mathrm{d}s}|j\rangle\right)|i\rangle\langle j|. \end{aligned}$$

For each $|i\rangle\langle j|$ we get an equation, four in total. One of these is redundant by the zero trace requirement. Writing $y_{i,j} \coloneqq \langle i|\rho|j\rangle)$ and $\omega_{ij} \coloneqq \langle i|\frac{\mathrm{d}}{\mathrm{d}s}|j\rangle$ the three remaining equations are

$$\begin{aligned} \tfrac{\mathrm{d}\rho_{00}}{\mathrm{d}s} &= -\rho_{10}\omega_{01} + \rho_{01}\omega_{10} \\ \tfrac{\mathrm{d}\rho_{01}}{\mathrm{d}s} &= -\Lambda\rho_{01} - (1 - \rho_{00})\omega_{01} + \rho_{00}\omega_{01} \\ \tfrac{\mathrm{d}\rho_{10}}{\mathrm{d}s} &= -\Lambda\rho_{10} - \rho_{00}\omega_{10} + (1 - \rho_{00})\omega_{10} \end{aligned}$$

Setting $\omega = \omega_{10} = -\omega_{01}$ and $y = \rho_{00} - 1/2$we obtain the equation

$$\frac{\mathrm{d}^2 y}{\mathrm{d}s^2} = \left(\frac{\frac{\mathrm{d}\omega}{\mathrm{d}s}}{\omega} - \Lambda\right)\frac{\mathrm{d}y}{\mathrm{d}s} - 4\omega^2 y.$$

Setting $\mathbf{y} = \begin{pmatrix} y \\ y' \end{pmatrix}$ this second order differential equation is equivalent to the first order system

$$\begin{pmatrix} y \\ y' \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -4 & -\Lambda/\omega \end{pmatrix},$$

which is in the form $\mathbf{y}' = A\mathbf{y}$. Now A is similar to $\begin{pmatrix} \omega/\Lambda & 0 \\ -4 & -\Lambda/\omega \end{pmatrix}$, so it is bounded if both Λ and Λ^{-1} are bounded functions.
Then by the Picard-Lindelöf theorem the initial value problem has a unique solution on $[0, 1]$. In addition let y_1 be the solution obtained using Λ_1 and y_2 using Λ_2. Then

$$\Lambda_1 \leq \Lambda_2 \implies y_1 \leq y_2.$$

To see this, we may first remark that

$$\{t \in [0,1] \mid y_1(t) \leq y_2(t) \wedge y_1'(t) \leq y_2'(t)\} = (y_2 - y_1)^{-1}[\,[0, +\infty[\,] \cap (y_2' - y_1')^{-1}[\,[0, +\infty[\,]$$

is a closed and bounded set that contains 0. Let t_1 be its supremum. This means that $y_1(t_1) \leq y_2(t_1)$ and $y_1'(t_1) \leq y_2'(t_1)$, but $y_1(t) > y_2(t)$ or $y_1'(t) > y_2'(t)$ on some open set $]t_1, t_1 + \delta[$.

Chapter 11

Origins and justification

Now we move onto the second of Lord Kelvin's clouds, which has to do with the equipartition theorem of classical statistical mechanics. One obvious way in which the classical theory fails is when we look at thermal radiation. Object at any temperature emit electromagnetic radiation. The hotter the object, the more it emits (it becomes "glowing hot"). Unfortunately classical theory predicts that all objects emit an infinite amount of radiation of very short wavelength (i.e. ultraviolet radiation). This is clearly not true and is known as the ultraviolet catastrophe. It was this problem that Planck solved in December of 1900. The way he solved it and the assumptions he made in order to solve it had profound implications and let to the development of quantum mechanics.

As with relativity we will first go through some of the seminal experiments for quantum mechanics in order to gain a sense of the phenomenology, then based on the phenomena we will have discussed, we will develop the theory of quantum mechanics.

Quantum mechanics is quite a bewildering field that makes very little intuitive sense. The *mathematics* of quantum mechanics is well understood and has yielded many thoroughly tested predictions. At this point it is a mature theory, so it's quite surprising we still have not quite figured out what exactly it *means*. There are several competing interpretations of quantum mechanics, which we shall very briefly mention below. These interpretations do not disagree on what they predict we observe, only on what is actually happening. Thus one may surmise that they are more philosophical in nature and not necessarily relevant to the student of physics. This instrumentalist view is prevalent in many quantum mechanics courses where it is much easier, rather than wade into the murky philosophical depths of our lack of understanding, to just quote David Mermin: "Shut up and calculate".

Unfortunately for us, this multitude of post hoc interpretations poses a serious problem for our project of Carnappian rational reconstruction: which conception of reality do we base our reconstruction on and why? Recently a possible answer to this quandary has started to emerge: quantum informational axiomatisations have emerged that do not rely on our accepting any one interpretation of quantum mechanics. This is a rather new field of research and not all problems have been ironed out yet, nevertheless it is an interesting if somewhat unsatisfactory avenue and we have a quick glance down it.

Bearing this in mind, this part on quantum mechanics is structured as follows: first we have a look at some of the strange phenomena that could not be explained classically and motivated the development of quantum mechanics; at the same time we will briefly mention some of the early ad hoc theories to explain these phenomena. These theories are now sometimes referred to as the old quantum theories and still find use as semi-classical approximations. We will then

have a look at the two theories that were invented in order to explain these phenomena: matrix mechanics and wave mechanics, with an emphasis on wave mechanics because it is conceptually simpler. The introduction of these theories is still somewhat ad hoc, but they do manage to explain all of the quantum phenomena. The next step is a generalization to the elegant modern formulation of quantum mechanics using separable Hilbert spaces and a statement of the Dirac-von Neumann axioms (or postulates as they are often called). We will do our best to justify why these axioms are natural but, for the reasons mentioned above, the results of this effort will remain somewhat unsatisfactory. At this point some of the interpretations of this theory we have just constructed will be mentioned.

11.1 Wave-particle duality

11.1.1 A history of particle and wave theories of light

The nature of light has long been the subject of debate, going all the way back to the ancient Greeks. In the 11^{th} century the Arabic scientist Ibn al-Haytham (also known as Alhazen) wrote the comprehensive Book of Optics describing such phenomena as reflection and diffraction. He asserted that light rays are composed of particles of light.
In 1630 championed the wave theory of light. All other known waves had to propagate through a medium, so he posited the existence of a universal medium, the luminiferous aether.
In Opticks, published in 1704, Isaac Newton defended his corpuscular (i.e. particle) theory. He argued that only particles could move in the perfectly straight lines of light rays.
Around the same time others (including Robert Hooke, Christiaan Huygens and Augustin-Jean Fresnel) were refining the wave viewpoint.
In 1801 Thomas Young proposed his infamous double-slit experiment demonstrating wave interference of light. This, in conjunction with the inability of the particle theory to explain polerisation, began to convince the scientific community and by the middle of the 19^{th} century the wave theory was generally accepted.
In the second half of the 19^{th} century James Clerk Maxwell discovered that his equations could be applied to describe self-propagating waves of the electric and magnetic fields. It quickly became apparent that visible light, ultraviolet light and infrared light were all just electromagnetic waves of different frequencies.

11.1.2 The particle nature of photons

11.1.2.1 Black body radiation

In the section on optics (TODO) we have seen the form the spectral distribution of black body radiation takes (figure TODO). We have also already given some results based on general thermodynamic results. In order to derive an actual expression for the spectral distribution function $\rho(\lambda, T)$ we need a more detailed model.
TODO
h gives quantum of action seen as a mathematical trick by Planck trick
Wall of cavity never in equilibrium with radiation: finite degrees of freedom v infinite degrees of freedom.

11.1.2.2 The photoelectric effect

While Heinrich Hertz was performing his celebrated experiments in 1887 producing and detecting electromagnetic waves, experimentally verifying Maxwell's theory, he discovered a peculiar

phenomenon: ultraviolet light falling on electrodes facilitates the passage of a spark. It was shown that this was due to electrons being ejected from the metallic surfaces when irradiated by high-frequency electromagnetic waves. This effect is called the photoelectric effect.
Clearly the incident electromagnetic radiation is exciting the electrons in the metal enough to break free of their bonds. The photoelectric effect has a number of features that are impossible to explain with the classical theory of electromagnetic waves. Remember that in classical theory the energy of electromagnetic radiation depends on the intensity of the beam, but is independent of the frequency. (TODO:?)

1. There is a cutoff frequency below which no emission of electrons takes place, no matter the intensity of the beam.

2. The maximum kinetic energy of the emitted electrons depends linearly on frequency and is independent of the intensity.

3. Emission takes place immediately when light shines on the surface, there is no detectable time delay between exposure to light and emission of electrons. Based on classical theory we would expect to have a delay when using low intensity beams because it would take time to pump enough energy into a region of space of atomic proportions.

A very the solution to these problems was proposed in 1905 by Einstein. It was for this paper that he got the Nobel prize. Einstein proposed that we take the quantization that Max Planck had introduced seriously, and not just view it as a mathematical trick. His assertion was that light consists of quanta, now called photons, each with an energy

$$E = h\nu = hc/\lambda.$$

The strange phenomena described above can then be explained as follows:

1. A single photon needs to carry enough energy to eject an electron from the surface. (The energy required to do that is called the work function W and depends on the metal being irradiated). Thus if the frequency is too low, no electrons are emitted, no matter how many photons are pumped into the system (i.e. how high the intensity is).

2. When a photon collides with an electron, it is absorbed and passes (almost) all of its energy on to the electron, thus the maximum kinetic energy of emitted electrons is the energy of the photons ($h\nu$) minus the work function.

3. If energy is not constantly being pumped into an electron, it quickly loses the energy it gained (I assume?? TODO). Thus, because the electrons receive energy from photons in small - quantized - bursts, there is no mechanism by which energy can accumulate to a point where the electron is able to break free from the bulk metal. This means emission is either instantaneous, or does not happen.

So in conclusion, the photoelectric effect provides compelling evidence for the corpuscular nature of light. Unfortunately there is also a large amount of evidence confirming that light travels in waves. Quantum mechanics gives us a model for photons that exhibits both wave-like and particle-like behaviour. This is know as “wave-particle duality”.

11.1.2.3 The Compton effect

Another phenomenon that helps confirm the particular nature of light is the Compton effect, named after Arthur Holly Compton. This effect occurs when irradiating matter (Compton used

graphite) with X-rays. Some of the X-rays are scattered omni-directionally with no change in wavelength. This is known as Thomson scattering and can be explained classically. Sometimes the X-rays are also scattered such that the scattered wave has a different wavelength, especially if the incident X-rays have a short wavelength. Even more strangely the wavelength of the scattered wave depends on the angle at which it is scattered. This is known as Compton scattering. The difference between the wavelength of the incident wave and the scattered wave is called the Compton shift.
Thomson scattering can be explained (classically) as follows: The incident oscillating electric field causes the atomic electrons to oscillate at the same frequency. These electron then radiate electromagnetic radiation in all directions at the same frequency.
Compton scattering can be explained if we assume the incident X-rays are made up of corpuscular X-rays that scatter the atomic electrons like particles. The process is very similar to the photoelectric effect, but Compton did not build on Einstein's work. The difference is that energy of the incident photon is higher and thus the photon is not absorbed, but merely relinquishes some of its energy to the electron during the scattering process and can be detected when it comes out of the sample of matter being irradiated. In fact the Compton and photoelectric effects are two of the three competing effects when photons interact with matter; the third is pair production and is dominant at even higher energies. (These effects will be discussed in more detail later).
Because scattered can only be detected if their energy is significantly higher than the binding energy of the electrons with the atoms, we can simplify our calculations by assuming the electrons are *free*, i.e. not bound to any atoms. This explains why the expression for the Compton shift is independent of the nature of the material used for the target. The scattering probability does of course depend on the target and is in particular directly proportional to the density of electrons.
Assuming the incident photon is a particle (TODO image), we can use conservation of relativistic energy and momentum to derive the following equation:

$$\lambda_s - \lambda_0 = \Delta\lambda = \frac{2h}{m_e c} \sin^2(\theta/2)$$

where λ_0 and λ_s are the wavelength of the incident and scattered X-rays respectively; m_e is the mass of an electron.
Thomson scattering (i.e. scattering with $\lambda_s = \lambda_0$) can be explained in this model by assuming it results from scattering by electrons bound so tightly that the whole atom recoils. In this case we should use the mass M of the whole atom and the Compton shift $\Delta\lambda$ becomes negligible.

11.1.3 The wave nature of matter: the de Broglie hypothesis

In 1923 it was generally accepted that electromagnetic radiation can exhibit both wave-like and particle-like behaviour. In 1923 / 1924 Louis de Broglie proposed that matter may also have this dual nature. He postulated that the Planck-Einstein formula for the energy of a photon

$$E = h\nu$$

is not a peculiar feature of photons but holds for all particles. This means there must be some way to associate a wavelength to a particle. This wavelength is called the de Broglie wavelength in his honor and is given by

$$\lambda = \frac{h}{p}.$$

What is unique to photons is the relation $\lambda\nu = c$ making this second relation redundant and setting $p = E/c$ for photons.

11.1.3.1 Electron diffraction

If we compare with the results of classical optics, we realise that we will only be able to verify de Broglie's hypothesis by observing interference and diffraction effects if the length of some relevant part of the measurement setup is comparable to the wavelength. So in order to hope to measure electron diffraction, it is useful to have an idea of the de Broglie wavelength of an electron.
Let us assume we are producing a beam of electrons by accelerating them across a potential difference V_0, so that the kinetic energy is $mv^2/2 = eV_0$. In the non-relativistic limit, the momentum is given by $p = mv$ and thus the de Broglie wave length is given by

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2meV_0}} \approx \frac{12.3}{\sqrt{V_0[\text{Volts}]}}\text{Å}$$

This is much smaller than any length we would usually come across, which is why we had not realised matter had a wave-like nature, but is comparable to the spacing between atoms in a crystal. Using the theory developed by Bragg, Laue and others for the diffraction of X-rays by crystals (which will be developed later in the section on solid state physics), George Paget Thomson (son of J.J. Thomson, who gave his name to Thomson scattering among many other things), and Clinton Davidson and Lester Germer independently gave experimental evidence that electrons do indeed exhibit a wave-like nature.
Subsequent experiments have shown that other particle - protons, neutrons and even molecules - can exhibit wave-like behaviour.

11.1.4 Wave-packets

11.1.4.1 Form of the wave

So we seem to be able to associate something that behaves like a wave to matter.
Meaning? Spread out in space.
TODO after fluid dynamics. Decomposition into packets.

11.1.4.2 Correspondence principle

TODO image of broad wave to Dirac delta.
For almost all macroscopic physics, classical physics is obviously still a very good approximation. Classically not spread out in space, but localised. thus with vanishing wavelength. Here we may draw the parallel with geometrical optics which is the short-wavelength limit of wave optics.

11.1.4.3 The Heisenberg uncertainty principle

De Broglie's original paper.

11.1.5 Double-slit experiments

Thomas Young's double-slit experiment can neatly be summed up by figure TODO (include labels of light source and screen). The image on the screen is then something like figure TODO(b). TODO: move to optics??
Similar experiments can be used to show the wave nature of electrons and other particles. In these cases it is much easier to record individual particles, and we get an image on the screen resembling figure TODO.
With light the interference pattern can be explained by the fact that

1. The amplitudes of the light-waves can be added linearly;
2. The intensity is given by the square of the amplitude.

In the classical view, the interpretation of these statements is relatively clear: the amplitudes of the light-waves are the amplitudes of the oscillations of the electric and magnetic fields. The intensity is the power per unit area and can be measured by our eyes (TODO: see optics!!).
If we try applying this to matter that is classically thought of as consisting of particles, we run into some problems: How come we can (only) measure single electrons? Electrons invariably show up as blips on our detectors, we seem not to be able to detect the waves directly.
In light of this we may think that electrons are in essence particles and the wave-like features we measure are simply the result of interactions between electrons we do not understand yet. This turns out not to be the case. Experiments have been performed with very weak sources over the course of weeks such that there is only ever at most one electron traveling through the double-slit apparatus. These experiments yield exactly the same interference patterns.
So it seems very likely that the two aspects of classical theory that allow these interference patterns to occur have quantum analogues.

11.1.5.1 Statistical interpretation

Max Born hypothesised an interpretation

11.1.5.2 Superposition principle

11.2 Quantised quantities

TODO: Planck postulated quantisation of energy. There are several other phenomena and experiments that can be explained much more easily if we assume certain quantities to be quantised.

11.2.1 Atomic spectra

11.2.1.1 Bohr model of the hydrogen atom

Quantised energy and angular momentum

11.2.2 Franck and Hertz experiment

Quantisation of energy.

11.2.3 Stern-Gerlach experiment

Angular momentum and spin

11.3 A history of quantum theories

In 1925 / 1926 two different (but ultimately equivalent) theories were put forward to describe the properties of quantum mechanical objects, *matrix mechanics* developed by Werner Heisenberg, Max Born and Pascual Jordan, and *wave mechanics* developed by Erwin Schödinger. Both are particular forms of a more general formulation of quantum mechanics developed by Paul Adrian Maurice Dirac in 1930.

11.4 Wave mechanics

11.4.1 Schrödinger equation

$$E = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x})$$

Make the following substitutions:

$$E \to i\frac{\partial}{\partial t}, \qquad \mathbf{p} \to -i\boldsymbol{\nabla}, \qquad \mathbf{x} \to \mathbf{x}$$

And we get Schrödinger's equation

$$\boxed{i\frac{\partial \psi}{\partial t} = \left(-\frac{\nabla^2}{2m} + V(x)\right)\psi} \equiv H\psi$$

We are going to study conservative systems, so H does not depend on time.

11.4.2 Bohr condition

$$\rho(\bar{x}, t) = |\psi(x)|^2 \geq 0$$
$$P(t) = \int_{\mathbb{R}^3} \mathrm{d}^3x \rho(x,t) = \int \mathrm{d}^3x \, |\psi(\bar{x}, t)|^2 = 1$$

So $\psi \in \mathcal{C}_2$

11.4.3 Is the probability constant in time?

$$\begin{aligned}
\tfrac{\mathrm{d}P}{\mathrm{d}t} &= \int_{\mathbb{R}^3} \mathrm{d}^3x \frac{\partial}{\partial t} |\psi(x,t)|^2 = \int_{\mathbb{R}^3} \mathrm{d}^3x \left(\frac{\partial \psi^*}{\partial t}\psi + \psi^* \frac{\partial \psi}{\partial t}\right) \\
&= \frac{i}{2M} \int_{\mathbb{R}^3} \mathrm{d}^3x \left(\psi^* \left(\bar{\nabla}^2 \psi\right) - \cancel{V\psi^*\psi \cdot 2M} - \left(\bar{\nabla}^2 \psi^*\right)\psi + \cancel{V\psi^*\psi \cdot 2M}\right) \\
&= \frac{i}{2M} \int_{\mathbb{R}^3} \mathrm{d}^3x \bar{\nabla} \cdot \left(\psi^* \left(\bar{\nabla}\psi\right) - \left(\bar{\nabla}\psi^*\right)\psi\right) \\
\mathbf{j}(\bar{x}, t) &\equiv \frac{-i}{2M} \left(\psi^* \left(\bar{\nabla}\psi\right) - \left(\bar{\nabla}\psi^*\right)\psi\right)
\end{aligned}$$

So we get

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\int_{\mathbb{R}^3} \boldsymbol{\nabla} \cdot \mathbf{j} \, \mathrm{d}^3x = -\oint_{\partial \mathbb{R}^3} \mathbf{j} \cdot \mathbf{u}_j \, \mathrm{d}^2\sigma = 0$$

Where

$$\begin{cases} \mathbf{j}(\bar{\infty}, t) = 0 \\ \psi(\bar{\infty}, t) = 0 \qquad \text{because} \quad \psi \in \mathcal{C}_2 \end{cases}$$

(Assumption of conservative system important!)

11.4.4 Continuity equation

$$\frac{\partial}{\partial t} |\psi(\bar{x}, t)|^2 + \boldsymbol{\nabla} \cdot \mathbf{j} = 0$$

Conservation of probability = continuity equation

11.4.5 One-dimensional examples

11.4.5.1 The linear quantum harmonic oscillator.

We have a potential of the form

$$V(x) = \frac{1}{2}kx^2$$

so the Schrödinger equation becomes

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + \frac{1}{2}kx^2\psi(x) = E\psi(x)$$

We can immediately see that all eigenfunctions correspond to bound states of positive energy.

11.4.6 Three-dimensional examples

11.4.7 General solution of the free Schrödinger equation

$$\boxed{i\frac{\partial}{\partial t}\psi(x,t) = -\frac{\boldsymbol{\nabla}^2}{2M}\psi(x,t)}$$

11.4.7.1 Derive a basis of eigenstates

of $H = -\frac{\nabla^2}{2M}$.
Let's suppose that there is a solution of the form $\psi(x,t) = \chi(t)\varphi(\mathbf{x})$. Then

$$i\frac{\partial}{\partial t}\psi = \varphi(x)i\frac{\partial\chi(t)}{\partial t} = -\chi(t)\frac{\nabla^2}{2M}\varphi(x) = E\chi(t)\varphi(x)$$

$$i\frac{1}{\chi(t)}\frac{\partial\chi}{\partial t} = -\frac{1}{\varphi}\frac{\boldsymbol{\nabla}^2}{2M}\varphi = E = \text{const.}$$

$$\begin{cases} \frac{\partial\chi}{\partial t} = -iE\chi & \to \qquad \chi(t) = \chi(0)e^{-iEt} \\ \boldsymbol{\nabla}^2\varphi = -2ME\varphi & \to \qquad \varphi(x) = \varphi(0)e^{i\mathbf{k}\cdot\mathbf{x}} \end{cases}$$

Dispersion relation: $|\mathbf{k}|^2 = 2ME, \qquad E_k = \frac{|\mathbf{k}|^2}{2M}$

$$\psi(x,t) \sim e^{-i(E_k t - \mathbf{k}\cdot\mathbf{x})}$$

11.4.7.2 General solution of S.E.

$$\psi(\bar{x},t) = \frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3k\tilde{\psi}(k)e^{-i(E_k t-\mathbf{k}\cdot\mathbf{x})}\Big|_{E_k=\frac{|\mathbf{k}|^2}{2M}}$$

11.4.8 The Schrödinger equation is nonrelativistic

The SE contains a first order derivative in time, but a second order derivative in space.

$$\partial_\mu = (\partial_0, \boldsymbol{\nabla}) \qquad \to \qquad \partial_\mu\partial^\mu = \Box$$

11.5 Matrix mechanics

11.6 Synthesis by Dirac and von Neumann

TODO + losing physical intuition

Chapter 12

The formalism of quantum mechanics

12.1 Postulates

Based on the exposition so far, a set of axioms may be postulated that provide a purely mathematical foundation of quantum mechanics. These are known is the Dirac-von Neumann axioms. They may be formulated in many equivalent ways.

12.1.1 The space of states

At any given time, a quantum system is in a state. We postulate that the space of states, $\mathcal{H}$, is a complex Hilbert space. This means it is a complex inner product space that is also a complete metric space with respect to the norm induced by the inner product.
Additionally any two elements in the Hilbert space that are multiples of each other refer to the same state. Thus a quantum state can be seen as the equivalence class of all vectors in the Hilbert space that are (non-zero) multiples of each other. These equivalence classes are elements of the associated projective Hilbert space.
It is however enough to postulate the existence of a Hilbert space. The fact that we are really dealing with a projective Hilbert space follows from the other postulates.
So the first postulate reads:

The space of states is a (projective) Hilbert space.

12.1.1.1 The Dirac bra-ket notation

The inner product defined on the Hilbert space is conventionally written using Dirac's bra-ket notation:

$$\langle \cdot | \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$

This can be seen as the multiplication of two objects, the bra $\langle \phi |$ and the ket $|\psi\rangle$, where ψ and ϕ are arbitrary vectors in the Hilbert space.
TODO: using dual spaces. TODO: $\langle \psi | \hat{A} | \psi \rangle$

12.1.2 Measuring observables

One of the main features of quantum mechanics is its acknowledgment of the importance of the measurement process. In general the act of measuring can change the state of the system. Sometimes we can infer things about the system purely from the things it sends out, but often we need to probe the system by sending something in and seeing what comes out. The probe we send in can obviously alter the system.
Quantum mechanics also makes quite a surprising assertion:

If the same observable is measured a second time, the system does not change state a second time.

Many of the weird aspects of quantum mechanics come from this postulate, often together with the following one:

Every observable A has a linear transformation $\hat{A}$ on the Hilbert space associated to it such that the expectation value of A measured in a system in state ψ is a real number given by

$$\langle A \rangle = \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle}$$

The linear transformation $\hat{A} : \mathcal{H} \to \mathcal{H}$ is usually called the operator. The state $\hat{A}\psi$ is in general **NOT** the result of a measurement of A.
We now explore some of the consequences of these two postulates.

- **The operators are Hermitian**. The postulate states that the expectation value is real. If the numerator is real, then it must equal its complex conjugate. Since the inner product is conjugate symmetric, the following equalities hold:

  $$\langle \psi | \hat{A} | \psi \rangle = \langle \psi | \hat{A} \psi \rangle = \overline{\langle \psi | \hat{A} \psi \rangle} = \langle \hat{A} \psi | \psi \rangle$$

  which assert the very definition of a Hermitian operator.

- **Any state can be written as a sum of eigenstates**. Thanks to the spectral theorem we know that every Hermitian operator has a complete orthonormal set of eigenvectors $\{a_i\}_{i \in I}$. TODO ?

- **After a measurement of A the system is in an eigenstate of $\hat{A}$.** Assume the system was first in a state $|\psi\rangle$ and TODO

- TODO probability to find it in any eigenstate.

- TODO no need for projectivity

12.1.3 Time evolution

12.2 Important theorems

12.2.1 Existence and categoricity of quantum theories

12.2.2 Equivalent representations

Say Ψ, Ψ' equivalent representations (same probabilities), then $\Psi' = U\Psi$ with U unitary and linear:

$$(U\Psi_1, U\Psi_2) = (\Psi_1, \Psi_2)$$
$$U(c_1\Psi_1 + c_2\Psi_2) = c_1 U\Psi_1 + c_2 U\Psi_2$$

or antiunitary and antilinear:

$$(U\Psi_1, U\Psi_2) = (\Psi_1, \Psi_2)^*$$
$$U(c_1\Psi_1 + c_2\Psi_2) = c_1^* U\Psi_1 + c_2^(U\Psi_2$$

For antilinear operators we redefine adjoint $A^\dagger$.

- Linear:
$$(\Psi_1, L^\dagger\Psi_2) \equiv (L\Psi_1, \Psi_2)$$
- Antilinear:
$$(\Psi_1, A^\dagger\Psi_2) \equiv (A\Psi_1, \Psi_2)^* = (\Psi_2, A\Psi_1)$$

Then the conditions of unitarity and antiunitarity are both expressed as

$$U^\dagger = U^{-1}$$

In particular the identity operator $U = 1$ is unitary and linear.

12.2.3 Symmetries

TODO see Weinberg
Operator on ray vs operator on state.
Changing phase. Also superselection.
Connected Lie groups must be represented in the physical Hilbert space by unitary (rather than antiunitary) operators $U(T(\theta))$

12.3 A reconstruction of quantum mechanics via quantum information

12.4 The Hamiltonian

TODO Minimal coupling principle

12.5 Angular momentum

12.6 Pictures

Say we have a quantum mechanical system with the following initial conditions:

$$\begin{cases} |\psi(t_0)\rangle = |\psi_0\rangle \\ A(t_0) = A_0 \end{cases}$$

where A is an operator

12.6.1 Schrödinger picture

In the Schrödinger picture the time dependence is carried by the states, according to the Schrödinger equation.

$$\begin{cases} |\psi(t)\rangle_s = U(t,t_0)\,|\psi(t_0)\rangle_s = U(t,t_0)\,|psi_0\rangle \\ A_s(t) = A_s(t_0) = A_0 \end{cases}$$

(For an arbitrary operator A (?))

$$\textbf{The Schrödinger eq: } i\tfrac{\mathrm{d}}{\mathrm{d}t}\,|\psi(t)\rangle_s = H\,|\psi(t)\rangle_s$$

So the evolution operator $U(t,t_0)$

$$\boxed{i\tfrac{\mathrm{d}U}{\mathrm{d}t}(t,t_0) = HU(t,t_0)}$$

$$\begin{cases} U(t_0,t_0) = \mathbb{1} \\ U(t,t_0) = e^{-iH(t-t_0)} \end{cases} \qquad \rightarrow \qquad [U(t,t_0),H] = 0$$

12.6.2 Heisenberg picture

$$\begin{cases} |\psi(t)\rangle_H = |\psi(t_0)\rangle_H = |\psi_0\rangle \\ A_H(t) = U^\dagger(t,t_0)A_H(t_0)U(t,t_0) = U^\dagger(t,t_0)A_0U(t,t_0) \end{cases}$$

Which is equivalent with requiring

$$i\frac{\mathrm{d}}{\mathrm{d}t}A_H(t) = i\left(\tfrac{\mathrm{d}U^\dagger}{\mathrm{d}t}A_0U + U^\dagger A_0\tfrac{\mathrm{d}U}{\mathrm{d}t}\right) = [A_H(t),H]$$

Schrödinger and Heisenberg are equivalent because related by a unitary transformation.

$$\begin{cases} |\psi(t)\rangle_H = H^\dagger(t,t_0)\,|\psi(t)\rangle_s \\ A_H(t) = U^\dagger(t,t_0)A_s(t)U(t,t_0) \end{cases}$$

$$\boxed{\langle\psi|A|\psi\rangle_H = \langle\psi|A|\psi\rangle_S}$$

So still same physics

12.6.3 The interaction picture

Assume that the Hamiltonian is time-independent and can be expressed as the sum

$$H = H_0 + H_I$$

where H_0 acting alone yields a soluble problem.
We define the state vector in the interaction picture as

$$|\Psi_I(t)\rangle \equiv e^{iH_0 t/\hbar} |\Psi_S(t)\rangle$$

We can do this because $e^{iH_0 t/\hbar}$ is a unitary transformation carried out at time t.
An arbitrary matrix element in the Schrödinger picture can be written as

$$\langle \Psi'_S(t)|\hat{O}_S|\Psi_S(t)\rangle = \langle \Psi'_I(t)|e^{iH_0 t/\hbar}\hat{O}_S e^{-iH_0 t/\hbar}|\Psi_S(t)\rangle$$

for an arbitrary operator $\hat{O}_S$ in the Schrödinger picture. This suggests defining the operator $\hat{O}_I$ in the interaction picture as

$$\hat{O}_I(t) \equiv e^{iH_0 t/\hbar}\hat{O}_S e^{-iH_0 t/\hbar}$$

12.6.4 The interaction picture

$$H = H_0 + H_{\text{int}}$$

$$\begin{cases} U_0(t,t_0) = e^{-iH_0(t-t_0)} \\ U(t,t_0) = e^{-iH(t-t_0)} \end{cases}$$

The picture is defined (in terms of Schrödinger picture, but that is arbitrary) by

$$\begin{cases} |\psi(t)\rangle_I = U_0^\dagger(t,0)\,|\psi(t)\rangle_s \\ A_I(t) = U_0^\dagger(t,0) A_s(t) U_0(t,0) \end{cases} \qquad \rightarrow \qquad \text{interaction picture}$$

Now we write the evolution eq. of $|\psi\rangle, A$ in interaction picture.

$$\begin{aligned} i\tfrac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle_I &= i\tfrac{\mathrm{d}}{\mathrm{d}t}\left(U_0^\dagger(t,0)U(t,t_0)\right)|\psi_0\rangle \\ &= i\left(U_0^\dagger(t,0)(H-H_0)U(t,t_0)\right)|\psi_0\rangle \\ &= U_0^\dagger(t,0)H_{\text{int}}\,|\psi(t)\rangle_s \\ &= \underbrace{U_0^\dagger H_{\text{int}} U_0(t,0)}_{\equiv H^I_{\text{int}}(t)}|\psi(t)\rangle_I \end{aligned}$$

Evolution equation of $|\psi(t)\rangle_I$

$$\boxed{i\tfrac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle_I = H^I_{\text{int}}(t)\,|\psi(t)\rangle_I}$$

In general

$$\left[H^I_{\text{int}}(t), H\right] \neq 0 \neq \left[H^I_{\text{int}}(t), U(t,t_0)\right]$$

12.6.4.1 Evolution of operators in the interaction picture

$$\begin{aligned} i\tfrac{\mathrm{d}A_I(t)}{\mathrm{d}t}(t) &= i\tfrac{\mathrm{d}}{\mathrm{d}t}\left(U_0^\dagger(t,0)A_s(t_0)U_0(t,0)\right) \\ &= -U_0^\dagger(t,0)H_0A_0U_0(t,0) + U_0^\dagger(t,0)A_0H_0U_0(t,0) \\ &= \left[U_0^\dagger(t,0)A_0U_0(t,0), H_0\right] \equiv [A_I(t), H_0] \end{aligned}$$

Evolution of operators in interaction picture

$$\boxed{i\tfrac{\mathrm{d}}{\mathrm{d}t}A_I(t) = [A_I(t), H_0]}$$

Now consider $\varphi_I(t)$. Then

$$i\frac{\mathrm{d}}{\mathrm{d}t}\varphi_I(t) = [\varphi_I(t), H_0]$$

Which is the equation for the free theory in the Heisenberg picture. So the field in the interaction picture satisfies the free field equation.

$$\begin{aligned} \left(\Box + m^2\right)\varphi_H(t) &= -\lambda\varphi_H^3(t) \qquad \rightarrow \qquad \text{Heisenberg picture} \\ \left(\Box + m^2\right)\varphi_I(t) &= 0 \qquad \rightarrow \qquad \varphi_I \sim ae^{-ikx} + a^\dagger e^{ikx} \end{aligned}$$

Operators evolve like free theory. Now all of the complication is in evolution of states.

12.6.4.2 States are not constant

$$\begin{aligned} |\psi(t)\rangle_I &= U_0^\dagger(t,0)U(t,t_0)\,|\psi(t_0)\rangle_s \\ &= U_0^\dagger(t,0)U(t,t_0)U_0(t_0)\,|\psi(t_0)\rangle_I \\ &\equiv U_I(t,t_0)\,|\psi(t_0)\rangle_I \end{aligned}$$

So

$$\begin{aligned} U_I(t,t_0) &\equiv U_0^\dagger U(t,t_0)U_0(t_0,0) \\ &= e^{iH_0t}e^{-iH(t-t_0)}e^{-iH_0t_0} \end{aligned}$$

In general

$$[H, H_0] \neq 0 \neq [H_0, H_{\text{int}}]$$

So

$$i\frac{\mathrm{d}}{\mathrm{d}t}U_I(t,t_0) = H_{\text{int}}^I(t)U_I(t,t_0)$$

$$\begin{cases} U_I(t_0,t_0) = \mathbb{1} \\ U_I(t,t_0) = \mathbb{1} - i\int_{t_0}^t \mathrm{d}\tau H_{\text{int}}^I(\tau)U_I(\tau,t_0) \end{cases}$$

12.6.4.3 Perturbative solution of U_I

Let's suppose that $H_{\text{int}} \propto \lambda$ with $\lambda \ll 1$.

$$U_I(t,t_0) = \sum_{u=0}^{\infty} C_u(t,t_0) \qquad \to \qquad C_u \sim \lambda^u$$

$U_I^{(N)}$ is the approx solution up λ^N

$$U_I(t,t_0) = U_I^{(N)} + (O)(\lambda^{N+1})$$

$$U_I^{(N)}(t,t_0) = \mathbb{1} - i\int_{t_0}^{t} \mathrm{d}\tau H_{\text{int}}^I(\tau) U_I^{(N)}(\tau,t_0)$$

But H_{int}^I depends on λ, so the right side is of order $N+1$, so we actually have

$$U_I^{(N)}(t,t_0) = \mathbb{1} - i\int_{t_0}^{t} \mathrm{d}\tau H_{\text{int}}^I(\tau) U_I^{(N-1)}(\tau,t_0)$$

This is a recursive relation. If we know the solution at $\mathcal{O}(N-1)$, we can obtain the solution at $\mathcal{O}(N)$

1. Solution of order 0 ($\lambda = 0, H_{\text{int}} = 0$)

$$U_I^{(0)}(t,t_0) = C_0 = \mathbb{1}$$

2. Solution of order 1

$$U_I^{(1)}(t,t_0) = \mathbb{1} - i\int_{t_0}^{t} \mathrm{d}\tau H_{\text{int}}^I(\tau)\cdot\mathbb{1} = C_0 + C_1$$

3. Second order term

$$\begin{aligned} U_I^{(2)}(t,t_0) &= \mathbb{1} - i\int_{t_0}^{t} \mathrm{d}\tau H_{\text{int}}^I(\tau) U_I^{(1)}(\tau,t_0) \\ &= \mathbb{1} - i\int_{t_0}^{t} \mathrm{d}\tau H_{\text{int}}^I(\tau)\left(\mathbb{1} - i\int_{t_0}^{t} \mathrm{d}\tau_1 H_{\text{int}}^I(\tau_1)\right) \\ &= \mathbb{1} - i\int_{t_0}^{t} \mathrm{d}\tau H_{\text{int}}^I(\tau) + (-i)^2\int_{t_0}^{t} \mathrm{d}\tau \int_{t_0}^{\tau} \mathrm{d}\tau_1 H_{\text{int}}^I(\tau) H_{\text{int}}^I(\tau_1) \\ &= C_0 + C_1 + C_2 \end{aligned}$$

So in general the u^{th} order term in the expansion is given by

$$C_u = (-i)^u \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{t_0}^{\tau_1} \mathrm{d}\tau_2 \ldots \int_{t_0}^{\tau_{u-1}} \mathrm{d}\tau_u H_{\text{int}}^I(\tau_1)\ldots H_{\text{int}}^I(\tau_u)$$

And we have

$$U_I^{(N)}(t,t_0) = \sum_{u=0}^{N} C_u \qquad U_I(t,t_0) = \sum_{u=0}^{\infty} C_u$$

12.6.4.4 Put it in a better formalism

$$C_2(t,t_0) = (-i)^2 \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{t_0}^{\tau_1} \mathrm{d}\tau_2 H^I_{\mathrm{int}}(\tau_1) H^I_{\mathrm{int}}(\tau_2)$$

[Picture horizontal lines, plot t1 vs. t2]

$$C_2(t,t_0) = (-i)^2 \int_{t_0}^{t} \mathrm{d}\tau_2 \int_{\tau_2}^{t} \mathrm{d}\tau_1 H^I_{\mathrm{int}}(\tau_1) H^I_{\mathrm{int}}(\tau_2)$$

[Picture vertical lines t2 vs. t1]

$$C_2(t,t_0) = (-i)^2 \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{\tau_1}^{t} \mathrm{d}\tau_2 H^I_{\mathrm{int}}(\tau_2) H^I_{\mathrm{int}}(\tau_1)$$

I can define

$$C_2 = \frac{(-i)^2}{2} \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{\tau_1}^{t} \mathrm{d}\tau_2 H^I_{\mathrm{int}}(\tau_2) H^I_{\mathrm{int}}(\tau_1) + \frac{(-i)^2}{2} \int_{t_0}^{t} \mathrm{d}\tau_2 \int_{\tau_2}^{t} \mathrm{d}\tau_1 H^I_{\mathrm{int}}(\tau_1) H^I_{\mathrm{int}}(\tau_2)$$

$$\begin{aligned} C_2 &= \frac{(-i)^2}{2!} \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{t_0}^{t} \mathrm{d}\tau_2 \left\{ \theta(\tau_1 - \tau_2) H^I_{\mathrm{int}}(\tau_1) H^I_{\mathrm{int}}(\tau_2) + \theta(\tau_1 - \tau_2) H^I_{\mathrm{int}}(\tau_2) H^I_{\mathrm{int}}(\tau_1) \right\} \\ &= \frac{(-i)^2}{2!} \int_{t_0}^{t} \mathrm{d}\tau_1 \, \mathrm{d}\tau_2 T \left(H^I_{\mathrm{int}}(\tau_1) H^I_{\mathrm{int}}(\tau_2) \right) \end{aligned}$$

Lecture 14/11 Interaction picture

$$\begin{cases} i \frac{\mathrm{d}}{\mathrm{d}t} A_I(t) = [A_I(t), H_0] & \rightarrow \qquad \text{free eq.} \\ i \frac{\mathrm{d}}{\mathrm{d}t} \left|\psi(t)\right\rangle_I = H^I_{\mathrm{int}}(t) \left|\psi(t)\right\rangle_I \end{cases}$$

$$\left|\psi(t)\right\rangle_I = U_I(t,t_0) \left|\psi(t_0)\right\rangle_I$$

$$i \frac{\mathrm{d}U_I(t)}{\mathrm{d}t} = H^I_{\mathrm{int}} U_I(t,t_0)$$

Would be a simple differential equation if it commuted. Solution:

$$\begin{aligned} U_I(t,t_0) &= T \left(\exp -i \int_{t_0}^{t} \mathrm{d}\tau H^I_{\mathrm{int}}(\tau) \right) \\ &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{t_0}^{t} \mathrm{d}\tau_1 \ldots \mathrm{d}\tau_n T \left(H^I_{\mathrm{int}}(\tau_1) \ldots H^I_{\mathrm{int}}(\tau_n) \right) \end{aligned}$$

12.7 Path integral formulation

12.8 Other axiomatizations

12.8.1 Dirac - von Neumann axioms in terms of a C^*-algebra

Chapter 13

Measurement

https://arxiv.org/pdf/0911.2539.pdf https://arxiv.org/pdf/1708.00769.pdf https://chaos.if.uj.edu.pl/~karol/geometry.htm

13.1 Adiabatic theorem

https://arxiv.org/pdf/2003.03063.pdf https://link.springer.com/content/pdf/10.1134/1.1352758.pdf https://arxiv.org/pdf/quant-ph/0603175.pdf

13.2 Projective measurement

13.3 Generalised measurement

13.4 Continuous measurement

https://arxiv.org/pdf/quant-ph/0611067.pdf Book: Quantum Trajectories and Measurements in Continuous Time

13.4.1 Quantum zeno

https://www.fi.muni.cz/usr/buzek/zaujimave/home.pdf https://iopscience.iop.org/article/10.1088/1742-6596/196/1/012018/pdf

Chapter 14

Quantum statistical mechanics

We will be considering a quantum mechanical system of N particles. In the Schrödinger picture and the position representation, the system is described by a wave function $\Psi(q_1, q_2, \ldots, q_N, t)$ which satisfies the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(q_1, q_2, \ldots, q_N, t) = H\Psi(q_1, q_2, \ldots, q_N, t)$$

14.1 Bosons and fermions

14.1.1 Totally symmetric and antisymmetric wave functions

14.1.2 Pauli exclusion principle

14.2 Fermi gas

TODO fermi temperature

Chapter 15

Approximation methods

15.1 For stationary problems

15.2 For time-dependent problems

Chapter 16

Interactions of quantum systems

Chapter 17

Relativistic wave equations

In this section we follow the mostly minus convention

$$\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1).$$

We will also be setting c and $\hbar$ to 1.

17.1 Klein-Gordon equation

17.1.1 Derivation

We start from the relativistic equation

$$p^\mu p_\mu = m^2 = E^2 - |\mathbf{p}|^2$$

which we rewrite as

$$E^2 = m^2 + |\mathbf{p}|^2 .$$

We quantise as usual by substituting the classical quantities with the usual operators:

$$-\frac{\partial^2}{\partial t^2}\psi = \left(-\boldsymbol{\nabla}^2 + m^2\right)\psi$$

This is equivalent to writing

$$\left[\left(\partial_0^2 - \boldsymbol{\nabla}^2\right) + m^2\right]\psi = 0 = \left[\partial_\mu \partial^\mu + m^2\right]\psi$$

and so we get the famous Klein-Gordon equation

$$\boxed{\left(\Box + M^2\right)\psi = 0}.$$

This equation is manifestly covariant. All the components of the equation are scalars. TODO rest of section in chosen notation
$\psi(x)$ is a scalar function

$$\psi'(x') \qquad \rightarrow \qquad \psi(x)$$

And M and $\Box$ are scalars

$$\begin{cases} M & \rightarrow \quad M' = M \\ \Box & \rightarrow \quad \Box' = \Box \end{cases}$$

So the Klein-Gordon equation transforms in the following way:

$$\left(\Box + M^2\right)\psi(x) = 0 \qquad \overset{\text{Poincaré}}{\rightarrow} \qquad \left(\Box' + M'^2\right)\psi'(x) = \left(\Box + M^2\right)\psi(x) = 0$$

17.1.2 The continuity equation

Any function satisfying the Klein-Gordon equation must also satisfy a continuity equation

$$\partial_0 \rho + \boldsymbol{\nabla} \mathbf{j} = 0$$

In order to prove this, we start by asserting the Klein-Gordon equation holds for both ψ and its complex conjugate ψ^*. We can then multiply both sides of the equations with either ψ or ψ^* to get

$$\begin{cases} \psi^* \left(\Box + m^2\right) \psi = 0 \\ \psi \left(\Box + m^2\right) \psi^* = 0 \end{cases}$$

Subtracting these two equations gives

$$\begin{aligned} 0 &= \psi^* \Box \psi + \cancel{m^2 \psi^* \psi} - \psi \Box \psi^* - \cancel{m^2 \psi \psi^*} \\ &= \psi^* (\Box \psi) - \psi (\Box \psi^*) \end{aligned}$$

Integrating by parts and expanding ∂_μ as $(\partial_0, -\boldsymbol{\nabla})$ gives

$$\begin{aligned} 0 &= \partial_0 \left[\psi^* (\partial_0 \psi) - (\partial_0 \psi^*) \psi\right] + \cancel{(\partial_0 \psi)(\partial_0 \psi^*)} - \cancel{(\partial_0 \psi^*)(\partial_0 \psi)} - \boldsymbol{\nabla} \left[\psi^* (\boldsymbol{\nabla} \psi) - (\boldsymbol{\nabla} \psi^*) \psi\right] - \cancel{(\boldsymbol{\nabla} \psi)(\boldsymbol{\nabla} \psi^*)} + \cancel{(\boldsymbol{\nabla} \psi^*)(\boldsymbol{\nabla} \psi)} \\ &= \partial_0 \left[\psi^* (\partial_0 \psi) - (\partial_0 \psi^*) \psi\right] - \boldsymbol{\nabla} \left[\psi^* (\boldsymbol{\nabla} \psi) - (\boldsymbol{\nabla} \psi^*) \psi\right] \end{aligned}$$

Defining the quantities

$$\begin{cases} \rho \equiv \frac{1}{2} \left[\psi^* (\partial_0 \psi) - (\partial_0 \psi^*) \psi\right] \equiv \frac{i}{2} \psi^* (\overset{\leftrightarrow}{\partial}_0) \psi \\ \mathbf{j} = -\frac{1}{2} \left[\psi^* (\boldsymbol{\nabla} \psi) - (\boldsymbol{\nabla} \psi^*) \psi\right] \equiv -\frac{2}{i} \psi^* \overset{\leftrightarrow}{\boldsymbol{\nabla}} \psi \end{cases}$$

where we have also introduced the notation $\overset{\leftrightarrow}{\partial}_0$ and $\overset{\leftrightarrow}{\boldsymbol{\nabla}}$, we see that we have derived a continuity equation

$$\boxed{\partial_0 \rho + \boldsymbol{\nabla} \mathbf{j} = 0}$$

Which can also be written as

$$\partial_\mu j^\mu = 0$$

where $j^\mu = (\rho, \mathbf{j})$. For any continuity equation there is a conserved quantity. In casu

$$Q = \int_{\mathbb{R}^3} \mathrm{d}^3 x \rho(x, t)$$

is conserved:

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \int_{\mathbb{R}^3} \mathrm{d}^3 x \frac{\partial}{\partial t} \rho(x, t) = -\int_{\mathbb{R}^3} \mathrm{d}^3 x \, \boldsymbol{\nabla} \cdot \mathbf{j} = -\oint_{\partial \mathbb{R}^3} \mathrm{d}^2 \sigma \mathbf{j} \cdot \mathbf{u}_\sigma = 0$$

TODO redo with chosen notation Because ρ can be positive or negative, we cannot associate Q to a probability.

17.1.3 General solution

The goal will be to write the solutions to the Klein-Gordon equation in a more illuminating form. A scalar solution $\varphi(x)$ can be written as the Fourier transform of $\tilde{\varphi}(k)$.

$$\varphi(x) = \frac{1}{(2\pi)^2} \int \mathrm{d}^4 k e^{-i k^\mu x_\mu} \tilde{\varphi}(k)$$

Conversely $\tilde{\varphi}(k)$ can be written as

$$\tilde{\varphi}(k) = \frac{1}{(2\pi)^2}\int \mathrm{d}^4 k e^{ik^\mu x_\mu}\varphi(x)$$

This is still a scalar function because d^4k and $k^\mu x_\mu$ are invariant and $\varphi(x)$ was required to be scalar.
Now we require that $\varphi(x)$ is a solution of the Klein-Gordon equation

$$(\Box + m^2)\varphi(x) = \int \frac{\mathrm{d}^4k}{(2\pi)^2}(-k^2+m^2)e^{-ik^\mu x_\mu}\tilde{\varphi}(k) = 0$$

This is equivalent to requiring

$$\int_{k^2\neq m^2} \frac{\mathrm{d}^4k}{(2\pi)^2} e^{-ik^\mu x_\mu}\tilde{\varphi}(k) = 0.$$

Using this we can rewrite the expression for $\varphi(x)$

$$\begin{aligned}
\varphi(x) &= \frac{1}{(2\pi)^2}\int \mathrm{d}^4 k e^{-ik^\mu x_\mu}\tilde{\varphi}(k) \\
&= \frac{1}{(2\pi)^2}\int_{k^2\neq m^2} \mathrm{d}^4 k e^{-ik^\mu x_\mu}\tilde{\varphi}(k) + \frac{1}{(2\pi)^2}\int \mathrm{d}^4 k e^{-ik^\mu x_\mu}\delta(k^2-m^2)\tilde{\varphi}(k) \\
&= \frac{1}{(2\pi)^2}\int \mathrm{d}^4 k e^{-ik^\mu x_\mu}\delta(k^2-m^2)\tilde{\varphi}(k)
\end{aligned}$$

Using the uniqueness of the Fourier transform, we can write $\tilde{\varphi}(k)$ as

$$\tilde{\varphi}(k) = \delta(k^2-m^2)\tilde{f}(k)$$

for a function $\tilde{f}(k)$.
The next step is to rewrite the delta function.

$$\begin{aligned}
\delta(k^2-m^2) &= \delta(k_0^2 - |\mathbf{k}|^2 - m^2) \\
&= \delta(k_0^2-\omega_k^2) = \frac{\delta(k_0-\omega_k)+\delta(k_0+\omega_k)}{2\omega_k}
\end{aligned}$$

Where we have defined $\omega_k \equiv (m^2+|\mathbf{k}|^2)^{1/2}$. This is the relativistic energy. The notation (writing it as a frequency) makes sense because we have set $\hbar = 1$.
The last equality follows from the result of composing the δ-function with a smooth, continuously differentiable function g:

$$\delta(g(x)) = \sum_i \frac{\delta(x-x_i)}{|g'(x_i)|}$$

where x_i are the simple roots of g. In this case of course k_0 is the variable and $\pm\omega_k$ are the roots.
Plugging all of this back into the original equation yields

$$\begin{aligned}
\varphi(x) &= \frac{1}{(2\pi)^2}\int \frac{\mathrm{d}^4k}{2\omega_k}\left(e^{-ik^\mu x_\mu}\delta(k_0-\omega_k)\tilde{f}(k) + e^{-ik^\mu x_\mu}\delta(k_0+\omega_k)\tilde{f}(k)\right) \\
&= \frac{1}{(2\pi)^2}\int \frac{\mathrm{d}^3k}{2\omega_k}\left(e^{-i(\omega_k x_0 - \mathbf{k}\cdot\mathbf{x})}\tilde{f}(\omega_k,\mathbf{k}) + e^{i(\omega_k x_0+\mathbf{k}\cdot\mathbf{x})}\tilde{f}(-\omega_k,\mathbf{k})\right)
\end{aligned}$$

Seeing as we are integrating over all **k**s anyway, We can apply the transformation $\mathbf{k} \to -\mathbf{k}$

$$\varphi(x) = \frac{1}{(2\pi)^2} \int \frac{\mathrm{d}^3 k}{2\omega_k} \left(e^{-i(\omega_k x_0 + \mathbf{k}\cdot\mathbf{x})} \tilde{f}(\omega_k, -\mathbf{k}) + e^{i(\omega_k x_0 - \mathbf{k}\cdot\mathbf{x})} \tilde{f}(-\omega_k, -\mathbf{k}) \right)$$

We now define the following functions:

$$\begin{cases} a(\mathbf{k}) = \frac{\tilde{f}(\omega_k, -\mathbf{k})}{\sqrt{(2\pi)(2\omega_k)}} \\ b^*(\mathbf{k}) = \frac{\tilde{f}(-\omega_k, -\mathbf{k})}{\sqrt{(2\pi)(2\omega_k)}} \end{cases}$$

so we get the following general solution of the Klein-Gordon equation:

$$\varphi(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \left(a(k) e^{-ikx} + b^*(k) e^{ikx} \right)_{k_0 = \omega_k}$$

Some remarks about this solution:

1. The function $\varphi(x)$ is a scalar, even if it does not look like one. We started with a scalar function and only performed algebraic manipulations.
2. We associate $a(k)e^{-ikx}$ with positive energy and $b^*(k)e^{ikx}$ with negative energy, so we can perform the following decomposition:

$$\varphi(x) = \varphi_+(x) + \varphi_-(x) \sim e^{-ikx} + e^{ikx}$$

3. We can further illustrate the association of $\varphi_\pm$ with energies:

$$\begin{aligned} i\partial_0 \varphi_+ &\approx i\partial_0 (e^{-ikx} a(k)) = +\omega_k (e^{-ikx} a(k)) \\ i\partial_0 \varphi_- &\approx i\partial_0 (e^{ikx} b^*(k)) = -\omega_k (e^{ikx} b^*(k)) \end{aligned}$$

4. We solved the Klein-Gordon equation generally for complex φ. If we want that φ be real ($\varphi(x) = \varphi^*(x)$), we require the following:

$$\begin{aligned} \varphi(x)^* &= \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \left(a^*(k) e^{ikx} + b(k) e^{-ikx} \right)_{k_0 = \omega_k} \\ &= \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \left(a(k) e^{-ikx} + b^*(k) e^{ikx} \right)_{k_0 = \omega_k} = \varphi(x) \end{aligned}$$

   which means that $a(k) = b(k)$ and $a^*(k) = b^*(k)$, so that

$$\varphi_{\mathbb{R}}(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \left(a(k) e^{-ikx} + a^*(k) e^{ikx} \right)_{k_0 = \omega_k}$$

17.1.3.1 Expressions for $a(k)$ and $a^*(k)$

In a real field.

$$\begin{aligned} a(p) &= \int \frac{\mathrm{d}^3 x}{\sqrt{2\omega_p}} \left(i\partial_0 \varphi + \omega_p \varphi \right) \Big|_{p_0 = \omega_p} \\ a^*(p) &= \int \frac{\mathrm{d}^3 x}{\sqrt{2\omega_p}} \left(-i\partial_0 \varphi + \omega_p \varphi \right) e^{-ipx} \Big|_{p_0 = \omega_p} \end{aligned}$$

In a complex field. The complex case is totally analogous TODO

17.1.4 Interpretation of the Klein-Gordon equation

TODO

Complex scalar solution of KG eq. can be associated to a CHARGED (EM) SPIN 0 particle

17.1.4.1 Coupling to an external electromagnetic field

External (given, fixed) EM field

$$A^\mu = (A^0, \bar{A})$$

$$\begin{cases} \bar{E} = -\bar{\nabla} A_0 - \partial_0 \bar{A} \\ \bar{B} = \bar{\nabla} \times \bar{A} \end{cases}$$

The minimal coupling with an (external) EM field A^μ is obtained by the following substitution:

$$\partial_\mu \to D_\mu \equiv (\partial_m u + iqA_\mu)$$

We call D_μ the covariant derivative

The minimal coupling

$$P_\mu \to P_\mu - qA_\mu \qquad \begin{cases} \omega_p \to \omega_p - qA_0 \\ \bar{p} \to \bar{p} - q\bar{A} \end{cases}$$

S.Eq. for non relativistic particle:

$$i\frac{\partial \psi}{\partial t} = \frac{P^2}{2M}\psi \to i\frac{\partial \psi}{\partial t} = \left(\frac{(\bar{p} - q\bar{A})^2}{2M} + qA_0\right)\psi$$

Coupled Klein-Gordon equation:

$$(\partial_\mu \partial^\mu)\varphi = 0 \to (D_\mu D^\mu + M^2)\varphi = 0$$

$$(D_\mu D^\mu + M^2)\varphi = \left[(\partial_0 + iqA_0)^2 - (\bar{\nabla} - iq\bar{A})^2 + M^2\right]\varphi = 0$$

Note: $(\partial_i + iqA_i)(\partial^i + iqA^i) \to (\bar{\nabla} - iq\bar{A})(-\bar{\nabla} + iq\bar{A})$ (spatial part of $D^\mu D_\mu$)

$$(D^\mu D_\mu + M^2)\varphi = [(\partial_\mu + iqA_\mu)(\partial^\mu + iqA^\mu) + M^2]\varphi$$

$$[\Box + 2iqA^\mu \partial_\mu - q^2 A^2 + iq(\partial_\mu A^\mu) + M^2]\varphi = 0$$

Lorentz gauge: $\partial_\mu A^\mu = 0$.
$\varphi = \varphi_+ + \varphi_-$

$$\begin{cases} D^\mu D_\mu \varphi_+ \approx [(\omega_k - qA_0)^2 - (\bar{k} - q\bar{A})^2]\varphi_+ = M^2\varphi_+ \\ D^\mu D_\mu \varphi_- \approx [(\omega_k + qA_0)^2 - (\bar{k} + q\bar{A})^2]\varphi_- = M^2\varphi_- \end{cases}$$

$$\begin{pmatrix} \varphi_+ \\ q \end{pmatrix} \leftrightarrow \begin{pmatrix} \varphi_- \\ -q \end{pmatrix}$$

Physical meaning:

$$\begin{cases}\varphi_+(E>0) \to \text{particle} \\ \varphi_-(E<0) \to \text{anti-particle}\end{cases}$$

Take non-relativistic limit of coupled Klein-Gordon equation

$$E = (M^2 + |\bar{p}|^2)^{1/2} = M(1 + \frac{|p|}{M^2})^{1/2} \approx M + \frac{\bar{p}^2}{2M} + \dots$$

Non-relativistic limit: $E \approx M$ or $M \gg |\bar{p}|$, $M >> E_k = \frac{\bar{p}^2}{2M}$.
Do following redefinition

$$\varphi(x,t) = e^{i\omega_p t} = e^{-iMt}\varphi'(x,t) = e^{-iMt}e^{-iE_k t}$$

$$[(\partial_0 + iqA_0)^2 + M^2]e^{-iMt}\varphi' = e^{-iMt}[(\bar{\nabla} - iq\bar{A})^2]\varphi'$$

$$e^{-iMt}[\partial_0 - M2 - 2iM\partial_0 + 2iqA_0\partial_0 + 2iq(\partial_0 A^0) + 2qMA_0 - q^2 + M^2]\varphi' = e^{-iMt}[(\bar{\nabla} - iq\bar{A})]\varphi'$$

In the non-relativistic limit:

$$\left|\frac{\partial_0\varphi'}{\varphi'}\right| \ll M, \qquad |qA_0| << M, \qquad \left|\frac{\partial_0 A_0}{A_0} << M\right|$$

Then $2iM\partial_0$ and $2qMA_0$ dominant (single time derivative)

$$i\partial_0\varphi' = \left[-\frac{1}{2M}(\bar{\nabla} - iq\bar{A})^2 + qA_0\right]\varphi'$$

$\to$ SL ofcoupled EM

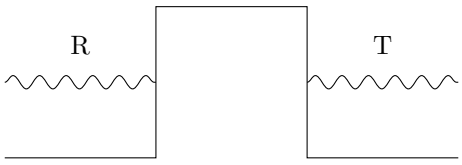
17.1.4.2 Problems with interpreting the equation

Two pajor problems with interpretation of probability:

1. Positive and negative energies:

$$E^2 = M^2 + |P|^2 \qquad \to \qquad \omega_p = \pm\left(M^2 + |P|^2\right)^{1/2}$$

2. Interpretation of probabilities. Take a quantum barrier.



Continuity equation $\int \mathrm{d}^3x\rho(x,t) >< 0$

Klein paradox:

$$R + T = 1 \qquad \begin{cases}R < 0 \\ T > 1\end{cases}$$

17.2 Dirac equation

17.2.1 The search for a linear equation

The Klein-Gordon equation has some problems. It is also quadratic, not linear like the Schrödinger equation. This prompted Dirac to look for a linear relativistic wave equation. Starting from the relativistic energy equation, we can try to factorise it like this

$$p^\mu p_\mu - m^2 = (a^\mu p_\mu + m)(b^\lambda p_\lambda - m) = 0$$

with a^μ and b^λ some constants. This is satisfied if

$$(a^\mu p_\mu + m) = a_0 E - \mathbf{a}\cdot\mathbf{p} + m = 0.$$

Rewriting in terms of E we get

$$E = \boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m$$

With $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and β still just some constants. Applying the quantisation recipe, we then get

$$i\frac{\partial}{\partial t}\psi = (-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\psi = H\psi$$

We have two requirements for this equation. These impose restrictions on $\boldsymbol{\alpha}$ and β.

1. H is a hermitian operator;
2. The energy E must still satisfy $E^2 = m^2 + |\mathbf{p}|^2$. In other words the Dirac equation has to be consistent with the Klein-Gordon equation.

So enforcing those, we get:

1. $H^\dagger = H$ implies $\alpha_i^\dagger = \alpha_i$ and $\beta^\dagger = \beta$
2. For consistency with the Klein-Gordon equation we assert that $\frac{\partial^2\psi}{\partial t^2} = (\boldsymbol{\nabla}^2 - m^2)\psi$:

$$\begin{aligned}
\frac{\partial^2\psi}{\partial t^2} &= -i\frac{\partial}{\partial t}\left(i\frac{\partial\psi}{\partial t}\right) = -i\frac{\partial}{\partial t}\left(-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m\right)\psi \\
&= -\left(-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m\right)\left(-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m\right)\psi \\
&= \sum_{i,j}\left(\alpha_i\alpha_j\partial_i\partial_j + i(\alpha_i\beta + \beta\alpha_i)m\partial_i - \beta^2 m\right)\psi \\
&= (\boldsymbol{\nabla}^2 - m^2)\psi
\end{aligned}$$

Comparing the last equality term by term, it imposes three conditions:

(a) The term

$$\begin{aligned}\sum_{i,j}\alpha_i\alpha_j\partial_i\partial_j &= \frac{1}{2}\left[\sum_{i,j}\alpha_i\alpha_j\partial_i\partial_j + \sum_{i,j}\alpha_i\alpha_j\partial_i\partial_j\right]\\ &= \frac{1}{2}\left[\sum_{i,j}\alpha_i\alpha_j\partial_i\partial_j + \sum_{i,j}\alpha_j\alpha_i\partial_j\partial_i\right]\\ &= \frac{1}{2}\left[\sum_{i,j}\alpha_i\alpha_j\partial_i\partial_j + \sum_{i,j}\alpha_j\alpha_i\partial_i\partial_j\right]\\ &= \sum_{i,j}\frac{1}{2}\left(\alpha_i\alpha_j+\alpha_j\alpha_i\right)\partial_i\partial_j\end{aligned}$$

must equal

$$\boldsymbol{\nabla}^2 = \sum_i \partial_i^2 = \sum_{i,j}\delta_{ij}\partial_i\partial_j$$

So consequently

$$\frac{1}{2}\left(\alpha_i\alpha_j+\alpha_j\alpha_i\right) = \frac{1}{2}\{\alpha_i,\alpha_j\} = \delta_{ij}$$

(b) The term $(\alpha_i\beta+\beta\alpha_i)$ must vanish. In other words

$$\{\alpha_i,\beta\} = 0$$

(c) Lastly we have the condition

$$\beta^2 = 1$$

It quite quickly becomes apparent that no real numbers α_i, β can satisfy these conditions. While this is a setback, this does not quite doom our project yet. We may try making α_i and β matrices. This would make ψ a column-vector. We have not yet considered how such objects might transform, so we are hesitant to use the word vector, instead we call them (Dirac) spinors. They are elements of an n dimensional spinorial space.
Now is a good time to explore what kind of properties the matrices α_i, β and the spinorial space have:

- We have been considering all possible products of two of the matrices α_i, β. For this to make any sense (i.e. if we want both $\alpha_i\alpha_j$ and $\alpha_j\alpha_i$ to make sense), we need all the matrices to be square with the same dimensions. They are all $n \times n$ matrices.
- Because we want the Hamiltonian to be Hermitian, we have already required that α_i, β be Hermitian.
- Because β is Hermitian, we know it must be unitarily diagonalisable:

  $$\beta^2 = (UDU^*)^2 = UD^2U^* = \mathbb{1}$$

  where D is a diagonal matrix with the eigenvalues of β on the diagonal. From this equality we see that $D^2 = \mathbb{1}$ and thus the eigenvalues of β must be 1 an -1. The same holds true for the α_is.

- All of the matrices anti-commute, so for $i \neq j$

$$\mathrm{Tr}(\alpha_i) = \mathrm{Tr}(\alpha_i \alpha_j^2) = \mathrm{Tr}(\alpha_i \alpha_j \alpha_j) = -\mathrm{Tr}(\alpha_j \alpha_i \alpha_j) = -\mathrm{Tr}(\alpha_i)$$

  where the last equality holds because the trace is cyclic. Consequently

$$\mathrm{Tr}(\alpha_i) = 0 = \mathrm{Tr}(\beta)$$

- When diagonalising, the trace of the diagonal matrix is the same as that of the original one, in this case zero. Because there can only be 1 and -1 on the diagonal, there must be an equal number of 1s and -1s, so the dimension n must be even.

Now the question is, can we actually find matrices that satisfy these conditions? Searching in the lowest even dimension $n = 2$, we realise that we can always only find three linearly independent traceless Hermitian matrices (for example the Pauli matrices σ_i). We need four linearly independent matrices to satisfy the anti-commutator relations, so we will need to try a higher dimension.

> On a side-note, if we assume the mass m is zero, we no longer need the β matrix and then there is a two dimensional spinorial representation. This is the Weyl equation.

For $n = 4$ we have no problem finding matrices that fulfill all the requirements:

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$$

So the lowest possible spinorial dimension for the Dirac spinors is $n = 4$.

17.2.1.1 Dirac γ-matrices in the Dirac / Pauli representation

TODO why called representation?
Based on our choice of α_i and β, we define the γ-matrices.

$$\gamma^0 \equiv \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^i \equiv \beta\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}$$

We can put these γ-matrices together in a four-vector-like notation

$$\gamma^\mu \equiv (\gamma^0, \gamma^i).$$

But it is important to remember that these matrices were introduced as constants, are still constants and thus **do not transform under Poincaré transformations** like four-vectors do (TODO ??????).
The properties of the matrices α_i, β translate to the following properties of the γ-matrices:

1. $\gamma^0 = (\gamma^0)^\dagger, \qquad \gamma^i = -(\gamma^i)^\dagger$
2. $(\gamma^0)^2 = \mathbb{1}, \qquad (\gamma^i)^2 = -\mathbb{1}$
3. $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$

In fact any matrices that satisfy these properties would work. Almost all results will be derived from these properties, so that they remain valid any representation.
In order to rewrite the Dirac equation in terms of γ-matrices, we multiply by β:

$$i\beta\partial_0\psi = \left(-i\beta\alpha_i\partial_i + \beta^2 m\right)\psi = \left(-i\gamma^i\partial_i + m\right)\psi$$

Rearranging we get

$$\left(i\gamma^0 + i\gamma^i\partial_i - m\right)\psi\left(i\gamma^\mu\partial_\mu - m\right)\psi = 0$$

This can be written even more succinctly with the slashed notation

$$\not{a} \equiv \gamma^\mu a_\mu = \gamma^0 a_0 + \gamma^i a_i$$

for any four-vector a^μ. Using this notation, we get the most famous form of the Dirac equation:

$$\boxed{(i\not{\partial} - m)\psi = 0}$$

The Dirac equation in fact gives four equations with components

$$\sum_\beta \left(i\gamma^\mu_{\alpha\beta}\partial_\mu - M\delta_{\alpha\beta}\right)\psi_\beta = 0$$

where α and β are spinorial indices ranging from one to four and μ is a Lorentz index.
TODO: calculus of spinor indices

17.2.1.2 Other representations of γ-matrices

The choice of the γ representation is not unique. With a non-singular unitary matrix C, one can always obtain a new representation

$$\tilde{\gamma}^\mu = C^{-1}\gamma^\mu C$$

17.2.2 Covariance of Dirac equation

Having posited the Dirac equation, we now want to verify the covariance of the equation.
TODO theoretical grounds for transformation of spinors (coordinate based construction / basis??? Essential essence of spinors)
Considering an arbitrary Lorentz transformation $\Lambda \in \mathrm{SO}(1,3)$:

$$x^\mu \qquad \rightarrow \qquad x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

which, by construction, gives

$$\partial_\mu \qquad \rightarrow \qquad \partial'_\mu = \Lambda_\mu{}^\nu \partial_\nu.$$

Clearly $\psi(x)$ is impacted by the coordinate transformation and becomes $\psi'(x')$. We may not just assume that each component of the spinor is a scalar. So in general $\psi(x) \neq \psi'(x')$, but we may write

$$\psi(x) \qquad \rightarrow \qquad \psi'(x') = S(\Lambda)\psi(x).$$

Some properties of $S(\Lambda)$ are immediately apparent (TODO):

1. Action

2. Linear map, so representation.

So $S(\Lambda)$ is an element of the spinorial representation of Lorentz group. The question is now which one.
We want the Dirac equation to be covariant, meaning it still holds under a Lorentz transformation

$$(i\gamma^\mu\partial_\mu - M)\psi = 0 \qquad \overset{\Lambda}{\to} \qquad (i\gamma^\mu\partial'_\mu - M)\psi'(x') = 0$$

where we have used that γ^μ does not transform in Λ.
Filling in the transformed quantities, we get

$$\begin{aligned} 0 &= (i\gamma^\mu {\Lambda_\mu}^\nu \partial_\nu - M)S(\Lambda)\psi(x) \\ &= S(\Lambda)[i{\Lambda_\mu}^\nu S(\Lambda)^{-1}\gamma^\mu S(\Lambda)\partial_\mu - M]\psi \end{aligned}$$

this holds if

$${\Lambda_\mu}^\nu S(\Lambda)^{-1}\gamma^\mu S(\Lambda) = \gamma^\mu$$

or, equivalently

$$S(\Lambda)^{-1}\gamma^\mu S(\Lambda) = {\Lambda^\mu}_\nu \gamma^\mu.$$

17.2.2.1 Spinorial ($N = 4$) representation of the Lorentz algebra

$$\begin{cases} {\Lambda^\mu}_\nu = \delta^\mu_\nu + {\omega^\mu}_\nu \\ S(\Lambda) = \mathbb{1} - \frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu} \end{cases}$$

($\Sigma^{\mu\nu}$ are generators of representation)
Using the "covariance" relation

$$\left(\mathbb{1} + \frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}\right)\gamma^\mu\left(\mathbb{1} - \frac{i}{2}\omega_{\rho\sigma}\Sigma^{\rho\sigma}\right) = \gamma^\mu + {\omega^\mu}_\nu\gamma^\nu$$

$$\gamma^\mu - \frac{i}{2}\omega_{\rho\sigma}[\gamma^\mu, \Sigma^{\rho\sigma}] + \mathcal{O}(\omega^2) = \gamma^\mu + \omega_{\rho\sigma}(\eta^{\mu\rho}\gamma^\sigma)$$

$$[\gamma^\mu, \Sigma^{\rho\sigma}] = i)(\eta^{\mu\rho}\gamma^\sigma - \eta^{\mu\sigma}\gamma^\rho)$$

So $\Sigma^{\rho\sigma}$ are the following matrices

$$\Sigma^{\mu\nu} \equiv \frac{i}{4}[\gamma^\mu, \gamma^\nu] \equiv \frac{1}{2}\sigma^{\mu\nu}$$

Example

Exercise: verify

$$[\gamma^\mu, \gamma^\rho\gamma^\sigma] = \{\gamma^\mu, \gamma^\rho\}\gamma^\sigma - \gamma^\rho\{\gamma^\mu, \gamma^\sigma\}$$

Example

Exercise: Verify $\Sigma^{\rho\sigma}$ satisfy the commutator algebra of $\mathrm{SO}(1,3)$.

$\Sigma^{\mu\nu}$ is the ($D = 4$) spinorial representation of $\mathrm{SO}(1,3)$

Lecture 16/10
ix) Definition of γ^5

$$\gamma^5 \equiv -\frac{i}{4}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma = i\gamma^0\gamma^1\gamma^2\gamma^3$$

Properties:

- $\left(\gamma^5\right)^2 = \mathbb{1}_4$
- $\gamma^5 = \left(\gamma^5\right)^\dagger$
- $\{\gamma^\mu, \gamma^5\} = 0$

Note: THe γ^5 matrix satisfies

$$[\gamma^5, \Sigma^{\mu\nu}] = 0$$

The $D = 4$ representation of SO(1.3) is reducible

Chirality projectors :

$$P_L \equiv \left(\frac{1-\gamma_T}{2}\right), \qquad P_R \equiv \left(\frac{1+\gamma_T}{2}\right)$$

Chiral components of ψ:

$$\psi_L = P_L\psi, \qquad \psi_R = P_R\psi$$

(What are projectors ????) Prove that:

$$\gamma^5\psi_L = -\psi_L, \qquad \gamma^5\psi_R = +\psi_R$$

ψ_L, ψ_R are (the) irreducible????

$$\psi \sim \psi_L \oplus \psi_R \sim (\tfrac{1}{2}, 0) \oplus (0, \tfrac{1}{2})$$

γ^5 matrix in Dirac / Pauli representation

$$\gamma^5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$$

Spin and helicity operators in ($D = 4$):

$$K_i \equiv \Sigma^{i0} = -\frac{i}{2}\begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}$$

$$\Sigma^i \equiv \frac{1}{2}\epsilon_{ijk}\Sigma^{jk} = \frac{1}{2}\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}$$

Spin is not a good quantum number, because it does not commute with Hamiltonian.

$$[H_D, \Sigma_3] = i(\alpha_1 P_2 - \alpha_2 P_1) = i\epsilon_{ij3}\alpha_i P_j \neq 0$$

Helicity is projection of spin operator in direction of momentum.

The helicity operator

$$\sigma_P \equiv \frac{\bar{\Sigma}\cdot\bar{P}}{|\bar{P}|}$$

The helicity is a good quantum number:

$$[H_D, \sigma_P] = i\epsilon_{ijk}\alpha_i P_j P_k = 0$$

17.2.3 Pauli-Lubansky vector and helicity

$$\omega^\mu \equiv \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}J_{\nu\rho}P_\sigma = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\Sigma_{\nu\rho}P_\sigma$$

Where $J_{\nu\rho} = L_{\nu\rho} + \Sigma_{\nu\rho}$ and $L_{\nu\rho} = X_\nu P_\rho - X_\rho P_\nu$
Now let's calculate ω^μ for $D = 4$ spinorial representation in the rest frame $P^\mu = (M, 0)$

$$\omega^\mu \overset{R.F.}{=} \frac{M}{2}\epsilon^{\mu\nu\rho\sigma}\Sigma_{\nu\rho} \to \begin{cases}\omega^0 = 0\\ \omega^i = M\Sigma^i\end{cases}$$

$$\begin{aligned}\frac{\omega^\mu\omega_\mu}{M^2} \overset{R.F.}{=} &= \Sigma^i\Sigma_i = -|\bar{\Sigma}|^2\\ &= -\left(\Sigma_1^2 + \Sigma_2^2 + \Sigma_3^2\right)^2 = -\frac{3}{4}\mathbb{1}_4\\ &= -s(s+1)4 \qquad \text{with} s = \tfrac{1}{2}\end{aligned}$$

So ω^2 is the Casimir of the Poincaré group. Now the connection between ω^μ and σ_P

$$n_P^\mu \equiv \left(\frac{|\bar{P}|}{M}, \frac{\omega_P}{M}\frac{\bar{P}}{|\bar{P}|}\right)$$

$$\frac{\omega^\mu n_\mu^P}{M} = -\frac{1}{2}\sigma_P = -\frac{1}{2}\begin{pmatrix}\frac{\bar{\sigma}\cdot\bar{P}}{|\bar{P}|} & 0\\ 0 & \frac{\bar{\sigma}\cdot\bar{P}}{|\bar{P}|}\end{pmatrix} = -\frac{1}{2}\frac{\bar{\Sigma}\cdot\bar{P}}{|\bar{P}|}$$

17.2.4 Interpretation of the Dirac equation

17.2.4.1 Dirac conjugate spinor

$$0 = [(i\partial\!\!\!/ - M)\psi]^\dagger = -\psi^\dagger(i(\gamma^\mu)^\dagger\overleftarrow{\partial}_\mu + M)$$

Using $\gamma^0(\gamma^\mu)^\dagger\gamma^0 = \gamma^\mu$

$$= -\psi^\dagger(i\gamma^0\gamma^\mu\gamma^0\overleftarrow{\partial}_\mu + M) = -(\psi^\dagger\gamma^0)\left(i\gamma^\mu\overleftarrow{\partial}_\mu + M\right)\gamma^0$$

$$\bar{\psi}\left(i\overleftarrow{\partial\!\!\!/} + M\right) = 0$$

With $\bar{\psi} \equiv \psi^\dagger\gamma^0$ the Dirac conjugate.
So

$$\psi'(x') = S(\Lambda)\psi(x)$$

$$\bar{\psi}'(x') = (\psi')^\dagger(x')\gamma^0 = \psi^\dagger(x)\left(\gamma^0\right)^2 S^\dagger(\Lambda)\gamma^0 = \bar{\psi}\gamma^0 S^\dagger(\Lambda)\gamma^0 = \bar{\psi}(x)S(\Lambda)^{-1}$$

17.2.4.2 Continuity equation

$$\bar{\psi}(i\vec{\not\partial} - M)\psi = 0$$
$$\bar{\psi}(i\vec{\not\partial} + M)\psi = 0$$

$$\begin{aligned} 0 &= \bar{\psi}\left(i\vec{\not\partial} - M + i\overleftarrow{\not\partial} + M\right)\psi \\ &= \bar{\psi}\left(i\gamma^\mu\vec{\partial}_\mu + i\gamma^\mu\overleftarrow{\partial}_\mu\right)\psi \\ &= i\left[\bar{\psi}\gamma^\mu(\partial_\mu\psi) + (\partial_\mu\bar{\psi})\gamma^\mu\psi\right] \\ &= i\partial_\mu[\bar{\psi}\gamma^\mu\psi] = 0 \end{aligned}$$

Continuity equation for Dirac:

$$\partial_\mu J^\mu = 0$$

With

$$\begin{aligned} J^\mu &\equiv \bar{\psi}\gamma^\mu\psi \\ &= (\psi\bar{\gamma^0}\psi, \psi\bar{\gamma^i}\psi) \\ &= (\psi^\dagger\psi, \psi^\dagger\partial_i\psi) \\ &= (\rho, \bar{j}) \end{aligned}$$

So we get

$$\partial_0\rho + \partial_i j^i = 0$$

Now define

$$\Theta = \int \mathrm{d}^3x\rho(x,t) \qquad \rightarrow \qquad \frac{\mathrm{d}\Theta}{\mathrm{d}t} = 0$$

So Θ is conserved.

$$Q = \int \mathrm{d}^3x\psi^\dagger\psi \qquad > 0$$

17.2.4.3 Problems with the interpretation

PROBLEM: $Q > 0, H_D >< 0$ Klein paradox (see above) So we still have all the problems of KG.
Physical objects $\bar{\psi}\Gamma\psi$ are bilinear.
How do bilinear transform under Lorentz?

$$\bar{\psi}\mathbb{1}\psi \xrightarrow{\Lambda} \bar{\psi}'\psi' = \bar{\psi}S(\Lambda)^{-1}S(\Lambda)\psi$$

So scalar.

$$\begin{aligned} J^\mu = \bar{\psi}\gamma^\mu\psi \xrightarrow{\Lambda} \bar{\psi}'\gamma^\mu\psi' &= \bar{\psi}S(\Lambda)^{-1}\gamma^\mu S(\Lambda)\psi \\ &= \Lambda^\mu_\nu\bar{\psi}\gamma^\nu\psi \\ &= \Lambda^\mu_\nu J^\nu \end{aligned}$$

17.2.5 General solution of the free Dirac equation

$$(i\partial\!\!\!/ - M)\psi(x) = 0$$
$$(\Box + M^2)\psi(x) = 0 \qquad (k^2 = M^2, \quad k_0 = \omega_k)$$

A general solution of the Klein-Gordon equation

$$\psi = \psi_+ + \psi_- = e^{-ikx}u(k) + e^{ikx}v(k)\big|_{k_0=\omega_k}$$

Now

$$(i\partial\!\!\!/ - M)\psi_+ \approx e^{-ikx}(k\!\!\!/ - M)u(k) = 0$$
$$(i\partial\!\!\!/ - M)\psi_- \approx e^{ikx}(k\!\!\!/ + M)v(k) = 0$$

$$\begin{cases}(k\!\!\!/ - M)u(k) = 0\\(k\!\!\!/ + M)v(k) = 0\end{cases} \qquad \text{Dirac eq in momentum space}$$

Find solution in the rest frame: $k^\mu \overset{R.F.}{=} (M, 0)$ (Mass different from 0)

$$\begin{cases}(k\!\!\!/ - M)u(k) \overset{R.F.}{\to} M(\gamma^0 - \mathbb{1}_4)u(M) = 0\\(k\!\!\!/ + M)v(k) \overset{R.F.}{\to} M(\gamma^0 + \mathbb{1}_4)v(M) = 0\end{cases}$$

In Dirac / Pauli representation

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

$$\begin{cases}M\left[\begin{pmatrix}\mathbb{1} & 0\\0 & -\mathbb{1}\end{pmatrix} - \begin{pmatrix}\mathbb{1} & 0\\0 & \mathbb{1}\end{pmatrix}\right]u(M) = M\begin{pmatrix}0 & 0\\0 & -2\end{pmatrix}u(M) = 0\\M\left[\begin{pmatrix}\mathbb{1} & 0\\0 & -\mathbb{1}\end{pmatrix} + \begin{pmatrix}\mathbb{1} & 0\\0 & \mathbb{1}\end{pmatrix}\right]v(M) = M\begin{pmatrix}2 & 0\\0 & 0\end{pmatrix}v(M) = 0\end{cases}$$

So we get

$$u_i(M) \equiv \sqrt{2M}\begin{pmatrix}\xi_i\\0\end{pmatrix}$$
$$v_i(M) \equiv \sqrt{2M}\begin{pmatrix}0\\\xi_i\end{pmatrix}$$

With $\xi_1 = \begin{pmatrix}1\\0\end{pmatrix}$ and $\xi_2 = \begin{pmatrix}0\\1\end{pmatrix}$.
Usually not useful to expand fully into 4 components.
Solution in a general frame.
Notice that

$$(k\!\!\!/ - M)(k\!\!\!/ + M) = \left(k\!\!\!/^2 + Mk\!\!\!/ - Mk\!\!\!/ + M^2\right) = (k^2 - M^2) = 0$$

We can therefore write

$$\begin{cases}u(k) = c(k\!\!\!/ + m)u(m)\\u(k) = c(k\!\!\!/ - m)v(m)\end{cases}$$

Using the boost operation explicitly, we can see that actually this is proportional to $k_i u_r$
Normalisation

$$\begin{cases} \bar{u}_r(k)u_s(k) = 2M\delta_{rs} \\ \bar{v}_r(k)v_s(k) = -2M\delta_{rs} \end{cases}$$

General solution in momentum space:

$$u_r(k) = \frac{(\not{k} + M)}{\sqrt{2M(M_{\omega_k} + \omega_k)}} u_r(M) = \sqrt{M + \omega_k} \begin{pmatrix} \xi_r \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{k}}{M+\omega_k}\xi_r \end{pmatrix}$$
$$v_r(k) = \frac{(-\not{k} + M)}{\sqrt{2M(M_{\omega_k} + \omega_k)}} v_r(M) = \sqrt{M + \omega_k} \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{k}}{M+\omega_k}\xi_r \\ \xi_r \end{pmatrix}$$

General solution in coordinate space:

$$\begin{aligned} \psi(x,t) &= \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k\varphi_R}} \sum_{r=1}^{2} \left(c_r(k)u_r(k)e^{-ikx} + d_r^*(k)v_r(k)e^{ikx} \right)_{k_0=\omega_k} \\ &= \bar{\psi}(x,t) \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k\varphi_R}} \sum_{r=1}^{2} \left(d_r(k)\bar{v}_r(k)e^{-ikx} + c_r^*(k)\bar{u}_r(k)e^{ikx} \right)_{k_0=\omega_k} \end{aligned}$$

Be careful manipulating the blockmatrices:
Lecture 17/10/18

Example

Dirac conjugate:

$$\bar{u}_r(k) = \bar{u}_r(M) \frac{\not{k} + M}{\sqrt{2M(M + \omega_k)}} = \left((M + \omega_k)^{1/2} \xi_1^\intercal, \dots \right)$$

Missed rest.??

Watch out: $\bar{\psi}\psi\gamma^\mu$ has no meaning.

$$\Pi_\pm \equiv \frac{\mathbb{1} \pm \sigma_p}{2} = \frac{\mathbb{1} \mp \omega^\mu n_\mu^1}{2}$$

With $W_0 = 0$, $W^i = \Sigma^i$,$\bar{p} = (0, 0, \bar{p})$.

$$\begin{cases} \pi_+ u_1(M) = +u_1(M) \\ \pi_- u_1(M) = 0 \\ \pi_+ u_2(M) = 0 \\ \pi_- u_2(M) = +u_2(M) \end{cases}$$

$$\begin{cases} \pi_+ v_1(M) = 0 \\ \pi_- v_1(M) = -v_1(M) \\ \pi_+ v_2(M) = +v_1(M) \\ \pi_- v_2(M) = 0 \end{cases}$$

Projectors over $\pm$ energy

$$\Lambda_\pm(k) = \pm\frac{\not{k} + M}{2M} \qquad \begin{pmatrix}\Lambda_+ v(k) = 0\\ \Lambda_- u(k) = 0\end{pmatrix}$$

$\Lambda_\pm$ are projectors (verify def!)
I can rewrite the projectors as

$$\Lambda_+(k) = \sum_{r=1}^{2} \frac{u_r(k) u_1(k)}{2M}$$

$$\Lambda_-(k) = -\sum_{r=1}^{2} \frac{v_r(k) v_1(k)}{2M}$$

(?)

17.2.6 Dirac equation coupled with an electromagnetic field

minimal coupling prescription

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$$

So

$$(i\not{\partial} - M)\psi = 0 \to (i\not{D} - M)\psi = 0$$

$$\to \psi_+ \text{and} \psi_- \qquad \begin{cases}\psi_+ & \text{particle of } (+q)\\ \psi_- & \text{antiparticle of } (-q)\end{cases}$$

Now we study the nonrelativistic limit of the Dirac equation

$$(i\partial_0 - qA_0)\psi = \left[-i\bar{\alpha}(\bar{\nabla} - iq\bar{A}) + \beta M\right]\psi$$

$$\left[\psi(x,t) = e^{-iMt}\psi'(x,t)\right]$$

$$e^{-iMt}\left(-\partial_0 + M - qA_0\right)\psi' = e^{-iMt}\left(-\bar{\alpha}(\bar{\nabla} - iq\bar{A}) + \beta M\right)\psi'$$

We write the 4 dim ψ' as

$$\psi' = \begin{pmatrix}\varphi'\\ \chi'\end{pmatrix}$$

So we get two 2 dim equations

$$\begin{cases} i\partial\varphi' = qA_0\varphi' - i\bar{\sigma}\cdot\left(\bar{\nabla} - iq\bar{A}\right)\chi' \\ i\partial\chi' = (qA_0 - 2M)\chi' - i\bar{\sigma}\cdot\left(\bar{\nabla} - iq\bar{A}\right)\varphi' \end{cases}$$

In nonrelativistic limit one can assume

$$\left|\frac{\partial_0\chi'}{\chi'}\right| \ll M, \qquad |qA_0| \ll M$$

The second equation becomes

$$\chi' = -\frac{i}{2M}\bar{\sigma}\cdot\left(\bar{\nabla} - iq\bar{A}\right)\varphi' \ll \varphi' \qquad \text{CONSTRAINT}$$

Then we get from the first equation

$$i\partial_0\varphi' = qA_0\varphi' - \frac{1}{2M}\left(\bar{\sigma}\cdot\bar{\nabla} - iq\bar{\sigma}\cdot\bar{A}\right)^2\varphi'$$

Which is the Pauli equation. We now work towards the Schrödinger equation.

$$\begin{aligned}
\left(\bar{\sigma}\cdot\bar{\nabla} - iq\bar{\sigma}\cdot\bar{A}\right)^2\varphi' &= \sigma_i\sigma_j(\partial_i + iqA_i)(\partial_j + iqA_j)\varphi' \\
&= \left(\frac{1}{2}\{\sigma_i,\sigma_j\} + \frac{1}{2}[\sigma_i,\sigma_j]\right)(\partial_i + iqA_i)(\partial_j + iqA_j)\varphi' \\
&= \left(\frac{1}{2}\{\sigma_i,\sigma_j\} + \frac{1}{2}[\sigma_i,\sigma_j]\right)\left(\partial_i\partial_j - q^2A_iA_j + iqA_j\partial_j + iqA_j\partial_i + iq(\partial_iA_j)\right)\varphi' \\
&= \left(\nabla^2 - q^2\bar{A}^2 - 2iq\bar{A}\cdot\bar{\nabla} - iq(\bar{\nabla}\cdot\bar{A}) + q\bar{\sigma}\cdot\bar{B}\right)\varphi' \\
&= \left[\left(\bar{\nabla} - iq\bar{A}\right)^2 + q\bar{\sigma}\cdot\bar{B}\right]\varphi'
\end{aligned}$$

$$iq\epsilon_{ijk}(\partial_iA_j)\sigma_k = -q(\bar{\nabla}\times\bar{A})_k\sigma_k = +q\bar{B}\cdot\bar{\sigma}$$

$$i\frac{\partial\varphi'}{\partial t} = \left\{-\frac{1}{2M}\left(\bar{\nabla} - iq\bar{A}\right)^2 + qA_0 - \frac{q}{2M}\bar{\sigma}\cdot\bar{B}\right\}\varphi'$$

$$i\frac{\partial\varphi'}{\partial t} = \left\{\frac{1}{2M}\left(\bar{P} - q\bar{A}\right)^2 + qA_0 - \frac{q}{2M}\bar{\sigma}\cdot\bar{B}\right\}\varphi'$$

Dipole term $\bar{\mu}_s = -\frac{q}{2M}\bar{\sigma}$

$$H_{\mathrm{DIP}} = -\bar{\mu}_s\cdot\bar{B} = -\frac{q}{2M}\bar{\sigma}\cdot\bar{B}$$

For an orbital momentum

$$\bar{\mu}_L = \frac{q}{2M}\bar{L} \qquad \rightarrow \qquad \left|\frac{\bar{\mu}_L}{\bar{L}}\right| = \frac{q}{2M}$$

$$\bar{\mu}_S = \frac{q}{2M}g_c\left(\frac{\bar{\sigma}}{2}\right) \qquad \rightarrow \qquad \left|\frac{\bar{\mu}_L}{\bar{\Sigma}_{(2)}}\right| = \frac{q}{2M}g_c$$

With $\bar{\Sigma}_{(2)} = \frac{\bar{\sigma}}{2}$.
With g_c is the gyromagnetic faction for electrons (spin 1/2).
So Dirac equation predicts $g_c = 2$. (Dirac 1938)

$$g_c^{\mathrm{exp}} = 1.99 \pm 0.2$$

Chapter 18

Measurement and interpretation

Part XII

Quantum field theory

Chapter 1

Roadmap

In our quest for a theory of everything we have come a long way. In the previous section we studied some wave equations that are both quantum mechanical in nature *and* consistent with special relativity.
TODO: Weinberg: QFT is the way it is because (with certain qualifications) this is the only way to reconcile quantum mechanics with special relativity.
There were however some problems with the interpretation of these equations, such as the prediction of states with negative energies. Also standard quantum mechanics does not allow for transmutations of particles into other particles. We know such precesses happen, so we need a theory to explain them. As an added bonus quantum field theory makes it much easier to deal with many particle systems.
Quantum field theory gives us a new way to conceive of, and thus (mathematically model), reality. We are inspired by the photon: in the section on electromagnetism we described electromagnetic waves as propagating fluctuations in the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$. When writing the laws of electrodynamics in a manifestly relativistic form, we saw that the potential A^μ satisfied the wave equation and thus photons could be seen as waves in the 4-vector field. Later we saw that in some cases it made sense to view these waves as particles. De Broglie then made the bold statement that all particles could be seen as waves. This led to the development of wave mechanics.
In quantum field theory we take the analogy one step further: might it be possible to describe other particles as waves propagating through a field of some kind? It turns out that we can, and this way of looking at things provides the theoretical framework for the whole of the standard model.
Within this framework we can formulate many theories. Some may describe events that may actually occur. This can be compared to the frameworks of Newtonian mechanics or "classical" quantum mechanics: on their own they cannot make any predictions, they require supplementary theories to provide, respectively, the relevant forces and Hamiltonian. For quantum field theory this input will usually be given in the form of a Lagrangian.

1.1 Approaches

Once we have accepted this conceptual model of reality, we still need to turn it into a mathematical one. There are two main approaches to actually doing calculations in quantum field theories:

- the canonical quantization approach;
- the path integral approach.

These two approaches are somewhat analogous to the two corresponding approaches in regular quantum mechanics. We will develop the canonical quantization approach first.
We will also consider some alternative approaches like axiomatic quantum field theory and quantum field theory on a lattice, instead of a continuous field.

1.2 How to *do* quantum field theory

Just like in the frameworks Newtonian mechanics and quantum mechanics, there are established ways to perform calculations. Quantum field theory is no different. In this section a *very* brief overview of the general methodology within the canonical quantization approach will be given.

1.2.1 Relativity

As input for our quantum field theory we need to decide on both a type of field (e.g. real, complex, vectorial, of Dirac spinors etc.) and a Lagrangian. If we take both the field and the Lagrangian to be invariant under Poincaré transformations, we know that the whole theory will be compatible with relativity. (TODO: Why + only Lagrangian enough?)

1.2.2 From particles to fields

1.2.2.1 Wigner correspondence

1.2.3 Quantization

Now we need to make sure our theory is also consistent with quantum mechanics. To do that we will, just as in quantum mechanics, apply a sort of quantization recipe. In doing so, the field dynamical variables become field non-commuting operators.
This quantization of fields is called second quantization. This is as opposed to the quantization of particles theories, which is called first quantization.
In order to quantize fields, we need to generalize the particle quantization recipe. This is easier if we reformulate it in a new, but still equivalent way.

1.2.3.1 First quantization

It turns out that if we apply our known recipe for particle quantization, the commutators of the operators bear a striking resemblance to the Poisson brackets of the original dynamical properties. For example, the commutator of the position and momentum operator is as follows:

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

where the subscripts refer to the components along different axes. Comparing this to the Poisson bracket before quantization

$$\{x_i, p_j\} = \delta_{ij}$$

suggests the following recipe for quantization

$$\{\ ,\ \} \qquad \rightarrow \qquad -\frac{i}{\hbar}[\ ,\ ].$$

It turns out that this transformation of Poisson brackets into commutators is equivalent with our previous recipe.

1.2.3.2 Second quantization

For the second quantization we quite simply use the same recipe as above, but this time on the Poisson brackets of fields.

1.2.3.3 Creation and annihilation operators

TODO!

1.2.4 Interactions in quantum field theories

TODO!

Chapter 2

Elements of field theories

2.1 Lagrangian and Hamiltonian formalisms for fields

2.1.1 Review of analytical mechanics

This is a quick summary of the relevant results from the chapter on analytical mechanics.

2.1.1.1 Lagrangian formulation

Suppose we have a system described by N parameters $q_i(t)$. We can calculate how the system evolves in time from the initial conditions $q_i(t_0)$ and $\dot{q}_i(t_0)$, and the Lagrangian

$$L(q, \dot{q}, t) \qquad \begin{cases} q = \{q_1, \dots, q_N\} \\ \dot{q} = \{\dot{q}_1, \dots, \dot{q}_N\} \end{cases}$$

In a conservative system we have $L = L(q, \dot{q})$.
We can obtain equations of motion from the Euler-Lagrange equation.

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = 0$$

The Lagrangian of a system is not uniquely defined.

Example

Show that L, L' have the same Euler-Lagrange equation (TODO replace as in Gregory):

$$\begin{cases} L(q(t), \dot{q}(t)) \\ L'(q(t), \dot{q}(t)) = L(q(t), \dot{q}(t)) + \frac{\mathrm{d}\, f(q(t))}{\mathrm{d}t} \end{cases}$$

2.1.1.2 Hamiltonian formulation

We define the conjugate canonical momentum p

$$p \equiv \frac{\partial L}{\partial \dot{q}}$$

The Hamiltonian is the Legendre transformation of L. The independent variables become p and q.

$$H(p,q,t) \equiv p\dot{q} - L(q,\dot{q},t)|_{\dot{q}=\dot{q}(p,q)}$$

We can obtain equations of motion from the Hamilton equations (assuming a conservative system: $H = H(p,q)$).

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases}$$

We now define the Poisson bracket of $f(p(t),q(t)), g(p(t),q(t))$

$$\begin{aligned} \{f,g\}_t &\equiv \left(\frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}\right)_t \\ &\equiv \sum_{i=1}^{N}\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right)_t \end{aligned}$$

The subscript t means we evaluate everything at time t. This will be important later.
We can now rewrite the Hamiltonian equations of motion

$$\dot{q}(t) = \{q(t), H\}_t, \qquad \dot{p}(t) = \{p(t), H\}_t$$

We can further see that the Poisson brackets of p and q are

$$\begin{cases} \{q_i(t), p_j(t)\} = \delta_{ij} \\ \{q_i(t), q_j(t)\} = 0 = \{p_i(t), p_j(t)\} \end{cases}$$

2.2 Lagrangian and Hamiltonian for fields

In effect we can view a field as a system with infinite degrees of freedom: one for each point in space. If we are considering a field of real scalars that is, otherwise we need multiple degrees of freedom for each point in space.
We will develop the theory here for a real scalar field $\varphi(\mathbf{x},t)$. The generalisation for other types of field is straightforward.
It will prove useful to work with the Lagrangian density TODO!
We go from finite $(q_i(t))$ to continuous $(\varphi_x(t) \equiv \varphi(\overline{x},t) \quad \overline{x} \in \mathbb{R}^3)$

Finite		continuous	
$q_i(t)$	$\varphi_x(t) \equiv \varphi(\overline{x},t)$		$\overline{x} \in \mathbb{R}^3$
$\sum_{i=1}^{N}()$		$\int \mathrm{d}^3x\,()$	
$L(t)$		$\mathcal{L}(x,t)$	

We call $\mathcal{L}(\overline{x},t)$ the Lagrangian density.

$$L(t) = \int \mathrm{d}^3x \mathcal{L}(\overline{x},t)$$

Example

$$\text{Discrete} \qquad L(t) = \frac{1}{2}\sum m_i q_i - V(q^i)$$

$$\text{Continuous} \qquad L(t) = \int \mathrm{d}^3x \left(\frac{1}{2}\rho(x)(\partial_0\varphi)^2 - V(\varphi^2)\right)$$

With $\mathcal{L}(\overline{x},t) = \frac{1}{2}\rho(x)(\partial_0\varphi)^2 - V(\varphi^2)$

The action for our system ($x \equiv (\overline{x},t)$)

$$S[\varphi, \partial D_4] \equiv \int \mathrm{d}t L(t) \equiv \int \mathrm{d}^4x \mathcal{L}(x)$$

$$\mathcal{L}(x) \equiv \mathcal{L}(\varphi(x), \partial_\mu\varphi(x)) \qquad \to \qquad \text{conservative}$$

The least action principle: the physical field configurations (fixing boundary conditions) are the ones that minimizes the action.

$$\textbf{Necessary condition: } \delta S = 0 \qquad \begin{pmatrix} \delta_0\varphi = \varphi'(x) - \varphi(x) \\ \delta_0\varphi(\partial D_4) = 0 \end{pmatrix}$$

$$\begin{aligned}
\delta S[\varphi] &= S[\varphi + \delta_0\varphi] - S[\varphi] \\
&= \int_{D_4} \mathrm{d}^4x \mathcal{L}(\varphi + \delta_0\varphi, \partial_\mu\varphi + \partial_\mu\delta_0\varphi) - \int \mathrm{d}^4x \mathcal{L}(\varphi, \partial_\mu\varphi) \\
&= \int_{D_4} \mathrm{d}^4x \delta_0\mathcal{L}(\varphi, \partial_\mu\varphi) \qquad \delta_0\partial_\mu\varphi = \partial_\mu\delta_0\varphi \\
&= \int_{D_4} \mathrm{d}^4x \left(\frac{\partial\mathcal{L}}{\partial\varphi}\delta_0\varphi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta_0\partial_\mu\varphi\right)
\end{aligned}$$

$$\delta S = ???$$

Note: The Lagrangian density $\mathcal{L}$ is not unique.

Example

Prove that $\mathcal{L}, \mathcal{L}'$ give the same E.L. eq

$$\begin{cases} \mathcal{L}(\varphi, \partial_\mu\varphi) \\ \mathcal{L}'(\varphi, \partial_\mu\varphi) \equiv \mathcal{L}(\varphi, \partial_\mu\varphi) + \partial_\mu K^\mu(\varphi) \end{cases}$$

2.2.1 Hamiltonian description

- Conjugate momentum $\pi(x)$

$$\pi(x) \equiv \frac{\partial\mathcal{L}}{\partial\partial_0\varphi} \qquad \partial_0\varphi \equiv \partial_0\varphi(\pi, \varphi)$$

- Hamiltonian density legendre transformation

$$\mathcal{H}(\pi,\varphi) = \pi\partial_0\varphi - \mathcal{L}(\varphi,\partial_\mu\varphi)|_{\partial_0\varphi=\partial_0\varphi(\pi,\varphi)}$$

So we get the Hamiltonian

$$H \equiv \int \mathrm{d}^3x\mathcal{H}(\overline{x},t)$$

Lecture 22/10

$$\begin{aligned} \varphi_x &\equiv \varphi(\overline{x},t) \qquad \text{Field} \\ \mathcal{L}(\varphi,\partial_\mu\varphi) &\equiv \mathcal{L}(\overline{x},t) \end{aligned}$$

$$\delta D = 0 \quad \Leftrightarrow \quad \boxed{\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}}$$

We define

$$\pi = \frac{\partial\mathcal{L}}{\partial\partial_0\varphi} \qquad \text{and} \qquad H = \pi\partial_0\varphi - \mathcal{L} = \sum_i \pi_i\partial_0\varphi_i - \mathcal{L} \equiv \mathcal{H}(\pi,\varphi)$$

Then we have the Hamiltonian equations

$$\boxed{\partial_0\varphi = \frac{\partial\mathcal{H}}{\partial\pi} \qquad | \qquad \partial_0\pi = -\frac{\partial\mathcal{H}}{\partial\varphi}}$$

Example

Exercise: Verify the Hamiltonian equations:

$$\begin{aligned} \frac{\partial\mathcal{H}}{\partial\pi} &= \frac{\partial}{\partial\pi}\left(\pi\partial_0\varphi - \mathcal{L}\right) \\ &= \partial_0\varphi + \pi\frac{\partial\partial_0\varphi}{\partial\pi} - \frac{\partial\mathcal{L}}{\partial\pi}\frac{\partial\varphi}{\partial\pi} - \frac{\partial\mathcal{L}}{\partial\partial_i\pi}\frac{\partial\partial_i\varphi}{\partial\pi} \\ &= \dot{\varphi} + \pi\frac{\partial\dot{\varphi}}{\partial\pi} - \frac{\partial\mathcal{L}}{\partial\partial_0\varphi}\frac{\partial\partial_0\varphi}{\partial\pi} - \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\frac{\partial\partial_i\varphi}{\partial\pi} \\ &= \dot{\varphi} + \pi\frac{\partial\dot{\varphi}}{\partial\pi} \end{aligned}$$

Also

$$\varphi(\overline{x},t) = \int \mathrm{d}^3y\delta^3(\overline{x}-\overline{y})\varphi(\overline{y},t)$$

And

$$\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\frac{\partial\partial_i\varphi}{\partial\pi}\right)_{\overline{x}} = \int \mathrm{d}^3y\frac{\partial x}{\partial\partial_i\varphi_x}\frac{\partial\varphi(y)}{\partial\pi}\partial_i^x\delta^3(\overline{x}-\overline{y}) = 0$$

Because $\frac{\partial\varphi(y)}{\partial\pi} = 0$. The other equation is homework.

2.2.2 Poisson brackets for fields

Let's compute the functionals $F[\varphi, \pi], G[\varphi, \pi]$

$$\{F, G\}_t \equiv \int \mathrm{d}^3x \left(\frac{\delta F}{\delta \varphi(\overline{x})} \frac{\delta G}{\delta \pi(\overline{x})} - \frac{\delta F}{\delta \pi(\overline{x})} \frac{\delta G}{\delta \varphi(\overline{x})} \right)_t$$

So the Hamilton equations become

$$\dot{\varphi}(\overline{x}, t) = \{\varphi(\overline{x}, t), H\}$$
$$\dot{\pi}(\overline{x}, t) = \{\pi(\overline{x}, t), H\}$$

Poisson brackets on fields:

$$\{\pi(\overline{x}, t), \pi(\overline{y}, t)\} = 0$$
$$\{\varphi(\overline{x}, t), \pi(\overline{y}, t)\} = \delta^3(\overline{x} - \overline{y})$$
$$\{\varphi(\overline{x}, t), \varphi(\overline{y}, t)\} = 0$$

2.3 Summary of functionals

2.3.0.1 Definition of functionals

A functionals F is a map between $\mathcal{C}$, the space of functions, and $\mathbb{R}$ (or $\mathbb{C}$)

$$F : \mathcal{C} \to \mathbb{R}(\mathbb{C}) : f \mapsto F[f]$$

2.3.0.2 Variation of a functional

Let's take $g, \delta g \in \mathcal{C}$

$$\delta_0 F \equiv F[g + \delta_0 g] - F[g]$$

2.3.0.3 Functional derivative

Let's take $g, \delta g \in \mathcal{C}$

$$\delta_0 F \equiv F[g + \delta_0 g] - F[g] \equiv \int \mathrm{d}x \left(\frac{\delta F}{\delta g(x)} \right) \delta_0 g$$

We call $\frac{\delta F}{\delta g(x)}$ the functional derivative of F with respect to the function $g(x)$.

The functional derivative $\frac{\delta F}{\delta g}$ has all the good properties of a derivative.

Useful properties:

- Identity functional

$$g(y) \equiv \int \mathrm{d}x \delta(x - y) g(x) \equiv F_{\mathrm{ID}}[g]$$

SO we have

$$\frac{\delta F_{\text{ID}}}{g(x)}\frac{\delta g(y)}{\delta g(x)} = \delta(x-y)$$

Also

$$\begin{aligned}\delta_0 F_{\text{ID}} &= F_{\text{ID}}[g+\delta_0 g] - F_{\text{ID}}[g] = g + \delta g - g = \delta g \\ &= \int \mathrm{d}x \left(\frac{\delta F_{\text{ID}}}{\delta g}\right)\delta g = \int \mathrm{d}x \delta(x-y)\delta g(x)\end{aligned}$$

Example

Exercise: Calculate $\frac{\delta H}{\delta\varphi_t(\overline{x})}, \frac{\delta H}{\delta\pi_t(\overline{x})}$. We have

$$H[\varphi_t, \pi_t] = \int \mathrm{d}^3x \mathcal{H}(\varphi(\overline{x}, t), \pi(\overline{x}, t))$$

(remembering that time is fixed, so subscript) and

$$\begin{aligned}\delta_0 H(\varphi_t, \pi_t) &= H[\varphi_t + \delta\varphi_t, \pi_t + \delta\pi_t] - H[\varphi_t,]pi_t] \\ &= \int \mathrm{d}^3x \left(\frac{\delta H}{\delta_0\varphi_t}\delta_0\varphi_t + \frac{\delta H}{\delta_0\pi_t}\delta_0\pi_t\right)\end{aligned}$$

$$\begin{cases}\frac{\delta H}{\delta_0\varphi_t} = \frac{\delta\mathcal{H}}{\delta_0\varphi_t} \\ \frac{\delta H}{\delta_0\pi_t} = \frac{\delta\mathcal{H}}{\delta_0\pi_t}\end{cases}$$

Example

Exercise: Verify the Hamilton equations as Poisson brackets

$$\begin{aligned}\dot{\varphi} = \{\varphi_t(\overline{x}), H\}_t &= \int \mathrm{d}^3y \left(\tfrac{\delta\varphi_t(x)}{\delta\varphi_t(y)}\frac{\delta H}{\delta\pi_t(y)} - \frac{\delta\varphi_t(x)}{\delta\pi_t(y)}\frac{\delta H}{\delta\varphi_t(y)}\right) \\ &= \int \mathrm{d}^3y \delta^3(\overline{x}-\overline{y})\frac{\delta H}{\delta\pi_t(y)} = \frac{\delta H}{\delta\pi_t(x)} = \frac{\partial\mathcal{H}}{\partial\pi_t}\end{aligned}$$

The other one ($\dot{\pi} = \{\pi_t(\overline{x}, t), H\}$) is left as an exercise.

2.4 Symmetries and conserved quantities

First we give an operative definition of symmetry:

A symmetry of the theory is a transformation of fields and / or coordinates that leaves the action invariant:

$$\delta S = 0$$

TODO: O constant of motion: $[O, H] = 0$. + 2.4 in Mandl - Shaw

Symmetry transformations leave the equations of motion invariant

$$\begin{cases} \delta\mathcal{L} = 0 \\ \delta\mathcal{L} = \partial_\mu K^\mu \end{cases} \qquad \Rightarrow \qquad \delta S = 0$$

The symmetries form a group.
Classification:

- Discrete or continuous? We will be interested in continuous groups and make a further classification among continuous groups:
- Global(α_i) or local ($\alpha_i(x)$).
- Internal (only on fields, φ) or spacetime (φ and X).

	Global	Local
Internal	Leptonic, Hadronic number	Gauge (E.M.)
Spacetime	Lorentz, Poincaré	GR

In this lecture we will maintly be discussing global symmetries. Consider a theory with a continuous global group of symmetry transformations.

Noether's theorem: To any continuous and global symmetry transformation a conserved current $J^\mu_{(0)}$ is associated.

$$\partial_\mu J^\mu_{(0)} = 0$$

$$J^\mu_{(0)} = (\rho_{(0)}, \overline{j}_{(0)})$$

Conserved charge:

$$Q_0 = \int \mathrm{d}^3x \rho_{(0)}(\overline{x}, t) = \int \mathrm{d}^3x J^\mu_{(0)}(\overline{x}, t)$$

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0$$

2.4.1 Global internal group - symmetry

$$\begin{cases} x'^\mu = x^\mu \\ \varphi'(x) = \varphi(x) + \delta_0\varphi(x) \end{cases} \qquad \to \qquad \begin{cases} \delta x^\mu = 0 \\ \delta\varphi = \epsilon^a X_a(\varphi) \end{cases}$$

$$\begin{aligned} 0 = \delta S &= \int \mathrm{d}^4x \left\{ \left[\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} \right] \delta_0\varphi + \partial_\mu \left[\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} \delta_0\varphi \right] \right\} \\ &= \epsilon^{(a)} \int \mathrm{d}^4x \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} X_{(a)} \right) \qquad \to \qquad \partial_\mu J^\mu_{(a)} = 0 \end{aligned}$$

$$\begin{cases} J^\mu_{(a)} \equiv \frac{\partial\mathcal{L}}{\partial\partial_\nu\varphi} X_{(e)} \\ Q_{(e)} \equiv \mathrm{d}^3 \frac{\partial\mathcal{L}}{\partial\partial_0\varphi} X_a \end{cases}$$

Lecture 23/10

2.4.2 Symmetries "external" (fields + coord)

$$\begin{cases} x'^\mu = x^\mu + \delta x^\mu \\ \varphi'(x') = \varphi(x) + \delta_0\varphi(x) \end{cases} \quad \to \quad \begin{cases} \delta x^\mu = \epsilon^{(a)}\Xi^\mu_{(a)} \\ \delta\varphi = \epsilon^a X_a \end{cases}$$

There are two main differences:

1. The variation of the field is given by

$$\begin{aligned} \delta\varphi \equiv \varphi'(x') - \varphi(x) &= \varphi'(x') - \varphi(x') + \varphi(x') - \varphi(x) \\ &= \delta_0\varphi(x) + (\partial_\mu\varphi)\delta x^\mu + \mathcal{O}(\delta x^2) \end{aligned}$$

   The total variation is the synchronous variation plus a transport term (to lowest order).

2. $\mathrm{d}^4x' = (1 + \partial_\mu\delta x^\mu)\,\mathrm{d}^4x + \mathcal{O}(\delta x^2)$

Now we calculate

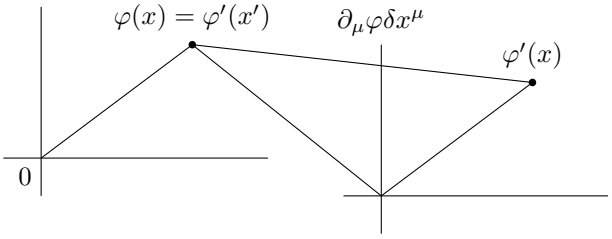
$$\begin{aligned} 0 = \delta S = \delta\int(\delta\,\mathrm{d}^4x)\mathcal{L} &= \int \mathrm{d}^4x\delta\mathcal{L} \\ &= \int \mathrm{d}^4x\,\{(\partial_\mu\delta x^\mu)\mathcal{L} + (\partial_\mu\mathcal{L})\delta x^\mu + \delta_0\mathcal{L}\} \\ &= \int \mathrm{d}^4x\left\{\partial_\mu(\mathcal{L}\delta x^\mu) + \left[\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\right]\delta_0\varphi + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta_0\varphi\right)\right\} \\ &\quad \varphi = \text{physical configuration} \\ &= \int \mathrm{d}^4x\partial_\mu\left\{\mathcal{L}\delta x^\mu + \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta_0\varphi\right\} \qquad \delta_0 \to (\delta - \partial_\mu) \\ &= \int \mathrm{d}^4x\partial_\mu\left\{\left[\eta^\mu_{\ \rho}\mathcal{L} - \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\rho\varphi\right]\delta x^\rho + \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta\varphi\right\} \\ &= \epsilon^{(a)}\int \mathrm{d}^4x\partial_\mu\left\{\left[\eta^\mu_{\ \rho}\mathcal{L}\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\rho\varphi\right]\Xi^\rho_{(a)} + \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}X_{(a)}\right\} = 0 \end{aligned}$$

So

$$\begin{cases} J^\mu_{(a)} = (\pm)\left[\eta^\mu_{\ \rho}\mathcal{L} - \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\rho\varphi\right]\Xi^\rho_{(a)} = \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}X_{(a)} \\ \boxed{\partial_\mu J^\mu_{(a)} = 0} \qquad \boxed{Q_{(a)} = \int \mathrm{d}^3x J^0_{(a)}} \end{cases}$$

2.4.2.1 Translational invariance

$$\begin{cases} x'^\mu = x^\mu + \epsilon^\mu \\ \varphi'(x') = \varphi(x) \end{cases} \quad \to \quad \begin{cases} \delta x^\mu = \eta^\mu_{\ \nu}\epsilon^\nu \\ \delta\varphi = 0 \end{cases} \quad \to \quad \Xi^\mu_{(\nu)} = \eta^\mu_{\ \nu}$$



$$J^{\mu}_{(\nu)} = \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi} \partial_\rho \varphi - \eta^{\mu}_{\ \nu} \mathcal{L} \equiv \tilde{T}^{\mu}_{\ \nu}$$

$$\tilde{T}^{\mu\nu} = \text{canonical energy-momentum tensor}$$

$$\begin{cases} T^{\mu\nu} = \text{covariant EM tensor} \\ \tilde{T}^{\mu\nu} \text{is not symmetric } \mu \to \nu \end{cases}$$

$$Q_{(\nu)} \equiv \int \mathrm{d}^3 x \tilde{T}^{0}_{\ \nu} = \int \mathrm{d}^3 x \mathcal{P}_\nu \equiv P_\nu$$

$$\begin{cases} Q_{(0)} = \int \mathrm{d}^3 x \mathcal{P}_0 = \int \mathrm{d}^3 x \left(\frac{\partial \mathcal{L}}{\partial \partial_0 \varphi} \partial_0 \varphi - \mathcal{L} \right) \equiv \int \mathrm{d}^3 x \mathcal{H} = H \\ Q_{(i)} = \int \mathrm{d}^3 x \mathcal{P}_i = \int \mathrm{d}^3 x \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi} \partial_i \varphi \end{cases}$$

2.4.2.2 Rotational invariance

$$\begin{cases} x'^{\mu} = x^\mu + \omega^{\mu}_{\ \nu} x^\nu \\ \varphi'(x') = \varphi(x) - \frac{i}{2} \omega^{\mu\nu} \Omega_{\mu\nu} \varphi(x) \end{cases} \quad \to \quad \begin{cases} \delta x^\mu = \frac{1}{2} \omega^{\rho\sigma} \Xi^{\mu}_{(\rho\sigma)} \\ \delta \varphi = \frac{1}{2} \omega^{\rho\sigma} X_{(\rho\sigma)} \end{cases}$$

$$\begin{cases} \Xi^{\mu}_{(\rho\sigma)} = \left(\eta^{\mu}_{\ \rho} \eta^{\nu}_{\ \sigma} - \eta^{\mu}_{\ \sigma} \eta^{\nu}_{\ \rho} \right) x^\rho \\ X_{(\rho\sigma)} = -i \Omega_{\rho\sigma} \varphi \end{cases} \quad \to \quad \begin{cases} \Omega_{\rho\sigma} = 0 \qquad \text{(scalar)} \\ \Omega_{\rho\sigma} = \Sigma'_{\rho\sigma} \qquad \text{(spinor)} \end{cases}$$

$$J^{\mu}_{\rho\sigma} = \left(x_\rho \tilde{T}^{\mu}_{\sigma} - X_\sigma \tilde{T}^{\mu}_{\rho} \right) + S^{\mu}_{\rho\sigma}$$

$$\begin{aligned} Q_{(\rho\sigma)} &= \int \mathrm{d}^3 x J^0_{(\rho\sigma)} = \int \mathrm{d}^3 x \left[(x_\rho \mathcal{P}_\sigma - x_\sigma \mathcal{P}_\rho) + \mathcal{P}_{\rho\sigma} \right] \\ &= M_{\rho\sigma} = L_{\rho\sigma} + S_{\rho\sigma} \end{aligned}$$

Which is the external angular momentum plus the spin. ($x_\rho \mathcal{P}_\sigma - x_\sigma \mathcal{P}_\rho = \mathcal{L}_{\rho\sigma}$ and $\mathcal{P}_{\rho\sigma} = S^0_{\rho\sigma}$)

2.5 Spin of a field

2.6 Dimensional analysis

In the rest of part we set $c = 1 = \hbar$

Chapter 3

Canonical quantization

It is now time to go from a classical to a quantum theory. This means quantization!

3.1 Quantization

3.1.1 First quantization

We start by exploring the quantization of a system with a finite number of degrees of freedom. We have already seen how to do this in the part on quantum mechanics. Here we give an alternative recipe (TODO: show that it is equivalent)

$$\begin{cases} p \equiv \frac{\partial L}{\partial \dot{q}} \\ H = p\,\dot{q} - L \end{cases}$$

The equations of motion are given by the Poisson brackets

$$\begin{cases} \{q_i, p_j\}_t = \delta_{ij} \\ \{q_i, q_j\} = 0 = \{p_i, p_j\} \end{cases} \qquad \begin{cases} \dot{q}_i = \{q_i, H\}_t \\ \dot{p}_i = \{p_i, H\}_t \end{cases}$$

We obtain the canonical quantization by applying the following substitutions:

- (q_i, p_i) which are variables become (X_i, P_i) which are operators;
- $\{\ ,\ \}_t \qquad \to \qquad -\frac{i}{\hbar}[\ ,\ ]_t$.

This gives us the following equations:

$$\begin{cases} [X_i, P_j]_t = i\hbar\delta_{ij} \\ [X_i, X_j] = 0 = [P_i, P_j] \end{cases} \qquad \begin{cases} \frac{\mathrm{d}X_i}{\mathrm{d}t} = -\frac{i}{\hbar}[X_i, H] \\ \frac{\mathrm{d}P_i}{\mathrm{d}t} = -\frac{i}{\hbar}[P_i, H] \end{cases}$$

As we can see, the operators evolve, so these are equations of motion in the Heisenberg picture. For a finite number of degrees of freedom in momentum space we have N independent oscillators.

$$\begin{cases} [a_p, a_k^\dagger] = \delta_{pk} \\ [a_p, a_k] = 0 = [a_p^\dagger, a_k^\dagger] \end{cases}$$

(90% of physics is harmonic oscillators because we expand to second order)

3.1.2 Second quantization

This procedure generalises quite nicely to fields with infinite degrees of freedom.

$$\begin{cases} \pi(x) \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi} \\ \mathcal{H} = \pi \cdot \partial_0 \varphi - \mathcal{L} \end{cases}$$

The Poisson brackets:

$$\begin{cases} \{\varphi(\overline{x},t), \pi(\overline{y},t)\}_t = \delta^3(\overline{x}-\overline{y}) \\ \{\varphi(\overline{x},t), \varphi(\overline{y},t)\}_t = 0 = \{\pi(\overline{x},t), \pi(\overline{y},t)\}_t \end{cases} \qquad \begin{cases} \dot{\varphi}(\overline{x},t) = \{\varphi(\overline{x},t), H\}_t \\ \dot{\pi}(\overline{x},t) = \{\pi(\overline{x},t), H\}_t \end{cases}$$

The quantization given by following substitution:

- (φ, π) = classical fields $\rightarrow$ $(\hat{\varphi}, \hat{\pi})$ = operator fields
- $\{\ ,\ \}_t \qquad \rightarrow \qquad -i[\ ,\ ]_t \qquad (\hbar = 1 = c)$

This gives us the following quantization:

$$\begin{cases} [\hat{\varphi}(\overline{x},t), \hat{\pi}(\overline{y},t)]_t = i\delta^3(\overline{x}-\overline{y}) \\ [\hat{\varphi}(\overline{x},t)\hat{\varphi}(\overline{y},t)]_t = 0 = [\hat{\pi}(\overline{x},t), \hat{\pi}(\overline{y},t)]_t \end{cases} \qquad \begin{cases} \dot{\hat{\varphi}}(\overline{x},t) = -i\left[\hat{\varphi}(\overline{x},t), \hat{H}\right]_t \\ \dot{\hat{\pi}}(\overline{x},t) = -i\left[\hat{\pi}(\overline{x},t), \hat{H}\right]_t \end{cases}$$

In momentum space we get an equivalent quantization condition:

$$\begin{cases} \left[a(p), a^\dagger(k)\right] = \delta^3(\mathbf{p}-\mathbf{k}) \\ [a(p), a(k)] = 0 = \left[a^\dagger(p), a^\dagger(k)\right] \end{cases}$$

Which is a continuous set of uncoupled harmonic oscillators.

3.2 Number (density) operator

TODO see p5-6 in Mandl-Shaw
The number density operator

$$\mathcal{N}(k) \equiv a^\dagger(k)a(k)$$

has the following properties

- $\mathcal{N}(k)^\dagger = \mathcal{N}(k)$
- $[\mathcal{N}(k), a(p)] = -a(p)\delta^3(\overline{k}-\overline{p})$
  $[\mathcal{N}(k), a^\dagger(p)] = +a^\dagger(p)\delta^3(\overline{k}-\overline{p})$

The number operator

$$N \equiv \int \mathrm{d}^3k \mathcal{N}(k)$$

has the following properties

- $N^\dagger = N$
- $[N, a(p)] = -a(p)$
  $[N, a^\dagger(p)] = +a(p)$

3.3 Normal ordering

3.4 What are the states in QFT?

So far we have operator fields, but no states. The space of states in quantum field theory is called the Fock Space.

3.4.1 Postulate of the existence of vacuum $= |0\rangle$

$$\boxed{a(k)\,|0\rangle = 0} \qquad \forall k$$

$$\begin{cases} \mathcal{N}(k)\,|0\rangle = 0 & \to \quad 0 \text{ occupation number} \\ N[H]\,|0\rangle = 0 & \to \quad 0 \text{ energy, momentum} \end{cases}$$

3.4.2 The Fock space

(Hilbert space) is the space with all possible states generated by applying $a^\dagger(k)$ on $|0\rangle$. So a generic state in Fock space can be written as

$$|n_1(k_1), \ldots, n_l(k_l)\rangle \quad \sim \quad a^\dagger(k_1)^{n_1} \ldots a^\dagger(k_l)^{n_l}\,|0\rangle$$

Example

We consider the simplest non vacuum state $|1(k)\rangle \sim a^\dagger(k)\,|0\rangle$.

$$\begin{aligned} N\,|1(k)\rangle &\sim \int \mathrm{d}^3p \mathcal{N}(p) a^\dagger(k)\,|0\rangle \\ &= \int \mathrm{d}^3p \left([\mathcal{N}(p), a^\dagger(k)] + a^\dagger(k)\mathcal{N}(p)\right)|0\rangle \\ &= \int \mathrm{d}^3p \delta^3(k-p) a^\dagger(k)\,|0\rangle + \cancel{a^\dagger(p)\mathcal{N}(k)\,|0\rangle} \\ &= \int \mathrm{d}^3p \left(1\delta^3(k-p) a^\dagger(p) + \cancel{a^\dagger(p)\mathcal{N}(k)}\right)|0\rangle \\ &\sim +1\,|1(k)\rangle \end{aligned}$$

Similarly

$$\begin{cases} H\,|1(k)\rangle = \omega_k\,|1(k)\rangle & \text{where} \quad \omega_k = \left(m^2 + |\overline{k}|^2\right)^{1/2} \\ P\,|1(k)\rangle = k\,|1(k)\rangle \end{cases}$$

From this we get for a state of n identical particles

$$\begin{aligned} |n(k)\rangle &\sim a^\dagger(k)^n\,|0\rangle \\ N\,|n(k)\rangle &= n\,|n(k)\rangle \\ H\,|n(k)\rangle &= n\omega_k\,|n(k)\rangle \\ \overline{P}\,|n(k)\rangle &= n\overline{k}\,|n(k)\rangle \end{aligned}$$

Also

$$\begin{aligned} |n_1(k_1), n_2(k_2)\rangle &\sim a^\dagger(k_1)^{n_1} a^\dagger(k_2)^{n_2}\,|0\rangle \\ N\,|n_1(k_1), n_2(k_2)\rangle &= (n_1+n_2)\,|n_1(k_1), n_2(k_2)\rangle \\ H\,|n_1(k_1), n_2(k_2)\rangle &= (n_1\omega_1+n_2\omega_2)\,|n_1(k_1), n_2(k_2)\rangle \\ P\,|n_1(k_1), n_2(k_2)\rangle &= (n_1k_1+n_2k_2)\,|n_1(k_1), n_2(k_2)\rangle \end{aligned}$$

From this example we learn that the Fock space of a real scalar field corresponds to a particle with Bose statistics, as we can add as many identical particles as we want.

Scalar field corresponds to spin 0

3.4.3 Normalization of states

$$\langle 1(p)|1(k)\rangle = \delta^3(\overline{k}-\overline{p})$$

Firstly

$$\langle 0|0\rangle = 1$$

Normally we would normalize the states as follows

$$\begin{cases} |1(k)\rangle = a^\dagger(k)\,|0\rangle \\ \langle 1(p)|1(k)\rangle = \delta^3(\overline{k}-\overline{p}) \end{cases}$$

This normalisation is not invariant under Lorentz transformations however.
So we use the "good" normalization

$$\begin{cases} |1(k)\rangle = (2\pi)^{3/2}\sqrt{2\omega_k}a^\dagger(k)\,|0\rangle \\ \langle 1(p)|1(k)\rangle = (2\pi)^3 2\omega_k\delta^3(\overline{k}-\overline{p}) \end{cases}$$

3.5 Covariant commutators and propagators

3.5.1 Feynman propagator

The Feynman propagator Δ_F is defined as

$$\Delta_F(x-y) \equiv \langle T\left(\varphi(x)\varphi(y)\right)\rangle\,0 = \varphi(x)\varphi(y)$$

$$\begin{aligned}\Delta_F(x-y) &= \theta(x_0-y_0)\,\langle T\,(\varphi(x)\varphi(y))\rangle\,0 + \theta(y_0-x_0)\,\langle T\,(\varphi(x)\varphi(y))\rangle\,0\\ &= \theta(x_0-y_0)\,\langle T\,(\varphi_+(x)\varphi_-(y))\rangle\,0 + \theta(y_0-x_0)\,\langle T\,(\varphi_+(x)\varphi_-(y))\rangle\,0\\ &= \theta(x_0-y_0)\,\langle[\varphi_+(x),\varphi_-(y)]\rangle\,0 - \theta(y_0-x_0)\,\langle[\varphi_-(x),\varphi_+(y)]\rangle\,0\\ &= \theta(x_0-y_0)D_+(x-y) - \theta(y_0-x_0)D_-(x-y)\end{aligned}$$

The following relation holds

$$T\,(\varphi(x)\varphi(y)) = N\,(\varphi(x)\varphi(y)) + \varphi(x)\varphi(y)$$

$$\varphi(x)\varphi(y) = T\,($$

$$-\theta(x_0-y_0)\,(\cancel{\varphi_-(x)\varphi_-(y)} + \cancel{\varphi_+(x)\varphi_+(y)} + \cancel{\varphi_-(x)\varphi_+(y)} + \varphi_-(y)\varphi_+(x)) - \theta(y_0-x_0)\,(\cancel{\varphi_-(x)\varphi_-(y)} + \cancel{\varphi_+(x)\varphi_+(y)} + \varphi_-$$

The Feynman propagator (complex scalar)

$$\begin{aligned}\Delta_F(x-y) &\equiv \left\langle T\left(\varphi(x)\varphi^\dagger(y)\right)\right\rangle 0 = \underbrace{\varphi(x)\varphi^\dagger(y)}\\ &= \theta(x_0-y_0)D_+(x-y) - \theta(y_0-x_0)D_-(x-y)\end{aligned}$$

The Feynman propagator for Dirac fields

$$\begin{aligned}S_+^{\alpha\beta}(x-y) &\equiv \left\langle T\left(\psi_\alpha(x)\overline{\psi}_\beta(y)\right)\right\rangle 0 = \underbrace{\psi_\alpha(x)\overline{\psi}_\beta(y)}\\ &= \theta(x_0-y_0)S_+^{\alpha\beta}(x-y) - \theta(y_0-x_0)S_-^{\alpha\beta}(x-y)\\ &= (i\partial\!\!\!/ + m)D_F(x-y)\end{aligned}$$

Ex.

3.5.2 Time-ordered product

The time-ordered product of two operators A and B

$$T\,(A(t_1)B(t_2)) = \begin{cases} A(t_1)B(t_2) & (t_1 > t_2)\\ \pm B(t_2)A(t_1) & (t_1 < t_2)\end{cases}$$

Where the $\pm$ is $+$ for bosons and $-$ for fermions. For n operators, the time-ordered product orders them by time and introduces a factor $(-1)^p$ where p is the number of fermionic exchanges necessary to achieve time ordering.
The time-ordered product is also taken to be linear, if addition were to turn up in the product.

3.5.3 Introducing Feynman diagrams

integral over time, no time ordering. time left to right. arrows and momenta

3.5.4 Microcausality

3.6 Spin statistics theorem

3.6.1 Quantization with anti-commutators

Chapter 4

Some examples of quantum field theories

We want an explicit expression for the Lagrangian density. We will try to guess its form.

1. The action is more or less an observable so we want it to be real (Hermitian)

$$\mathcal{L} = \textbf{real (Hermitian)}$$

2. The action is a scalar (under Poincaré transformation)

$$\mathcal{L} = \textbf{scalar density}$$

3. The action is adimensional in natural units

$$\hookrightarrow S = \int \mathrm{d}^4 x \mathcal{L} = [M]^{-4}[M]^4 = [M]^0$$

4. By equation of motion $\rightarrow$ relativistic equation = nice

$$\mathcal{L}(\varphi, \partial_\mu \varphi)$$

4.1 Klein-Gordon field (real scalar field)

As a first example, we consider the following Lagrangian for a real scalar field:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2\varphi^2$$

where m is a parameter that we will identify with the mass. This can be seen as the sum of kinetic and a mass term (TODO why?). Knowing that the Lagrangian must have a dimension of $[M^4]$ and that each derivative has dimension $[\partial] = [M^1]$, we see that

$$\begin{cases} [\varphi] = [M^1] \\ [m] = [M^1] \end{cases}$$

Luckily this is consistent with our identification of m as the mass.

Because we are dealing with a real and relativistic field, we impose the following conditions

$$\begin{cases} \varphi(x) = \varphi^*(x) \\ \varphi'(x') = \varphi(x) \end{cases}$$

where the prime means the quantity has been transformed by a Poincaré transformation.

4.1.1 Equation of motion

Applying the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi} = 0$$

We get the Klein-Gordon equation:

$$-m^2\varphi - \partial_\mu \partial^\mu \varphi = -(\Box + m^2)\varphi = 0$$

As we know this has the following general solution:

$$\varphi(x) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3k \left(a(k)e^{-ikx} + a^*(k)e^{ikx} \right)_{k_0=\omega_k}$$

4.1.2 Hamiltonian formulation

We now calculate the relevant quantities in the Hamiltonian formulation.

$$\pi \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi} = \frac{1}{2}(\partial_0\varphi + \partial^0\varphi) = \partial_0\varphi$$

$$\mathcal{H} = \pi\partial_0\varphi - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}(\overline{\nabla}\varphi)^2 + \frac{1}{2}m^2\varphi^2$$

We see that the Hamiltonian density is always positive. (TODO: as opposed to ...)

$$H = \int \mathrm{d}^3x \mathcal{H} > 0$$

Example

Exercise: Verify Poisson Brackets

$$\dot{\varphi} = \{\varphi(\overline{x}, t), H\}_t$$
$$\dot{\pi} = \{\pi(\overline{x}, t), H\}_t$$

$$\{\varphi(\overline{x}, t), \varphi(\overline{y}, t)\} = 0$$

4.1.3 Symmetries

4.1.3.1 Translational invariance

$$J^\mu_{(\nu)} = (\partial^\mu\varphi)(\partial_\nu\varphi) - \eta^\mu_{\ \nu}\mathcal{L}$$

$$\partial_\mu J^\mu_{(\nu)} = 0 \qquad \rightarrow \qquad \text{Exercise: prove}$$

The 4 conserved charges:

$$\begin{cases} P_0 \equiv H = \int \mathrm{d}^3x \left\{\pi\partial_0\varphi - \mathcal{L}\right\} = \int \mathrm{d}^3x \mathcal{H} \\ P_i = \int \mathrm{d}^3x \mathcal{P}_i = \int \mathrm{d}^3x \pi(\partial_i\varphi) \end{cases}$$

4.1.3.2 Rotational invariance

$$\begin{cases}\delta x^\mu = \omega^{\rho\sigma}\Xi^\mu_{(\rho\sigma)} & \to \qquad \Xi^\mu_{(\rho\sigma)} = \left(\eta^\mu_\rho\eta^\nu_\sigma - \eta^\mu_\sigma\eta^\nu_\rho\right)x^\nu \\ \delta\varphi = \frac{1}{2}\omega^{\rho\sigma}X_{(\rho\sigma)} = 0 & \to \qquad X_{(\rho\sigma)} = 0\end{cases}$$

Differences??

$$\frac{1}{2}, \nu\text{ipv}\rho$$

$$J^\mu_{(\rho\sigma)} = \left(x_\rho \tilde{T}^\mu_\sigma - x_\sigma \tilde{T}^\mu_\rho\right) + (X)$$

$$Q_{(\rho\sigma)} = \int \mathrm{d}^3x J^0_{(\rho\sigma)} = \int \mathrm{d}^3x \left(x_\rho \mathcal{P}_\sigma - x_\sigma \mathcal{P}_\rho\right) = L_{(\sigma\rho)??(\text{last equality})}$$

Example

Exercise: Write P^μ in the momentum space (as functions of $a(k), a^*(k)$)
Solution:

$$\varphi(x) = \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k}}\left(a(k)e^{-ikx} + a^*(k)e^{ikx}\right)_{k_0=\omega_k}$$

$$\begin{aligned}H &= \frac{1}{2}\int \mathrm{d}^3x \left(\pi^2 + (\overline{\nabla})^2 m^2\varphi^2\right) \\ &= \frac{1}{2}\int \frac{\mathrm{d}^3x}{(2\pi)^3}\int \frac{\mathrm{d}^3k\,\mathrm{d}^3p}{\sqrt{4\omega_k\omega_p - \omega_k}} \times \Big\{(-i\omega_k)(-i\omega_p)\Big[a(k)a(p)e^{-i(k+p)x} + a^*(k)a^*(p)e^{i(k+p)x} - a(k)a^*(p)e^{-i(k-p)x} - \\ &= \frac{1}{2}\int \frac{\mathrm{d}^3k\,\mathrm{d}^3p}{\sqrt{4\omega_k\omega_p - \omega_k}}\int \frac{\mathrm{d}^3x}{(2\pi)^3}\Big\{(m^2 - \omega_k\omega_k - \overline{k}\cdot\overline{p})\Big[a(k)a(p)e^{-i(k+p)x} + a^*(k)a^*(p)e^{i(k+p)x}\Big] + \Big[a(k)a^*(p)e^{-i(k-\end{aligned}$$

Let's integrate $\int \mathrm{d}^3x$:

$$H = \frac{1}{2}\int \frac{\mathrm{d}^3k\,\mathrm{d}^3p}{\sqrt{4\omega_k\omega_p}}\Big\{\delta^3(\overline{k}+\overline{p})(m^2 - \omega_k\omega_p - \overline{k}\cdot\overline{p})\Big[a(k)a(p)e^{-i(k_0+p_0)x_0} + a^*(k)a^*(p)e^{+i(k_0+p_0)x_0}\Big] + \delta^3(\overline{k}-\overline{p})(m^2 +$$

Let's integrate $\int \mathrm{d}^3p$:

$$\begin{aligned}H &= \frac{1}{2}\int \frac{\mathrm{d}^3k}{2\omega_k}\left\{(m^2 - \omega_k^2 - |\overline{k}|^2)[\ldots] + (m^2 + \omega_k^2 + |\overline{k}|^2)\left[a(k)a^*(p) + a^*(k)a(p)\right]\right\} \\ &= \frac{1}{2}\int \mathrm{d}^3k\omega_k(a(k)a^*(k) + a^*(k)a(k))\end{aligned}$$

We didn't use any "commutator property"

$$\boxed{H = \int \mathrm{d}^3k\omega_k a^*(k)a(k)} \qquad \text{total energy}$$

$$P^i = \int \mathrm{d}^3k k^i(a(k)a^*(k) + a^*(k)a(k))$$

$$P^i = \int \mathrm{d}^3k k^i a^*(k)a(k) \qquad \text{total free momentum}$$

4.1.4 Quantization

Following the recipe for the second quantization (TODO ref), we get

$$\begin{cases} [\hat{\varphi}(\mathbf{x},t), \hat{\pi}(\mathbf{y},t)]_t = i\delta^3(\mathbf{x}-\mathbf{y}) \\ [\hat{\varphi}(\mathbf{x},t)\hat{\varphi}(\mathbf{y},t)]_t = 0 = [\hat{\pi}(\mathbf{x},t), \hat{\pi}(\mathbf{y},t)]_t \end{cases}$$

We have already shown that the solutions of the Klein-Gordon equation are of the form

$$\begin{cases} \hat{\varphi}(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k}} \left(\hat{a}(k)e^{-ikx} + \hat{a}^\dagger(k)e^{ikx}\right)_{k_o=\omega_k} \\ \hat{\pi}(x) = \frac{-i}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k}} \omega_k \left(\hat{a}(k)e^{-ikx} - \hat{a}^\dagger(k)e^{ikx}\right)_{k_o=\omega_k} \end{cases}$$

where the integral is over all wave vectors allowed by the boundary conditions. Alternatively, we can write in momentum space

$$\begin{cases} \hat{a}(p) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3x}{\sqrt{2\omega_p}} \left(\omega_p\varphi(\mathbf{x},t) + i\pi(\mathbf{x},t)\right) e^{ipx}\Big|_{p_o=\omega_p} \\ \hat{a}^\dagger(p) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3x}{\sqrt{2\omega_p}} \left(\omega_p\varphi(\mathbf{x},t) - i\pi(\mathbf{x},t)\right) e^{-ipx}\Big|_{p_o=\omega_p} \end{cases}$$

In momentum space we get an equivalent quantization condition:

$$\begin{cases} \left[a(p), a^\dagger(k)\right] = \delta^3(\mathbf{p}-\mathbf{k}) \\ [a(p), a(k)] = 0 = \left[a^\dagger(p), a^\dagger(k)\right] \end{cases}$$

These are the harmonic oscillator commutation relations for a continuous set of uncoupled harmonic oscillators.

4.1.5 Hamiltonian (density)

The Hamiltonian density operator

$$\mathcal{H}(\varphi,\pi) \equiv \frac{1}{2}\pi^2 + \frac{1}{2}(\overline{\nabla}\varphi)^2 + \frac{1}{2}m^2\varphi^2$$

The Hamiltonian operator

$$\begin{aligned} H &= \int \mathrm{d}^3x\mathcal{H} = \frac{1}{2}\int \mathrm{d}^3k\omega_k \left(a^\dagger(k)a(k) + a(k)a^\dagger(k)\right) \\ &= \int \mathrm{d}^3k\omega_k \left(a^\dagger(k)a(k) + \frac{1}{2}[a(k), a^\dagger(k)]\right) \\ &= \underbrace{\int \mathrm{d}^3k\omega_k\mathcal{N}(k)}_{\Delta E\text{energy operator}} + \cancel{\frac{1}{2}\int \mathrm{d}^3k\omega_k\delta^3(\overline{k}-\overline{k})} \end{aligned}$$

$$N[H] = \int \mathrm{d}^3k\omega_k\mathcal{N}(k)$$

The dropped term corresponds to the constant energy of vacuum ($E_0 \to \infty$). We drop it here because it is constant (i.e. a number, not an operator), but it becomes relevant when quantizing gravity.

4.1.6 Three momentum (density) operator

$$\begin{aligned} P_i &= \int \mathrm{d}^3x \pi(\partial_i\varphi) = \dots \\ &= \frac{1}{2}\int \mathrm{d}^3k(k_i)\left(a^\dagger(k)a(k) + a(k)a^\dagger(k)\right) \\ &= \int \mathrm{d}^3k k_i \mathcal{N}(k) + \cancel{\infty = \text{number}} \end{aligned}$$

$$N[P_i] = \int \mathrm{d}^3k k_i \mathcal{N}(k)$$

The switching round of $a(k)$ and $a^\dagger(k)$ in the previous calculations motivates us to make the following definition:

A product is normal ordered when all creation operators are to the left of all annihilation operators in the product. I.e. when all $a^\dagger$ are on the left and all $a(k)$ are on the right

$$N[a^\dagger(k)a(k) + a(k)a^\dagger(k)] = 2a^\dagger(k)a(k)$$

Inside $N[\,]$ (for scalar fields) everything commutes.

Normal ordering does not touch any H.O. commutators

4.1.7 General solution and interpretation

From now on we drop the hats over the operators to make the notation less heavy. We also split

$$\begin{aligned} \varphi(x) &= \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k}}\left(a(k)e^{-ikx} + a^\dagger(k)e^{ikx}\right)_{k_0=\omega_k} \\ &= \varphi_+(x) + \varphi_-(x) \end{aligned}$$

The operators $\varphi_\pm(x)$ are the annihilation / creation operators of a particle with only momentum k at the point x.
One-particle states are linear superpositions of

$$a^\dagger(\mathbf{k})\,|0\rangle \qquad \text{for any} \qquad \mathbf{k}$$

Two-particle states are linear superpositions of

$$a^\dagger(\mathbf{k})a^\dagger(\mathbf{k'})\,|0\rangle \qquad \text{for any} \qquad \mathbf{k}, \mathbf{k'} \neq \mathbf{k}$$

and

$$\frac{1}{\sqrt{2}}(a^\dagger(\mathbf{k}))^2\,|0\rangle \qquad \text{for any} \qquad \mathbf{k}$$

These states are normalized if the vacuum state is normalized.

Example

$$\begin{aligned}\langle 0|\left(\varphi_+(x)\,|1(p)\rangle\right) &= \frac{1}{(2\pi)^{3/2}}\int\frac{\mathrm{d}^3k}{\sqrt{2\omega_k}}(2\pi)^{3/2}\sqrt{2\omega_p}\langle 0|a(k)a^\dagger(p)|0\rangle e^{-ikx}\\ &= \int \mathrm{d}^3k\frac{\sqrt{2\omega_p}}{\sqrt{2\omega_k}}\delta^3(\overline{k}-\overline{p})e^{-ikx}\langle 0|0\rangle = e^{-ipx}\end{aligned}$$

(This example also hints at the reason why the "good" normalization is good)

4.1.8 Covariant commutators

So far the story has been:

1. We defined the Lagrangian density, $\mathcal{L}$, which is a covariant object.
2. We then switched to the Hamiltonian formalism, with $\mathcal{H} = \pi\partial_0\varphi - \mathcal{L}$. This is the Legendre transformation of the Lagrangian density. The Hamiltonian density $\mathcal{H}$ is not manifestly covariant. (Which is why we stressed that we compute it at fixed time)
3. We then applied the canonical quantization to get $[\varphi(\overline{x},t),\pi(\overline{y},t)] = i\delta^3(\overline{x}-\overline{y})$, which is not manifestly covariant.

So we need now to verify that the canonical quantization is consistent with Lorentz covariance. In other words, we need to show that the commutation relations are covariant. To do that, we write the commutator $[\varphi(x),\varphi(y)]$ for two arbitrary spacetime points x and y.
First we note that

$$[\varphi_+(x),\varphi_+(y)] = [\varphi_-(x),\varphi_-(y)] = 0.$$

So

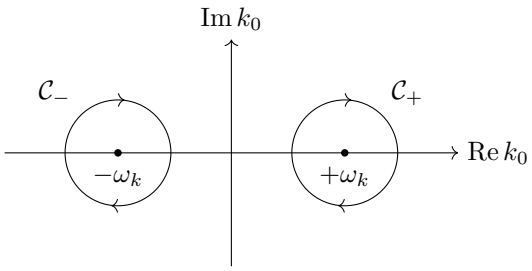
$$\begin{aligned}[\varphi(x),\varphi(y)] &= [\varphi_+(x)+\varphi_-(x),\varphi_+(y)+\varphi_-(y)]\\ &= [\varphi_+(x),\varphi_-(y)] + [\varphi_-(x),\varphi_+(y)]\\ &\equiv D_+(x,y) + D_-(x,y)\end{aligned}$$

Working out each part separately, we get

$$\begin{aligned}D_+(x,y) &\equiv [\varphi_+(x),\varphi_-(y)]\\ &= \frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3k}{\sqrt{2\omega_k}}\int\frac{\mathrm{d}^3k'}{\sqrt{2\omega_{k'}}}e^{-ikx}e^{ik'y}\left[a(k),a^\dagger(k')\right]\\ &= \frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3k}{2\omega_k}e^{-ik(x-y)}\Big|_{k_0=\omega_k}\end{aligned}$$

and

$$\begin{aligned}D_-(x,y) &\equiv [\varphi_-(x),\varphi_+(y)]\\ &= \frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3k}{\sqrt{2\omega_k}}\int\frac{\mathrm{d}^3k'}{\sqrt{2\omega_{k'}}}e^{ikx}e^{-ik'y}\left[a^\dagger(k),a(k')\right]\\ &= -\frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3k}{2\omega_k}e^{ik(x-y)}\Big|_{k_0=\omega_k}\end{aligned}$$



We notice that D_+ and D_- do not really depend on both x and y, but rather on their difference $x-y$. Thus we can define

$$\begin{aligned} D(x-y) &\equiv D_+(x-y) + D_-(x-y) \\ &= \frac{1}{(2\pi)^3}\int \frac{\mathrm{d}^3k}{2\omega_k}\left(e^{-ik(x-y)} - e^{ik(x-y)}\right)\Big|_{k_0=\omega_k} \\ &= -\frac{1}{(2\pi)^3}\int \frac{\mathrm{d}^3k}{\omega_k}\sinh(x-y)\Big|_{k_0=\omega_k} \end{aligned}$$

We can now write $D_\pm(x-y)$ in a covariant way:

$$D_\pm(x-y) \equiv i\int_{c_\pm} \frac{\mathrm{d}^4k}{(2\pi)^4}\frac{e^{-ik(x-y)}}{k^2-m^2}$$

Let's calculate $D_+(x-y)$

$$\begin{aligned} D_+(x-y) &= \frac{i}{(2\pi)^4}\int_{c_+}\mathrm{d}^4k\frac{e^{-ik(x-y)}}{k^2-m^2} \\ &= i\int \frac{\mathrm{d}^3k}{(2\pi)^3}e^{i\overline{k}(\overline{x}-\overline{y})}\cdot\int_{c_+}\frac{\mathrm{d}k_0}{2\pi}\frac{e^{ik_0(x_0-y_0)}}{(k_0-\omega_k)(k_0+\omega_k)} \end{aligned}$$

We define

$$f(k_0) = \frac{e^{ik_0(x_0-y_0)}}{2\pi(k_0-\omega_k)(k_0+\omega_k)}$$

Now, using the residue theorem

Residue theorem

$$\int_{c_+}\mathrm{d}k_0 f(k_0) = -2\pi i\,\mathrm{Res}(f(k_0))_{k_0=\omega_k}$$

we get

$$D_+(x-y) = i\int\frac{\mathrm{d}^3k}{(2\pi)^3}(-2\pi i)\frac{e^{-i\omega_k(x_0-y_0)}}{(2\pi)2\omega_k}e^{i\overline{k}(\overline{x}-\overline{y})} = \frac{1}{(2\pi)^3}e^{-ik(x-y)}\Big|_{k_0=\omega_k}$$

In conclusion the quantization conditions are consistent with covariance of the theory

$$D(x-y) = [\varphi(x),\varphi(y)]$$

Example

Exercise: $[\pi(x), \pi(y)], [\varphi(x), \pi(y)]$

4.1.9 Covariance and causality

- $D(x-y) \equiv [\varphi(x), \varphi(y)]$ is covariant.
- $D(x-y)|_{x_0=y_0} = [\varphi(x,t), \varphi(y,t)] = 0$

  So $D(x-y) = 0$ if $(x-y)^2 < 0$ (space line)

The creation / annihilation of a particle $(x_0, \overline{x})$ cannot be influenced by the creation / annihilation of a particle $(y_0, \overline{y})$ if $(x-y)^2 < 0$
Given A, B; if $[A, B] = 0$ there is no interference in the measurement
Microcausality.
Lecture 5/11
Recap Canonical quantization of complex scalar (free)

1. General form

$$\varphi(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{2\omega_k}() \equiv \varphi_+(x) + \varphi_-(x)$$

   ?

2. Hamiltonian formalism

$$\pi \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi} = \partial_0 \varphi^\dagger \qquad \rightarrow \qquad (\varphi, \pi)$$
$$\pi^\dagger \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi^\dagger} = \partial_0 \varphi \qquad \rightarrow \qquad (\varphi^\dagger, \pi^\dagger)$$
$$\mathcal{H} = \pi \partial_0 \varphi + \pi^\dagger \partial_0 \varphi^\dagger - \mathcal{L}$$

3. Canonical quantization

$$[\varphi(\overline{x}, t), \pi(\overline{y}, t)]_t = i\delta^3(\overline{x} - \overline{y}) = [\varphi^\dagger(\overline{x}, t), \pi^\dagger(\overline{y}, t)]$$
$$[\varphi(\overline{x}, t), \varphi(\overline{y}, t)] = 0 = [\varphi^\dagger(\overline{x}, t), \varphi^\dagger(\overline{y}, t)] = \ldots$$

   So equations of motion (Heisenberg picture)

$$\frac{\mathrm{d}\varphi^{(\dagger)}(\overline{x}, t)}{\mathrm{d}t} = -i[\varphi^{(\dagger)}(\overline{x}, t), H], \qquad \frac{\mathrm{d}\pi^{(\dagger)}(\overline{x}, t)}{\mathrm{d}t} = -i[\pi^{(\dagger)}(\overline{x}, t), H]$$

4. Quantization conditions in momentum space

$$[a(p), a^\dagger(k)] = \delta^3(\overline{p} - \overline{k}) \qquad = [b(p), p^\dagger(k)]$$
$$[a^{(\dagger)}(p), a^{(\dagger)}(k)] = 0 \qquad = [b^{(\dagger)}(p), b^{(\dagger)}(k)]$$

   So there are two sets of infinite ($\forall k$) H.O. (decoupled)

5. Observables

$$\begin{cases} \mathcal{N}_\dashv(k) \equiv a^\dagger(k)a(k) & N_a \equiv \int \mathrm{d}^3k \mathcal{N}_\dashv(k) \\ \mathcal{N}_\lfloor(k) \equiv b^\dagger(k)b(k) & N_b \equiv \int \mathrm{d}^3k \mathcal{N}_\lfloor(k) \end{cases}$$

$$N[H] \equiv \int \mathrm{d}^3x \mathcal{H} = \int \mathrm{d}^3k \omega_k \left(\mathcal{N}_a(k) + \mathcal{N}_b(k)\right) \qquad \geq 0$$

$$N[Q] \equiv q \int \mathrm{d}^3x J^0_{\mathrm{U}(1)} = \int \mathrm{d}^3k \left(q\mathcal{N}_a(k) + (-q)\mathcal{N}_b(k)\right) \qquad > \text{or} < 0$$

(New information of complex theory in charge + doubling of a and b)

6. Fock space for complex scalar field

$$|0\rangle \qquad \to \qquad \begin{cases} a(k)\,|0\rangle = 0 \\ b(k)\,|0\rangle = 0 \end{cases}$$

Any other state is obtained by applying $a^\dagger(k), b^\dagger(k)$

$$|n(k), \overline{n}(p)\rangle \propto (a^\dagger(k))^u (b^\dagger(p))^{\overline{u}}\,|0\rangle$$

$$\begin{cases} H\,|\,\rangle \propto (n\omega_k + \overline{n}\omega_p)\,|\,\rangle \\ Q\,|\,\rangle \propto (nq - \overline{n}q)\,|\,\rangle \end{cases}$$

How about the covariant commutators?

$$[\varphi(x), \varphi(y)] = \left[\varphi^\dagger(x), \varphi^\dagger(y)\right] = 0$$

because we only get

$$[a, a] \quad \text{or} \quad \left[b^\dagger, b^\dagger\right] \quad \text{or} \quad \left[a^\dagger, a^\dagger\right] \quad \text{or} \quad [b, b]$$

The only interesting commutator is

$$\begin{aligned} \left[\varphi(x), \varphi^\dagger(y)\right] &= \left[\varphi_+(x) + \varphi_-(x), \varphi_+^\dagger(y) + \varphi_-^\dagger(y)\right] \\ &= \left[\varphi_+(x), \varphi_-^\dagger(y)\right] + \left[\varphi_-(x), \varphi_+^\dagger(y)\right] = D_+(x-y) + D_-(x-y) \\ &= D(x-y) \end{aligned}$$

Which we've already shown to be covariant.

4.2 Complex Klein-Gordon field

We will now consider a field like the one in the previous section, except now we allow the field to be complex. This new field is very similar to the old one, except it admits charges.
So we start with the Lagrangian

$$\mathcal{L} = (\partial_\mu \varphi^*)(\partial^\mu \varphi) - m^2 \varphi^* \varphi$$

and the field has the following properties:

$$\begin{cases} \varphi^*(x) \neq \varphi(x) \\ \varphi'(x') = \varphi(x) \qquad \text{for Poincaré transformations} \end{cases}$$

The complex field can be split into two real fields φ_1 and φ_2.

$$\varphi = \frac{\varphi_1 + i\varphi_2}{\sqrt{2}}, \qquad \varphi^* = \frac{\varphi_1 - i\varphi_2}{\sqrt{2}}$$

So the Lagrangian becomes.

$$\begin{aligned}\mathcal{L} =& \frac{1}{2}(\partial_\mu \varphi_1)(\partial^\mu \varphi_1) - \frac{1}{2}m^2\varphi_1^2 \\ &+ \frac{1}{2}(\partial_\mu \varphi_2)(\partial^\mu \varphi_2) - \frac{1}{2}m^2\varphi_2^2\end{aligned}$$

Because φ_1 and φ_2 are real fields, the creation and annihilation operators associated with them cannot describe charged particles, only linear combinations of them can. For this reason it is often more natural to work directly with φ and φ^*

4.2.1 Equation of motion for φ and φ^*

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \varphi^*} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi^*} &= -m^2\varphi - \Box\varphi = 0 \\ \frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi} &= -m^2\varphi^* - \Box\varphi^* = 0\end{aligned}$$

So

$$(\Box + m^2)\varphi = 0 = (\Box + m^2)\varphi^*$$

4.2.2 Hamiltonian formulation

$$\begin{cases}\pi(x) \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi(x)} = \partial_0 \varphi^*(x) & \rightarrow \qquad (\pi, \varphi) \equiv (\partial_0 \varphi^*, \varphi) \\ \pi^*(x) \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi^*(x)} = \partial_0 \varphi(x) & \rightarrow \qquad (\pi^*, \varphi^*) \equiv (\partial_0 \varphi, \varphi^*)\end{cases}$$

So

$$\begin{aligned}\mathcal{H} &= \pi \partial_0 \varphi + \pi^* \partial_0 \varphi^* - \mathcal{L} \\ &= \pi^*\pi + (\nabla\varphi^*)(\nabla\varphi) + m^2\varphi^*\varphi\end{aligned}$$

Again the Hamiltonian density is positive.

$$H = \int \mathrm{d}^3x \mathcal{H} \quad \geq 0$$

Example

Hamilton equation of motion

$$\dot{\varphi} = \{\varphi, H\}, \qquad \dot{\varphi}^* = \{\varphi^*, H\}, \qquad \dot{\pi} = \{\pi, H\}, \qquad \dot{\pi}^* = \{\pi^*, H\}$$

Remember condensed notation!

$$\{F(\pi,\varphi),G(\pi,\varphi)\}=\sum_{i=1}^{N}\int \mathrm{d}^3x\left(\frac{\delta F}{\delta\varphi_i}\frac{\delta G}{\delta\pi_i}-\frac{\delta F}{\delta\pi_i}\frac{\delta G}{\delta\varphi_i}\right)$$

4.2.3 Conserved currents / charges

4.2.3.1 Translations

$$\begin{aligned}J^{\mu}_{(\nu)} &= \text{“ }\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\nu\varphi-\eta^{\mu}_{\nu}\mathcal{L}\text{ ”}=\sum_i\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi_i}-\eta^{\mu}_{\nu}\mathcal{L}\\ &=\frac{\partial_\mu\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\nu\varphi+\frac{\partial_\mu\partial\mathcal{L}}{\partial\partial_\mu\varphi^*}\partial_\nu\varphi^*-\eta^{\mu}_{\nu}\mathcal{L}\equiv\tilde{T}^{\mu}_{\nu}\\ &=(\partial^\mu\varphi^*)(\partial_\nu\varphi)+(\partial^\mu\varphi)(\partial_\nu\varphi^*)-\eta^{\mu}_{\nu}\mathcal{L}\end{aligned}$$

$$\begin{cases}P_0=\int\mathrm{d}^3x\tilde{T}^0_0=\int\mathrm{d}^3x\mathcal{H}\\ P_i=\int\mathrm{d}^3x\tilde{T}^0_i=\int\mathrm{d}^3x\left[(\partial_0\varphi^*)(\partial_i\varphi)+(\partial_0\varphi)(\partial_i\varphi^*)\right]\end{cases}$$

4.2.4 U(1) global symmetry

Charge!!
We are still considering the following Lagrangian density:

$$\mathcal{L}=(\partial_\mu\varphi^*)(\partial^\mu\varphi)-m^2\varphi^*\varphi$$

We are now going to apply a U(1) transformation

$$\begin{cases}\varphi'(x)=e^{i\alpha}\varphi(x)\approx(1+i\alpha1)\varphi\\ \varphi'^*(x)=e^{-i\alpha}\varphi^*(x)\approx(1-i\alpha1)\varphi\end{cases}$$

This does change the Lagrangian.

$$\begin{aligned}\mathcal{L}' &=(\partial_\mu\varphi'^*)(\partial^\mu\varphi')-m^2\varphi^{*\prime}\varphi'\\ &=e^{-i\alpha}e^{i\alpha}(\partial_\mu\varphi^*)(\partial^\mu\varphi)-m^2e^{-i\alpha}e^{i\alpha}\varphi^*\varphi=\mathcal{L}\end{aligned}$$

Thus there is a U(1) global internal symmetry of the theory.

4.2.4.1 Conserved current and charge

$$\begin{cases}x'^\mu=x^\mu\\ \varphi'(x)=(1+i\alpha)\varphi(x)\\ (\varphi'^*(x)=(1-i\alpha)\varphi^*(x))\end{cases}\quad\rightarrow\quad\begin{cases}\delta x^\mu=0\quad\rightarrow\quad\Xi=0\\ \delta_0\varphi=i\alpha\varphi\quad\rightarrow\quad x_\varphi=i\varphi\\ (\delta_0\varphi^*=-i\alpha\varphi^*\quad\rightarrow\quad x_{\varphi^*}=-i\varphi^*)\end{cases}$$

$$\begin{aligned}J^{\mu}_{\mathrm{U}(1)} &=\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}x_\varphi+\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi^*}x_{\varphi^*}\\ &=i(\partial_\mu\varphi^*)\varphi-i\varphi^*(\partial_\mu\varphi)=i\varphi^*\overleftrightarrow{\partial}_\mu\varphi\end{aligned}$$

$$Q_{\mathrm{U}(1)} \equiv q \int \mathrm{d}^3 x i \left(\varphi^* \overleftrightarrow{\partial}_0 \varphi\right)$$

Example

Exercise: For a complex field φ write P^μ, Q in the momentum space.

$$\varphi(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 x}{\sqrt{2\omega_k \delta^3}} \left(a(k) e^{-ikx} + b^*(k) e^{ikx}\right)_{k_0=\omega_k}$$
$$P^\mu = \int \mathrm{d}^3 k k^\mu \left(a^*(k) a(k) + b^*(k) b(k)\right)$$
$$Q_{\mathrm{U}(1)} = \int \mathrm{d}^3 k \left(q a^*(k) a(k) - q b^*(k) b(k)\right)$$

4.3 Dirac spinor field

As before we first guess the Lagrangian! The form

$$\mathcal{L}_D = \frac{i}{2} \left[\overline{\psi} \gamma^\mu (\partial_\mu \psi) - (\partial_\mu \overline{\psi}) \gamma^\mu \psi\right] - m \overline{\psi} \psi$$

seems promising (we want it to be linear in derivatives). We check dimensionality. Remembering that $[\mathcal{L}_D] = M^4$ and using $[m] = M^1$, we get

$$[\overline{\psi}\psi] = M^{4-1} \qquad \rightarrow \qquad [\psi] = M^{3/2}$$

Because $[\partial_\mu] = L^{-1} = M$ the dimensionality checks out.

$$\psi' = S(\Lambda) \rightarrow \mathcal{L} \textbf{ is Lorentz invariant}$$

$$\mathcal{L}_D \textbf{ is real (Hermitian)}$$

4.3.1 Euler-Lagrange equations of motion

Applying the Euler-Lagrange equations we firstly get

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \overline{\psi})} = \frac{i}{2} \gamma^\mu (\partial_\mu \psi) - m\psi + \frac{i}{2} \gamma^\mu \partial_\mu \psi = 0$$

Which gives us

$$\boxed{(i\overline{\partial\!\!\!/} - m)\psi = 0}$$

And secondly

$$\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} = -\frac{i}{2} \gamma^\mu (\partial_\mu \overline{\psi}) - m\overline{\psi} - \frac{i}{2} (\partial_\mu \overline{\psi}) \gamma^\mu = 0$$

$$\boxed{\overline{\psi}(\overleftarrow{\partial\!\!\!/} + m) = 0}$$

The usual $\mathcal{L}'_D$

$$\mathcal{L}'_D = \overline{\psi}(i\overrightarrow{\partial\!\!\!/} - m)\psi$$

Example

Exercise: Show that $\mathcal{L}_D$ is equivalent to $\mathcal{L}'_D$ (same eq of motion)
Exercise: Derive the EL eqs from $\mathcal{L}'_D$

4.3.2 General solution of Dirac equation

$$\begin{aligned}
\psi(x) &= \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k}} \sum_{r=1}^{2} \left(c_r(k) u_r(k) e^{-ikx} + d_r^\dagger(k) v_r(k) e^{ikx} \right)_{k_0=\omega_k} \\
&= \psi_+(x) + \psi_-(x) \\
\overline{\psi}(x) &= \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{\sqrt{2\omega_k}} \sum_{r=1}^{2} \left(d_r(k) \overline{v}_r(k) e^{-ikx} + c_r^\dagger(k) \overline{u}_r(k) e^{ikx} \right)_{k_0=\omega_k} \\
&= \overline{\psi}_+(x) + \overline{\psi}_-(x)
\end{aligned}$$

4.3.3 Conserved charges

4.3.3.1 Invariance under translation

$$J^\mu_{(\nu)} \equiv \tilde{T}^\mu_{(\nu)} = \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} (\partial_\nu \psi) + (\partial_\nu \overline{\psi}) \frac{\partial \mathcal{L}}{\partial \partial_\mu \overline{\psi}} - \underbrace{\cancel{\eta^\mu_\nu \mathcal{L}}}_{\text{by eq of motion}}$$

$$= \frac{i}{2} \left[\overline{\psi} \gamma^\mu (\partial_\nu \psi) - (\partial_\nu \overline{\psi} \gamma^\mu \psi) \right]$$

$$P_\nu = \int \mathrm{d}^3x J^0_{(\nu)} = \frac{i}{2} \int \mathrm{d}^3x (\psi^\dagger \overleftrightarrow{\partial_\nu} \psi)$$

$$\begin{cases} H \equiv P_0 = \frac{i}{2} \int \mathrm{d}^3x [\psi^\dagger (\partial_0 \psi) - (\partial_0 \psi^\dagger) \psi] \\ H' \equiv i \int \mathrm{d}^3x \psi^\dagger (\partial_0 \psi) \end{cases}$$

4.3.3.2 Invariance under rotations

$$x'^\mu = x^\mu + \omega^\mu_{\ \nu} x_\nu \qquad \rightarrow \qquad \delta x^\mu = \frac{1}{2} \omega^{\rho\sigma} \Xi^\mu_{(\rho\sigma)}$$

$$\begin{cases} \psi'(x') = \left(\mathbb{1} - \frac{i}{2} \omega^{\rho\sigma} \Sigma_{\rho\sigma} \right) \psi(x) \\ \overline{\psi}(x') = \overline{\psi}(x) \left(\mathbb{1} + \frac{i}{2} \omega^{\rho\sigma} \Sigma_{\rho\sigma} \right) \end{cases} \qquad \rightarrow \qquad \begin{cases} \delta\psi = \frac{1}{2} \omega^{\rho\sigma} X_{(\rho\sigma)} \\ \delta\overline{\psi} = \frac{1}{2} \omega^{\rho\sigma} \overline{X}_{(\rho\sigma)} \end{cases}$$

$$\begin{cases} X_{\rho\sigma} = i \Sigma_{\rho\sigma} \psi \qquad \overline{X}_{\rho\sigma} = -i \overline{\psi} \Sigma_{\rho\sigma} \\ \Xi^\mu_{\rho\sigma} = \left(\eta^\mu_\rho x_\sigma - \eta^\mu_\sigma x_\rho \right) \end{cases}$$

$$J^\mu_{(\rho\sigma)} = x_\rho \tilde{T}^\mu_{\ \sigma} - x_\sigma \tilde{T}^\mu_{\ \rho} + \left[\overline{\psi} \gamma^\mu \Sigma_{\rho\sigma} \psi \right] = L^\mu_{\rho\sigma} + S^\mu_{\rho\sigma} = M^\mu_{\rho\sigma}$$

$$J_{\rho\sigma} = \int \mathrm{d}^3x J^0_{(\rho\sigma)} = \int \mathrm{d}^3x \left\{ \underbrace{(x_\rho T_\sigma - x_\sigma T_\rho)}_{L_{\rho\sigma}} + \psi^\dagger \Sigma_{\rho\sigma} \psi \right\}$$
$$S_{\rho\sigma}$$

4.3.3.3 U(1) global symmetry

$$\begin{cases} x'^{\mu} = x^{\mu} \\ \psi'(x) = e^{i\alpha}\psi(x) \end{cases} \quad \rightarrow \quad \begin{cases} \Xi^{\mu} = 0 \\ X = i\psi\mathbb{1} \end{cases}$$

$$J^{\mu}_{\mathrm{U}(1)} = \frac{\partial \mathcal{L}}{\partial \partial_{\mu}\psi} X_{\psi} + \overline{X}_{\psi} \frac{\partial \mathcal{L}}{\partial \partial_{\mu}\overline{\psi}} = \overline{\psi}\gamma^{\mu}\psi$$

$$Q_{\mathrm{U}(1)} = q \int \mathrm{d}^3x J^0_{\mathrm{U}(1)} = \int \mathrm{d}^3x \psi^{\dagger}\psi \qquad \rightarrow \qquad \frac{\mathrm{d}Q}{\mathrm{d}t} = 0$$

4.3.4 Hamiltonian description

$$\begin{cases} \pi \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \psi} = \frac{i}{2}\psi^{\dagger} = \frac{i}{2}\overline{\psi} & (\psi, \pi) = (\psi, \psi^{\dagger}) \\ \pi^{\dagger} \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 \psi^{\dagger}} = -\frac{i}{2}\psi & (\psi^{\dagger}, \pi^{\dagger}) = (\psi^{\dagger}, -\psi) \end{cases}$$

$$\mathcal{H} = \pi\partial_0\psi + \partial_0\psi^{\dagger}\pi^{\dagger} - \mathcal{L}$$

$$\mathcal{L}' = \overline{\psi}(i\partial\!\!\!/ - m)\psi$$

$$\begin{cases} \pi^{\dagger} \equiv \frac{\partial \mathcal{L}'}{\partial \partial_0 \psi} = i\psi^{\dagger} \\ \pi^{\dagger\prime} \equiv \frac{\partial \mathcal{L}'}{\partial \partial_0 \psi^{\dagger}} = 0 \end{cases}$$

$$\mathcal{H}' = \pi'\partial_0\psi - \mathcal{L}'$$

Lecture 6/11 Recap Classical Dirac Field Theory

$$\begin{aligned} \mathcal{L}_D &= \frac{i}{2}\left[\overline{\psi}\gamma^{\mu}(\partial_{\mu}\psi) - (\partial_{\mu}\overline{\psi}\gamma^{\mu}\psi)\right] - m\overline{\psi}\psi \\ &= \overline{\psi}\overleftrightarrow{\partial\!\!\!/}\psi - m\overline{\psi}\psi \end{aligned}$$

Simplified Lagrangian

$$\mathcal{L}'_D = \overline{\psi}(i\overrightarrow{\partial\!\!\!/} - m)\psi$$

$$(i\partial\!\!\!/ - m)\psi = 0$$

Hamiltonian Formalism

$$\pi \equiv \frac{\partial \mathcal{L}'}{\partial \partial_0 \psi} = i\psi^{\dagger}$$

$$\begin{aligned} \mathcal{H}' &= \pi\partial_0\psi - \mathcal{L}' = \cancel{i\psi^{\dagger}\partial_0\psi} - \cancel{i\psi^{\dagger}\partial_0\psi} - i\psi^{\dagger}\alpha_i\partial_i\psi + m\overline{\psi}\psi \\ &= \psi^{\dagger}(H_D\psi) = i\psi^{\dagger}\partial_0\psi \end{aligned}$$

4.3.5 Poisson brackets

$$\begin{cases} \{\psi_{\alpha}(\overline{x}, t), \pi_{\beta}(\overline{y}, t)\} = \delta^3(\overline{x} - \overline{y})\delta_{\alpha\beta} \\ \{\psi_{\alpha}, \psi_{\beta}\} = 0 = \{\pi_{\alpha}, \pi_{\beta}\} \end{cases}$$

4.3.6 Canonical quantization (commutators)

$$\begin{aligned} (\psi, \pi) \quad &\rightarrow \quad (\hat{\psi}, \hat{\pi}) \\ \{\ ,\ \} \quad &\rightarrow \quad -i[\ ,\ ] \end{aligned}$$

Example

Guided exercise

1. Quantization condition ($\pi = i\psi^\dagger$)

$$[\psi_\alpha(\overline{x},t),\pi_\beta(\overline{y}t)] = i\delta(\overline{x}-\overline{y})\delta_{\alpha\beta}$$
$$[\psi_\alpha(\overline{x},t),\psi_\beta(\overline{y},t)] = 0 = [\pi_\alpha(\overline{x},t),\pi_\beta(\overline{y},t)]$$

2. Annihilation and creation ops:

$$\begin{cases} c_r(k) = \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}^3x}{\sqrt{2\omega_k}}\overline{u}_r(k)\gamma_0\psi(x)e^{ikx} \\ c_r^\dagger(k) = \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}^3x}{\sqrt{2\omega_k}}\overline{\psi}(x)\gamma_0 u_r(k)e^{-ikx} \end{cases}$$
$$\begin{cases} d_r(k) = \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}^3x}{\sqrt{2\omega_k}}\overline{\psi}(x)\gamma_0 v_r(k)e^{ikx} \\ d_r^\dagger(k) = \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}^3x}{\sqrt{2\omega_k}}\overline{v}_r(k)\gamma_0\psi(x)e^{-ikx} \end{cases}$$

3. Quantization condition in momentum space

$$\begin{cases} \left[c_r(k),c_s^\dagger(p)\right] = \delta^3(\overline{k}-\overline{p})\delta_{rs} \\ [c_r(k),c_s(p)] = 0 = \left[c_r^\dagger(k),c_s^\dagger(p)\right] \end{cases}$$
$$\begin{cases} \left[d_r(k),d_s^\dagger(p)\right] = \delta^3(\overline{k}-\overline{p})\delta_{rs} \\ [d_r(k),d_s(p)] = 0 = \left[d_r^\dagger(k),d_s^\dagger(p)\right] \end{cases}$$

4. Operator

$$\mathcal{N}_c^r(k) \equiv c_r^\dagger(k)c_r(k), \qquad \mathcal{N}_d^r(k) \equiv d_r^\dagger(k)d_r(k)$$
$$\left[\mathcal{N}_c^r(k),c_s^{(\dagger)}(p)\right] = \pm c_s^{(\dagger)}(p)\delta^3(\overline{k}-\overline{p})\delta_{rs}$$

$$\begin{cases} H = \int \mathrm{d}^3x\, i\psi^\dagger\partial_0\psi = \int \mathrm{d}^3k\,\omega_k \sum_{r=1}^2 \left(c_r^\dagger(k)c_r(k) - d_r(k)d_r^\dagger(k)\right) \\ N[H] = \int \mathrm{d}^3k\,\omega_k \sum_{r=1}^2 \left(\mathcal{N}_c^r(k) \underbrace{-}_{!!} \mathcal{N}_d^r\right) > \text{ or } < 0 \end{cases}$$
$$\begin{cases} Q_{\mathrm{U}(1)} = q\int \mathrm{d}^3x\,\psi^\dagger\psi = q\int \mathrm{d}^3k \sum_{r=1}^2 \left(c_r^\dagger(k)c_r(k) + d_r(k)d_r^\dagger(k)\right) \\ N[Q_{\mathrm{U}(1)}] = q\int \mathrm{d}^3k \sum_{r=1}^2 \left(\mathcal{N}_c^r(k) + \mathcal{N}_d^r\right) > 0 \end{cases}$$

5. Fock space

$$\begin{cases} c_r(k)\,|0\rangle \\ d_r(k)\,|0\rangle \end{cases}$$
$$|n(k),\overline{n}(p)\rangle \propto \left(c_r^\dagger(k)\right)^n \left(d_r^\dagger(p)\right)^{\overline{n}}|0\rangle$$

States with $n > 1$ identical particel (problem 2!!!!) $\rightarrow$ bosonic particle with spin $1/2$?

4.3.7 Canonical quantization with anticommutators

$$\begin{array}{ccc} (\psi,\pi) & \to & (\hat{\psi},\hat{\pi}) \\ \{\ ,\ \}_{PB} & \to & -i\{\ ,\ \}_{AC} \end{array}$$

4.3.7.1 Quantization conditions

$$\begin{cases} \{\psi_\alpha(\overline{x},t),\pi_\beta(\overline{y},t)\} = i\delta^3(\overline{x}-\overline{y})\delta_{\alpha\beta} \\ \{\psi_\alpha(\overline{x},t),\psi_\beta(\overline{y},t)\} = 0 = \{\pi_\alpha(\overline{x},t),\pi_\beta(\overline{y},t)\} \end{cases}$$

$$\begin{cases} \left\{c_r(k),c_s^\dagger(p)\right\} = \delta(\overline{k}-\overline{p})\delta_{rs} = \left\{d_r(k),d_s^\dagger(p)\right\} \\ \{c_r(k),c_s(p)\} = 0 = \dots \end{cases}$$

4.3.7.2 Commutation relations with N_c, N_d

$$\begin{aligned} [\mathcal{N}_c^r(k),c_s(p)] &= c_r^\dagger(k)\{c_r(k),c_s(p)\} - \left\{c_r^\dagger(k),c_s(p)\right\}c_r(k) \\ &= -c_r(k)\delta^3(\overline{k}-\overline{p})\delta_{rs} \\ \left[\mathcal{N}_c^r(k),c_s^\dagger(p)\right] &= c_r^\dagger(k)\delta^3(\overline{k}-\overline{p})\delta_{rs} \end{aligned}$$

4.3.7.3 Operators

$$\begin{aligned} H &= \int \mathrm{d}^3k\omega_k \sum_{r=1}^{2}\left(c_r^\dagger(k)c_r(k) - d_r(k)d_r^\dagger(k)\right) \\ &= int\,\mathrm{d}^3k\sum_{r=1}^{2}\left(\mathcal{N}_c^r(k)+\mathcal{N}_d^r(k)\right) + \underbrace{\cancel{\int \mathrm{d}^3k\omega_k\delta(0)}}_{=\infty\text{vacuum energy}} \end{aligned}$$

$$N[H] = \int \mathrm{d}^3k\omega_k\sum_{r=1}^{2}\left(\mathcal{N}_c^r(k)+\mathcal{N}_d^r(k)\right)$$

$$Q_{\mathrm{U}(1)} = \int \mathrm{d}^3k\sum_{r=1}^{2}\left(qc_r^\dagger(k)c_r(k) + qd_r(k)d_r^\dagger(k)\right)$$

$$N[Q_{\mathrm{U}(1)}] = \int \mathrm{d}^3k\sum_{r=1}^{2}\left(q\mathcal{N}_c^r(k) - q\mathcal{N}_d^r(k)\right)$$

When I have fermionic operators I have to count the number of permutations $dd^\dagger$ have for going in N.O.

$$N[d^\dagger d] = d^\dagger d, \qquad N[dd^\dagger] = -d^\dagger d$$

4.3.7.4 Fock space for fermions

$$\exists\,|0\rangle \quad \rightarrow \quad \begin{cases} c_r(k)\,|0\rangle = 0 \\ d_r(k)\,|0\rangle = 0 \end{cases} \quad \rightarrow \quad \begin{cases} \mathcal{N}_c^r(k)\,|0\rangle = 0 \\ \mathcal{N}_d^r(k)\,|0\rangle = 0 \end{cases}$$

The one particle states

$$|1_r(p)\rangle \propto c_r^\dagger(p)\,|0\rangle$$
$$\left|\overline{1}_r(p)\right\rangle \propto d_r^\dagger(p)\,|0\rangle$$

$$\begin{cases} N_c^s\,|1_r(p)\rangle \propto \int \mathrm{d}^3k \mathcal{N}_c^s(k) c_r^\dagger(p)\,|0\rangle = +1\delta_{rs}\,|1_r(p)\rangle \\ H\,|1_r(p)\rangle = \omega_p\,|1_r(p)\rangle \quad \rightarrow \quad \text{particle } +q \\ Q_{\mathrm{U}(1)}\,|1_r(p)\rangle = +q\,|1_r(p)\rangle \end{cases}$$

$$\begin{cases} N_d^s\left|\overline{1}_r(p)\right\rangle = +1\delta_{rs}\left|\overline{1}_r(p)\right\rangle \\ H\left|\overline{1}_r(p)\right\rangle = \omega_p\left|\overline{1}_r(p)\right\rangle \quad \rightarrow \quad \text{antiparticle } -q \\ Q_{\mathrm{U}(1)}\left|\overline{1}_r(p)\right\rangle = -q\left|\overline{1}_r(p)\right\rangle \end{cases}$$

Two particle state

$$\begin{aligned} |2_r(p)\rangle &\propto c_r^\dagger(p) c_r^\dagger(p)\,|0\rangle \\ &= \frac{1}{2}\left\{c_r^\dagger(p), c_r^\dagger(p)\right\}|0\rangle = 0 \end{aligned}$$

So states with identical particles do not exist. Dirac field theory has Fermi-Dirac statistic!

Spin-statistic theorem

1. Particles with integer spin are bosons.
2. Particles with semi-integer spin are fermions.

Example

Exercise: Show that scalar field theory is inconsistent if quantified with {}

4.3.8 Covariant anticomutators

$$\begin{aligned} S_{\alpha\beta}(x-y) &\equiv \left\{\psi_\alpha(x), \overline{\psi}_\beta(y)\right\} \\ &= \left\{\psi_{\alpha,+}(x), \overline{\psi}_{\beta,-}(y)\right\} + \left\{\psi_{\alpha,-}(x), \overline{\psi}_{\beta,+}(y)\right\} \\ &= S_+^{\alpha\beta}(x-y) + S_-^{\alpha\beta}(x-y) \end{aligned}$$

$$S_+(x-y) = \left\{\psi_+(x), \overline{\psi}_-(y)\right\} = \frac{1}{(2\pi)^3}\int \frac{\mathrm{d}^3k}{2\omega_k} \underbrace{\sum_{r=1}^{2} u_r(k)\overline{u}_r(k)}_{(\not{k}+m)}\, e^{-ik(x-y)}$$

$$S_-(x-y) = \left\{\psi_-(x), \overline{\psi}_+(y)\right\} = \frac{1}{(2\pi)^3}\int \frac{\mathrm{d}^3k}{2\omega_k} \underbrace{\sum_{r=1}^{2} v_r(k)\overline{v}_r(k)}_{(\not{k}-m)}\, e^{ik(x-y)}$$

$$S_+(x-y) = (i\partial\!\!\!/_x + m)\frac{1}{(2\pi)^3}\int \frac{\mathrm{d}^3k}{2\omega_k} e^{-ik(x-y)} \qquad D_+$$

$$S_-(x-y) = -(i\partial\!\!\!/_x + m)\frac{1}{(2\pi)^3}\int \frac{\mathrm{d}^3k}{2\omega_k} e^{ik(x-y)} \qquad D_-$$

$$\boxed{S(x-y)} = (i\partial\!\!\!/ + m)D(x-y)$$

$$\overline{\psi}S(x-y)\psi = \text{scalar}$$

$$S^{-1}\gamma^\mu S\Lambda_\mu^{\ \nu} = \gamma^\nu$$

Microcausal

4.4 Relativistic vector field (free)

We now generalise these prcedures to vector fields.

4.4.1 Real vector fields

$$\Phi = \begin{pmatrix}\varphi_1\\ \vdots \\ \varphi_N\end{pmatrix} \qquad \rightarrow \qquad \mathcal{L}_{\mathbb{R}} = \frac{1}{2}(\partial_\mu \Phi^\intercal)(\partial^\mu \Phi) - M\Phi^\intercal\Phi$$

We consider an O(2) transformation of a two dimensional real vector field.

$$\Phi = \begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} \qquad \Phi' = O_2\Phi \qquad O_2 \in \mathrm{O}(2)$$

gives

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \Phi^\intercal)O^\intercal O(\partial^\mu \Phi) - M\Phi^\intercal O^\intercal O\Phi = \mathcal{L}$$

Thus this is an internal symmetry of the theory.

$$\mathrm{O}(2) \approx \mathrm{SO}(2) \qquad \rightarrow \mathfrak{o}(2) = \{1\text{generator}\} \approx \mathfrak{u}(1)$$

$$N\text{real vector field} \rightarrow \mathrm{O}(N)$$
$$N\text{complex vector field} \rightarrow \mathrm{U}(N)$$

4.4.2 Complex vector field

$$\mathcal{L}_{\mathbb{C}} = \frac{1}{2}(\partial_\mu \Phi^\dagger)(\partial^\mu \Phi) - M\Phi^\dagger\Phi$$

Vector (spin 1) Transforms like vectors:

$$V^\mu(x) \qquad \rightarrow \qquad V^{\mu\prime}(x') = \Lambda^\mu_{\ \nu}V^\nu(x)$$

Vector fields transform under defining representation. What is a vector?

$$V^\mu(x) \equiv (V^0, \overline{V})$$

Potential + 3D vector = 4 d.o.f.

Spin 1 massive particle has 3 d.o.f.

Spin 1 massless (like photon) particle has 2 d.o.f. (like polerizations)

4.4.3 Massive vector field ($M \neq 0$) Real

As always we first guess the Lagrangian density using $V^\mu, \partial^\mu V^\nu$. We want a quadratic equation of motion (we want to be able to recover Maxwell when $M \to 0$)

$$\mathcal{L} = -\frac{1}{4} V^{\mu\nu} V_{\mu\nu} + \frac{1}{2} M^2 V^\mu V_\mu$$

Where $V^{\mu\nu} = \partial^\mu V^\nu - \partial^\nu V^\mu$ is the field strength. This structure is faced by $H > 0$

4.4.3.1 Equation of motion

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial V^\sigma} - \partial_\rho \frac{\partial \mathcal{L}}{\partial \partial_\rho V_\sigma} &= 0 \\ &= M^2 V^\sigma + \partial_\rho \partial^\rho V^\sigma - \partial^\sigma \partial_\rho V^\rho \\ &= \boxed{(\Box + M^2) V^\sigma - \partial^\sigma (\partial_\rho V^\rho) = 0} \qquad \text{Proca eq.} \end{aligned}$$

$$\frac{M}{2} \frac{\partial}{\partial V_\sigma} V^\mu V_\mu = M^2 V^\mu \frac{\partial V_\mu}{\partial V_\sigma} = M^2 V^\mu \delta_{\mu\sigma} M^2 V_\sigma$$

I missed rest of his "technical derivation" (1 eq)

$$\partial_\sigma \left[(\Box + M^2) V^\sigma - \partial^\sigma \partial_\rho V^\rho \right] = 0$$

$$\cancel{\Box(\partial_\sigma V^\sigma)} + M^2 (\partial_\sigma V^\sigma) - \cancel{\Box(\partial_\rho V^\rho)} = 0 \qquad \to \qquad \partial_\sigma V^\sigma = 0$$

So the Proca equation is equivalent with

$$\begin{cases} (\Box + M^2) V^\sigma = 0 \qquad \to \qquad \text{Klein-Gordon equation} \\ \partial_\sigma V^\sigma \qquad \to \qquad \text{Constraint} \end{cases}$$

4.4.3.2 General solution of Proca equation

$$V^\mu(x) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \left(f^\mu(k) e^{-ikx} + f^{\mu *}(k) e^{ikx} \right)$$

$$(\Box + M^2) V^\mu(x) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4} (-k^2 + M^2) \left(f^\mu(k) e^{-ikx} + f^{\mu *}(k) e^{ikx} \right) = 0$$

Let's define

$$f^\mu(k) \equiv (2\pi)^{5/2} \sqrt{2\omega_k} \delta(k^2 - M^2) \sum_{\lambda=0}^{3} \epsilon_\lambda^\mu(k) a_\lambda(k)$$

Where $\epsilon_\lambda^\mu(k)$ are polerization vectors for $\lambda = 0, 1, 2, 3$. So we get

$$V^\mu(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \sum_{\lambda=0}^{3} \left(\epsilon_\lambda^\mu(k) a_\lambda(k) e^{-ikx} + \epsilon_\lambda^{\mu *}(k) a_\lambda^*(k) e^{ikx} \right)_{k_0 = \omega_k}$$

And we still need the condition $\partial_\nu V^\mu(x)$, which translates to $k_\mu \epsilon_\lambda^\mu(k) = 0$. This effectively removes one independent polerization, so we have 3 independent polerization vectors.

Example

Exercise: Given $k^\mu = (\omega_k, 0, 0, k)$, prove that we can define the following basis of ϵ^μ_λ that satisfy the Proca eq.

$$\begin{cases} \epsilon^\mu_{(1)} = (0,1,0,0), \quad \epsilon^\mu_{(2)} = (0,0,1,0) & \rightarrow \quad \text{transverse polerizations} \\ \epsilon^\mu_{(1)} = (\frac{k}{M}, 0, 0, \frac{\omega_k}{M}) & \rightarrow \quad \text{Longitudinal polerizations} \end{cases}$$

Example

Exercise: Prove the following relations:

$$\begin{cases} \epsilon^\mu_{(\lambda)} \epsilon_\mu^{(\lambda')} = -\delta_{\lambda\lambda'} & \rightarrow \quad \text{orthogonal} \\ \sum_{\lambda=1}^3 \epsilon^\mu_{(\lambda)} \epsilon^\nu_{(\lambda)} = -\left(\eta^{\mu\nu} - \frac{k^\mu k^\nu}{M^2}\right) & \rightarrow \quad \text{Completeness} \end{cases}$$

Example

Ex: Conserved charges associated to inv translations.
Ex: Conserved charges inv under Lorentz group

$$\begin{cases} x'^\mu = x^\mu + \omega^\mu_{\ \nu} x^\nu \\ V'^\mu(x') = V^\mu(x) + \omega^\mu_{\ \nu} V^\nu(x) \end{cases} \rightarrow \quad \begin{cases} \delta x^\mu = \frac{1}{2}\omega^{\rho\sigma} \Xi^\mu_{(\rho\sigma)} \\ \delta V^\mu = \frac{1}{2}\omega^{\rho\sigma} X^\mu_{(\rho\sigma)} \end{cases}$$

Where

$$\begin{cases} \Xi^\mu_{(\rho\sigma)} = -i \left(\mathcal{J}^{\mu\nu}\right)_{\rho\sigma} x_\nu \\ X^\mu_{(\rho\sigma)} = -i \left(\mathcal{J}^{\mu\nu}\right)_{\rho\sigma} V_\nu(x) \end{cases} \qquad \left(\mathcal{J}^{\mu\nu}\right)_{\rho\sigma} (\text{or} J?) = i\left(\eta^\mu_{\ \rho}\eta^\nu_{\ \sigma} - \eta^\mu_{\ \sigma}\eta^\nu_{\ \rho}\right)$$

$$J^\mu_{(\rho\sigma)} = x_\rho \tilde{T}^\mu_{\ \sigma} - x_\sigma \tilde{T}^\mu_{\ \rho} + \left(\mathcal{J}_{\nu\lambda}\right)_{\rho\sigma} V^\mu \nu V^\lambda$$

$$\Theta_{(\rho\sigma)} \equiv J_{\rho\sigma} = \int \mathrm{d}^3 x J^0_{(\rho\sigma)} = L_{\rho\sigma} + \underbrace{S_{\rho\sigma}}_{\text{spin}}$$

Example

Exercise: calculate $W^2_{R.F.}$ in the rest frame $k^\mu = (M, 0, 0, 0)$

$$W_\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\mathcal{J}^{\nu\rho}P^\sigma \qquad \begin{cases} W_0^{R.F.} = 0 \\ W_i^{R.F.} = \frac{M}{2}\epsilon_{ijk}\mathcal{J}^{jk} = M\Sigma_i \end{cases}$$

$$W^2_{R.F} = -\left(W_i^{R.F.}\right)^2 = -M\overline{\Sigma}\cdot\overline{\Sigma} = -2M^2 \underbrace{\begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}}_{V_0, \overline{V}}$$

4.4.3.3 Hamiltonian formalism

$$\pi^\mu \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 V_\mu} = -V^{0\mu} \qquad \rightarrow \qquad \begin{cases} \pi^0 \\ \pi^i = -V^{0i} \end{cases}$$

$$\begin{cases} \text{The } V^0 \text{ component is not dynamical} & V_0 = \text{const} = 0 \\ \text{The } V^i \text{ components are dynamical fields} \end{cases}$$

$$\begin{aligned} \mathcal{H} &= \pi^\mu(\partial_0 V_\mu) - \mathcal{L} = \cancel{\pi^0(\partial_0 V_0)} + \pi^i(\partial_0 V_i) + \frac{1}{2}V_{0i}V^{0i} + \frac{1}{4}V^{ij}V_{ij} - \frac{1}{2}M^2\left(V_0^2 + V^i V_i\right) \\ &= \frac{1}{2}\pi_i^2 + \frac{1}{4}\left(V_{ij}\right)^2 + \frac{M^2}{2}V_i^2 + \left(-\frac{M}{2}V_i^2 - \pi_i\partial_i V_0\right) \\ &= \frac{1}{2}\pi_i^2 + \frac{1}{4}\left(V_{ij}\right)^2 + \frac{M^2}{2}V_i^2 + \left(\cancel{\frac{M^2}{2}V_0^2} + 2(\ldots)\right) \end{aligned}$$

$$H = \sum_{i=1}^{3}\int \mathrm{d}^3x \left(\frac{1}{2}\pi_i^2 + \frac{1}{2}\left(V_{ij}\right)^2 + \frac{M^2}{2}V_i^2\right) > 0$$

4.4.3.4 Poisson brackets

$$\begin{cases} \dot{V}_i = \{V_i(x,t), H\} \\ \dot{\pi}_i = \{\pi(x,t), H\} \end{cases}$$

$$\begin{cases} \{V_i, \pi_j\} = \delta_{ij}\delta(\overline{x} - \overline{y}) \\ \{V_i, V_j\} = \{\pi_i, \pi_j\} = 0 \end{cases}$$

But we have a problem. The theory only depends on space components and thus is no longer covariant.

4.4.4 Massless vector field ($M = 0$): E.M.

4.4.4.1 Guess the Lagrangian

$$\mathcal{L}_{EM} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \qquad \rightarrow \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

4.4.4.2 Eq. of motion

$$\begin{aligned} \frac{\partial \mathcal{L}_{EM}}{\partial A_\sigma} - \partial_\rho \frac{\partial \mathcal{L}_{EM}}{\partial \partial_\rho A_\sigma} &= \partial_\rho F^{\rho\sigma} = 0 \\ &= \Box A^\sigma - \partial^\sigma\left(\partial_\mu A^\mu\right) = 0 \end{aligned}$$

Free Maxwell eq!
Taking 4 div ∂_σ does NOT give any constraint! (But we still need to reduce d.o.f.)

4.4.4.3 U(1) Gauge invariance

A gauge transformation of A^μ

$$A^\mu(x) \quad \rightarrow \quad A^{\mu\prime}(x) = A^\mu(x) + \partial^\mu \alpha(x) \qquad (\alpha(x) \in \mathbb{R})$$

It is a local transformation $\alpha = \alpha(x)$

$$F^{\mu\nu}(x) \quad \rightarrow \quad F'^{\mu\nu}(x) = \partial^\mu A^{\nu\prime} - \partial^\nu A^{\mu\prime}$$
$$= \partial^\mu A^\nu - \partial^\nu A^\mu + \cancel{\partial^\mu \partial^\nu \alpha} - \cancel{\partial^\nu \partial^\mu \alpha} = F^{\mu\nu}(x)$$

$$\mathcal{L}_{EM} \quad \rightarrow \quad \mathcal{L}'_{EM} = -\frac{1}{4} F'^{\mu\nu} F'_{\mu\nu} = \mathcal{L}_{EM}$$

U(1)local symmetry $\leftrightarrow$ Physics is independent of the chosen gauge

A^μ with 4 d.o.f.. Add gauge fixing: 2 d.o.f.

4.4.4.4 Example of gauge fixing = Coulomb gauge

The Coulomb gauge is defined by

$$\overline{\nabla} \cdot \overline{A} = 0 \qquad A_0 = 0 \qquad \text{(Free theory)}$$

However not covariant (so not very useful)

Maxwell eq in Coulomb gauge

$$\Box A^\mu - \partial(\cancel{\partial_\rho A^\rho}) = 0$$

Which is the Klein-Gordon equation with $M = 0, \omega_k = |\overline{k}|$

General solution

$$A^\mu(x) \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \sum_{\lambda=0}^{3} \epsilon^\mu_\lambda(k) \left(a_\lambda(k) e^{-ikx} + a^*_\lambda(k) e^{ikx} \right)_{\omega_k = |\overline{k}| = k}$$

Coulomb gauge conditions

$$A^0(x) = 0 \quad \rightarrow \quad \epsilon^0_{(\lambda)} = 0$$
$$\overline{\nabla} \cdot \overline{A} = 0 \quad \rightarrow \quad \overline{k} \cdot \overline{\epsilon}_{(\lambda)}$$

From that we get

$$k^\mu = (|\overline{k}|, 0, 0, k)$$

and

$$\begin{cases} \epsilon^\mu_{(1)} = (0,1,0,0) \\ \epsilon^\mu_{(2)} = (0,0,1,0) \end{cases} \qquad \text{Transverse polerizations}$$

Lecture 8/11

$$\mathcal{L}_{EM} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

$$\partial_\mu F^{\mu\nu} = \Box A^\nu - \partial^\nu \left(\partial_\mu A^\mu\right) = 0$$

Local U(1) internal symmetry - gauge

$$A'^\mu(x) = A^\mu(x) + \partial_\mu \alpha(x)$$

1. Coulomb gauge

$$\overline{\nabla} \cdot \overline{A} = 0 \qquad A_0 = 0$$

Depends on reference frame (not covariant), not good for relativistic calculations

2. Lorentz gauge

The Lorentz gauge is defined imposing

$$\partial_\mu A^\mu(x) = 0 \qquad \rightarrow \qquad \text{covariant}$$

One can always go in the Lorentz gauge

$$A^{\mu\prime}(x) = A^\mu(x) + \partial^\mu \alpha$$

$$0 = \partial_\mu A'^\mu = \partial_\mu A^\mu(x) + \Box\alpha(x) \qquad \rightarrow \qquad \Box\alpha(x) = -\partial_\mu A^\mu(x)$$

The Lorentz gauge does not completely fix the gauge

$$A''^\mu(x) = A'^\mu(x) + \partial^\mu \beta(x)$$

$$0 = \partial_\mu A''^\mu = \cancel{\partial_\mu A'^\mu(x)} + \Box\beta(x) = 0 \qquad \rightarrow \qquad \Box\beta(x) = 0$$

Eq of motion in Lorentz gauge

$$\begin{cases} \Box A^\mu(x) = 0 & \left(k^2 = 0 \leftrightarrow \omega_k = |k|\right) \\ \partial_\mu A^\mu(x) = 0 \end{cases}$$

$$A^\mu(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \sum_{\lambda=1}^{3} \epsilon^\mu_\lambda(k) \left(a_\lambda(k) e^{-ikx} + a^*_\lambda(k) e^{ikx}\right)_{k_0=|k|}$$

$$\partial_\mu A^\mu = 0 \qquad \rightarrow \qquad k_\mu \epsilon^\mu_\lambda = 0$$

Hamiltonian formalism

$$\pi^\mu \equiv \frac{\partial \mathcal{L}}{\partial \partial_0 A_\mu} = -F^{0\mu} \qquad \rightarrow \qquad \begin{cases} \pi^0 & \rightarrow \quad A_0 = k \\ \pi^i = -F^{0i} \end{cases}$$

$$\mathcal{H} = (\ \ )$$

Not covariant quantization!! (Pity)

4.4.4.5 Add a gauge fixing term to $\mathcal{L}_{EM}$

$$\mathcal{L} = \mathcal{L}_{EM} + \mathcal{L}_{GF} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2$$

Where ξ is an arbitrary real parameter. $\mathcal{L}$ is not gauge invariant! So it is not an EM theory (?)

Eq of motion

$$\frac{\partial\mathcal{L}}{\partial A_\mu} - \partial_\nu\frac{\partial\mathcal{L}}{\partial\partial_\nu A_\mu} = \Box A^\nu - \left(\frac{\xi-1}{\xi}\right)\partial^\nu\left(\partial^\mu A_\mu\right) = 0$$

Feynman gauge fixing: $\xi = 1$

$$\Box A^\nu = 0 \qquad \rightarrow \qquad \text{Klein-Gorden eq. massless field}$$

$$A^\mu(x) = \frac{1}{(2\pi)^{3/2}}\int\frac{\mathrm{d}^3k}{\sqrt{2\omega_k}}\sum_{\lambda=1}^{3}\epsilon^\mu_\lambda(k)\left(a_\lambda(k)e^{-ikx} + a^*_\lambda(k)e^{ikx}\right)_{k_0=|k|}$$

ϵ is real and my choice.
A basis for polarization vectors satisfies

$$\begin{cases}\epsilon^\mu_{(\lambda)}(k)\epsilon_{\mu(\lambda')}(k) = \eta_{\lambda\lambda'} & \text{(orthogonality)}\\ \sum_{\lambda=0}^{3}\epsilon^\mu_{(\lambda)}(k)\epsilon^\nu_{(\lambda')}(k)\eta^{\lambda\lambda'} = \eta^{\mu\nu} & \text{(completeness)}\end{cases}$$

$$k^\mu = \left(|\overline{k}|, \overline{k}\right) \qquad (k^2 = 0)$$

$$n^\mu \qquad \rightarrow \qquad \begin{cases}n^\mu n_\mu = 1\\ n^\mu k_\mu \neq 0\end{cases}$$

$$\begin{cases}\epsilon^\mu_{(0)} = n^\mu & \text{Scalar / timelike}\\ \epsilon^\mu_{(3)} = \frac{k^\mu - (n\cdot k)n^\mu}{n\cdot k} & \text{Longitudinal}\\ \epsilon^\mu_{(1,2)} \rightarrow \left(k^\mu\epsilon^{(1,2)}_\mu = 0 = n^\mu\epsilon^{(1,2)}_\mu\right) + \epsilon^\mu_{(1,2)}\epsilon_{\mu(1,2)} = -1 & \text{Transverse}\end{cases}$$

Example

Exercise: Prove that $\epsilon^\mu_{(\lambda)}$ satisfy the conditions of orthogonality and completeness.
Exercise: Let's choose $k^\mu = (|k|, 0, 0, k), \quad n^\mu = (1, 0, 0, 0)$

$$\epsilon^\mu_{(0)} = (1,0,0,0) \qquad \rightarrow \qquad \text{Scalar / timelike}$$

$$\epsilon^\mu_{(3)} = (0,0,0,1) \qquad \rightarrow \qquad \text{Longitudinal}$$

$$\begin{cases}\epsilon^\mu_{(1)} = (0,1,0,0)\\ \epsilon^\mu_{(2)} = (0,0,1,0)\end{cases} \qquad \rightarrow \qquad \text{Transverse}$$

Hamiltonian formalism ($\xi = 1$)

$$\begin{cases}\pi^\mu \equiv \frac{\partial\mathcal{L}}{\partial\partial_0 A_\mu} = -\partial_0 A^\mu & \rightarrow \qquad \pi^\mu \neq 0\\ \mathcal{H} = \pi^\mu\partial_0 A_\mu - \mathcal{L} = \frac{1}{2}\left\{\pi_i^2 + (\partial_i A_j)^2 - \pi_0^2 - (\partial_i A_0)^2\right\}\end{cases}$$

So $\mathcal{H} \geq 0$ no more positive definite

Poisson brackets

$$\{A_\mu(\overline{x},t),\pi(\overline{y},t)\}=\eta_{\mu\nu}\delta^3(\overline{x}-\overline{y})$$
$$\{A_\mu(\overline{x},t),A_\nu(\overline{y},t)\}=\{\pi_\mu(\overline{x},t),\pi_\nu(\overline{y},t)\}=0$$

And the equation of motion

$$\dot{A}_\mu(x^\mu)=\{A_\mu(\overline{x},t),H\}_t$$
$$\pi A_\mu(x^\mu)=\{pi_\mu(\overline{y},t),H\}_t$$

4.4.5 Covariant quantization of E.M.

$$\begin{cases}\mathcal{L}=\mathcal{L}_{EM}+\mathcal{L}_{GF}=-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}-\frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2\\ \xi=1\end{cases}$$

4.4.5.1 Canonical quantization using commutators

Commutators

$$[A_\mu(\overline{x},t),\pi_\nu(\overline{y},t)]=i\eta_{\mu\nu}\delta^3(\overline{x}-\overline{y})$$
$$[A_\mu(\overline{x},t),A_\nu(\overline{y},t)]=0=[\pi_\mu(\overline{x},t),\pi_\nu(\overline{y},t)]$$
$$\begin{cases}[A_0(\overline{x},t),\partial_0 A_0(\overline{y},t)]=\underbrace{-}\,i\delta^3(\overline{x}-\overline{y}) & \text{problem}\\ [A_i(\overline{x},t),\partial_0 A_j(\overline{y},t)]=\underbrace{+}\,i\delta^3(\overline{x}-\overline{y})\delta_{ij}\end{cases}$$

Example

Exercise: Prove

$$a_0(k)=\underbrace{-}\,\frac{1}{(2\pi)^{3/2}}\int\frac{\mathrm{d}^3x}{\sqrt{2\omega_k}}\epsilon^\mu_{(0)}(k)\left(\omega_k A_\mu(x)+i\partial_0 A_\mu(x)\right)e^{ikx}\Big|_{k_0=|\overline{k}|}$$
$$a_i(k)=\underbrace{+}\,\frac{1}{(2\pi)^{3/2}}\int\frac{\mathrm{d}^3x}{\sqrt{2\omega_k}}\epsilon^\mu_{(i)}(k)\left(\omega_k A_\mu(x)+i\partial_0 A_\mu(x)\right)e^{ikx}\Big|_{k_0=|\overline{k}|}$$

Completeness / orthogonality relations

Quantization conditions in momentum space

$$\begin{cases}\left[a_\lambda(k),a^\dagger_{\lambda'}(p)\right]=-\eta_{\lambda\lambda'}\delta^3(\overline{k}-\overline{p})\\ [a_\lambda(k),a_{\lambda'}(p)]=0=\left[a^\dagger_\lambda(k),a^\dagger_{\lambda'}(p)\right]\end{cases}$$

$$\begin{cases}\left[a_0(k),a^\dagger_0(p)\right]=-\delta^3(\overline{k}-\overline{p})\\ \left[a_i(k),a^\dagger_i(p)\right]=+\delta^3(\overline{k}-\overline{p})\end{cases}$$

We define

$$\begin{cases}\mathcal{N}_i(k)=a^\dagger_i(k)a_i(k) & \rightarrow \quad N_i(k)=\int\mathrm{d}^3k a^\dagger_i(k)a_i(k)\\ \mathcal{N}_0(k)=-a^\dagger_0(k)a_0(k) & \rightarrow \quad N_0(k)=-\int\mathrm{d}^3k a^\dagger_0(k)a_0(k)\end{cases}$$

$$\begin{cases} \left[\mathcal{N}_0(k), a_0^\dagger(p)\right] = +a_0^\dagger(\overline{p})\delta^3(\overline{k}-\overline{p}) \\ [\mathcal{N}_0(k), a_0(p)] = -a_0(\overline{p})\delta^3(\overline{k}-\overline{p}) \end{cases}$$

$$H = \int \mathrm{d}^3k\,|\overline{k}| \sum_{\lambda=0}^{3} \mathcal{N}_\lambda(k) \qquad P_i = \int \mathrm{d}^3k k_i \sum_{\lambda=0}^{3} \mathcal{N}_\lambda(k)$$

Hamiltonian operator

$$N[H] = \int \mathrm{d}^3k\omega_k \sum_{\lambda=0}^{3} \mathcal{N}_\lambda(k)$$

$$N[\overline{P}] = \int \mathrm{d}^3k\overline{k} \sum_{\lambda=0}^{3} \mathcal{N}_\lambda(k)$$

4.4.5.2 Fock space

1. Def the vacuum state

$$\begin{array}{lll} |0\rangle & \to & a_{(\lambda)}(k) = 0 \\ & \to & \mathcal{N}_{(\lambda)}(k)\,|0\rangle = 0 \end{array}$$

2. Space components $|n_i(k)\rangle \propto (a_i^\dagger)^{n_i}\,|0\rangle$
3. Time components ($\lambda = 0$)

$$|1_0(k)\rangle \propto a_0^\dagger(k)\,|0\rangle \qquad \to \qquad N_0\,|1_0(k)\rangle = +1\,|1_0(k)\rangle$$

$$\begin{aligned} \langle 1_0(k)|1_0(k)\rangle &\propto \langle 0|a_0(k)a_0^\dagger(k)|0\rangle \\ &= \langle]|[a_0(k), a_0^\dagger(k)|1_0(p)\rangle 1_0(p) + \cancel{\langle a_0^\dagger(k)a_0(k)|0|a_0^\dagger(k)a_0(k)\rangle} \\ &= -1\delta(0) \quad < 0 \qquad \text{negative norm is unphysical} \end{aligned}$$

Physical states are a subset of the Fock space of my theory.

Example

$|1_0(p)\rangle$ has negative energy!

$$\langle 1_0(p)|H|1_0(p)\rangle = -\omega_p$$

4.4.5.3 Gupta-Bleuler condition

The physical states are the ones that satisfy

$$\boxed{\langle \varphi_{\mathrm{PHYS}}|\partial_\mu A^\mu(x)|\varphi'_{\mathrm{PHYS}}\rangle = 0} \qquad \left\{ \begin{matrix} |\varphi_{\mathrm{PHYS}}\rangle \\ |\varphi'_{\mathrm{PHYS}}\rangle \end{matrix} \right\}$$

This is equivalent with

$$\begin{cases} \partial_\mu A_+^\mu\,|\varphi_{\mathrm{PHYS}}\rangle = 0 \\ \langle \varphi_{\mathrm{PHYS}}|\,\partial_\mu A_-^\mu = 0 \end{cases}$$

$$\partial_\mu A^\mu_+(x) = \frac{-i}{(2]pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega_k}} \underbrace{\sum_{\lambda=0}^{4} k^\mu \epsilon^{(\lambda)}_\mu(k) a_\lambda(k)}_{\equiv L(k)} e^{-ikx}$$

So the Gupta-Bleuler condition in momentum space is

$$\boxed{L(k)\,|\varphi_{\mathrm{PHYS}}\rangle = 0}$$

In the specific basis for $\epsilon^\mu_{(\lambda)}(k)$ $\quad \epsilon^\mu_0 = (1,0,0,0), \epsilon^\mu_3 = (0,0,0,1)$

$$L(k) = i\,(a_3(k) - a_0(k))$$

So

$$L\,|\varphi_{\mathrm{PHYS}}\rangle = 0 \leftrightarrow a_3\,|\varphi_{\mathrm{PHYS}}\rangle = a_0\,|\varphi_{\mathrm{PHYS}}\rangle$$
$$\langle\varphi_{\mathrm{PHYS}}|\,a_3^\dagger = \langle\varphi_{\mathrm{PHYS}}|\,a_0^\dagger$$

1. Count $n_3 + n_0$

$$\langle \mathcal{N}_3(k) + \mathcal{N}_0(k)\rangle\,\varphi_{\mathrm{PHYS}} = \left\langle a_3^\dagger(k) a_3(k) - a_0^\dagger(k) a_0(k) \right\rangle \varphi_{\mathrm{PHYS}}$$
$$= \left\langle a_3^\dagger(k) \left(\underbrace{a_3(k) - a_0(k)}_{L} \right) \right\rangle \varphi_{\mathrm{PHYS}} = 0$$

2. φ_{PHYS} has positive energy

$$\langle H\rangle\,\varphi_{\mathrm{PHYS}} = \int \mathrm{d}^3 k \omega_k \sum_{\lambda=0}^{3} \langle \mathcal{N}_\lambda(k)\rangle\,\varphi_{\mathrm{PHYS}}$$
$$= \int \mathrm{d}^3 k \omega_k\,\langle \mathcal{N}_1(k) + \mathcal{N}_2(k)\rangle\,\varphi_{\mathrm{PHYS}}$$
$$= \int \mathrm{d}^3 k \omega_k\,(n_1(k) + n_2(k)) > 0$$

Lecture 12/11

4.4.6 Covariant commutators for real vector

$$[A^\mu(x), A^\nu(y)] = [A^\mu_+(x), A^\nu_-(y)] + [A^\mu_-(x), A^\nu_+(y)] = D^{\mu\nu}_+(x-y) + D^{\mu\nu}_-(x-y)$$

$$\begin{cases} \left[A^\mu_+(x), A^\nu_-(y)\right] = -\eta^{\mu\nu} D_+(x-y) \\ \left[A^\mu_-(x), A^\nu_+(y)\right] = -\eta^{\mu\nu} D_-(x-y) \end{cases}$$

Microcausality is satisfied

Real field no charge, complex has charge.

4.5 Other fields?

Overview so far:

1. Scalar field $(\mathbb{R}, \mathbb{C})$ $\rightarrow$ spin 0
2. Dirac field $\rightarrow$ spin 1/2
3. Vector field $\rightarrow$ spin 1

This is enough to describe all known fundamental particles.

Chapter 5

Interacting quantum fields

5.1 Interaction terms in the Lagrangian

Interacting fields can be studied by adding an interaction term to the Lagrangian.

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I$$

where $\mathcal{L}_0$ is the free field Lagrangian and $\mathcal{L}_I$ is the interaction term. The Hamiltonian can be split in the same way.

5.1.1 Scalar field self-interaction

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\text{int}} = \frac{1}{2}\left(\partial_\mu \varphi\right)\left(\partial^\mu \varphi\right) - \underbrace{\frac{1}{2}m^2\varphi^2 - V_{\text{int}}(\varphi)}_{V(\varphi)}$$

$$V(\varphi) = \underbrace{\frac{1}{2}m^2\varphi}_{\text{mass (?)}} + \underbrace{\frac{k}{3}\varphi^3 + \frac{\lambda}{4}\varphi^4 + \frac{\mu}{5}\varphi^5 + \dots}_{\text{coupling}}$$

Linear term minimum $\varphi \neq 0$. Check minimum (not maximum).

5.1.1.1 Dimension of coupling

$$\begin{aligned} [S] = M^0 \quad &\rightarrow \quad [\mathcal{L}] = M^4 \\ [\partial_\mu] = M^1, \quad &[\varphi] = M^2 \\ [m] = M^1, \quad &[k] = M^1, \quad [\lambda] = M^0, \quad [\mu] = M^{-1}, \quad \dots \end{aligned}$$

A theory is said to be renormalizable if it only has couplings M^α with $\alpha \geq 0$.

A renormalizable theory is valid for any scale

Complex case:

$$\mathcal{L} = \left(\partial_\mu \varphi^*\right)\left(\partial^\mu \varphi\right) - m^2\varphi^*\varphi - \lambda\left(\varphi^\dagger\varphi\right)^2 - \cancel{\left(\mu\left(\varphi^\dagger\varphi\right)^3 + \dots\right)}$$

5.1.2 Dirac field self-interactions

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\text{int}} = \overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi + \cancel{G\left(\overline{\psi}\Gamma\psi\right)\left(\overline{\psi}\Gamma\psi\right)}$$

Where G is called the Fermi constant.

$$[\psi] = M^{3/2}, \qquad [G] = M^{-2}$$

5.1.3 Interaction between scalar and Dirac field

$$\mathcal{L} = \mathcal{L}_0\mathcal{L}_{\text{int}} = \mathcal{L}_0 + y_s\varphi\overline{\psi}\psi + y_p\varphi\overline{\psi}\gamma_5\psi + \cancel{\left(\omega_1\varphi^2\overline{\psi}\psi + \ldots\right)}$$

$$[y_{s,p}] = M^0 \qquad \text{Yukawa coupling}$$
$$[\omega_1] = M^{-1}$$

Where s and p stand for scalar and pseudoscalar interactions.

$$\begin{cases}\varphi(-x) = \varphi(x) & \text{scalar}\\ \varphi(-x) = -\varphi(x) & \text{speudoscalar}\end{cases}$$

Parity not generally a symmetry of interactions. (?)

5.1.4 Interaction between Dirac and vector fields

$$\begin{aligned}\mathcal{L} = \mathcal{L}_0\mathcal{L}_{\text{int}} = \mathcal{L}_0 + g_v\overline{\psi}\gamma^\mu\psi V_\mu + g_a\overline{\psi}\gamma^\mu\gamma_5\psi V_\mu\\ + c_v\overline{\psi}\psi V^\mu V_\mu + ic_a\overline{\psi}\gamma_5\psi V^\mu V_\mu\\ + d_v\overline{\psi}\sigma^{\mu\nu}\psi V_{\mu\nu} + id_a\overline{\psi}\sigma^{\mu\nu}\psi V_{\mu\nu}\end{aligned}$$

$$[g_{v,a}] = M^0, \qquad [c_{v,a}] = M^{-1} = [d_{v,a}]$$

$$V^{\mu\nu}V_{\mu\nu} = \left(\partial^\mu V^\nu - \partial^\nu V^\mu\right)\left(\partial_\mu V_\nu - \partial_\nu V_\mu\right) \sim \partial^\mu V^\nu \partial_\mu V_\nu$$
$$[\partial_\mu] = M^1 \qquad [V^\mu] = M^1$$

Where v and a stand for vector and axial vector. (Missing γ^5?)

$$\begin{cases}V^\mu(-x) = V^\mu(x) & \text{vector}\\ V^\mu(-x) = -V^\mu(x) & \text{axial vector}\end{cases}$$

$$\begin{aligned}\mathcal{L}^2_{\text{int}} &= \left(g_v\overline{\psi}\gamma^\mu\psi + g_a\overline{\psi}\gamma^\mu\gamma_5\psi\right)V_\mu\\ &= \left(g_L\overline{\psi}\gamma^\mu_L\psi + g_R\overline{\psi}\gamma^\mu_R\psi\right)V_\mu\end{aligned}$$

Where

$$\begin{cases}g_L = g_v - g_a\\ g_R = g_v + g_a\end{cases} \qquad \begin{cases}\gamma^\mu_L = \gamma^\mu P_L\\ \gamma^\mu_R \equiv \gamma^\mu P_R\end{cases} \qquad \begin{cases}P_L = \frac{1-\gamma_5}{2}\\ P_R = \frac{1+\gamma_5}{2}\end{cases}$$

$$\mathcal{L}^2_{\text{int}} = \left(g_L\overline{\psi}_L\gamma^\mu\psi_L + g_R\overline{\psi}_R\gamma^\mu\psi_R\right)V_\mu$$

$$\begin{cases}\psi_L \equiv P_L\psi\\ \overline{\psi}_L \equiv \overline{(\psi_L)} = \psi^\dagger_L\gamma_0 = \psi^\dagger P_L\gamma^0 = \psi^\dagger\gamma^0 P_R = \overline{\psi}P_R\end{cases} \qquad \begin{cases}\psi_R = P_R\psi\\ \overline{\psi}_R = \overline{\psi}P_L\end{cases}$$

$$\overline{\psi}R_R\gamma^\mu P_L\psi = \overline{\psi}\gamma^\mu P^2_L\psi = \overline{\psi}\gamma^\mu_L\psi$$

- Theory is vector like if $g_A = 0$ and $g_L = g_R$
- Interaction is chiral if $g_L \neq g_R \quad (g_v, g_a \neq 0)$

5.1.5 Classical electrodynamics

Interaction between fermions and photons.
Looking at the interaction between a photon and an electron we realize:

- Photon is a vector field (not axial field)
- ED is vector like theory $(g_v \neq 0, g_a = 0, g_L = g_R)$

The most general renormalizable Lagrangian is

$$\begin{aligned}\mathcal{L}_{\text{ED}} &= \mathcal{L}_0 + g_v \overline{\psi}\gamma^\mu \psi A_\mu \\ &= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi - q\overline{\psi}\gamma^\mu\psi A_\mu \\ &= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\psi}\left(iD\!\!\!/ - m\right)\psi\end{aligned}$$

Where $D_\mu \equiv \partial_\mu + iqA_\mu$. This Lagrangian is the same as we guessed before from Schrödinger eq.
QED has a U(1) local (gauge) symmetry

$$\begin{cases} A'^\mu(x) = A^\mu(x) + \partial^\mu\alpha(x) \\ \psi'(x) = e^{-iq\alpha(x)}\psi(x) \end{cases}$$

The EM free Lagrangian is invariant $(F'^{\mu\nu}(x) = F^{\mu\nu}(x))$, so to check the symmetry we check the fermion interaction $(\overline{\psi}\left(iD\!\!\!/ - m\right)\psi)$.

$$\begin{aligned} m\overline{\psi}'\psi' &= m\cancel{e^{iq\alpha}}\cancel{e^{-iq\alpha}}\overline{\psi}\psi \\ (D_\mu\psi)' &= \left(\partial_\mu + iqA'_\mu\right)\psi'(x) \\ &= \left(\partial_\mu + iqA_\mu + iq(\partial_\mu\alpha)\right)e^{-iq\alpha(x)}\psi(x) \\ &= e^{-iq\alpha(x)}\left(\partial_\mu + iqA_\mu + \cancel{iq(\partial_\mu\alpha)} - \cancel{iq(\partial_\mu\alpha)}\right)\psi \\ &= e^{-iq\alpha(x)}\left(D_\mu\psi\right)\end{aligned}$$

So

$$\begin{aligned}\overline{\psi}iD\!\!\!/\psi \qquad \rightarrow \qquad \overline{\psi}'i\left(D\!\!\!/\psi\right)' &= \overline{\psi}e^{iq\alpha(x)}e^{-iq\alpha(x)}\left(D\!\!\!/\psi\right) \\ &= \overline{\psi}i\left(D\!\!\!/\psi\right)\end{aligned}$$

From which follows that the QED Lagrangian is invariant under the U(1) gauge symmetry.

Example

Exercise: Complex scalar field + photon

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_{\text{int}} \\ &= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \left(D_\mu\varphi^*\right)\left(D^\nu\varphi\right) - m^2\varphi^*\varphi + V_{\text{int}}(\varphi^*\varphi)\end{aligned}$$

$$D_\mu \qquad \rightarrow \qquad D_\mu = \partial_\mu + iqA_\mu$$

U(1) local invariance

$$\begin{cases} A^{\mu\prime}(x) = A^\mu(x) + \partial^\mu\alpha(x) \\ \varphi'(x) = e^{-iq\alpha(x)}\varphi(x) \end{cases}$$

5.2 Particles in interacting theories

Real vs bare particles
Adiabatically switching on interaction. p.93 Mandl Shaw

5.3 Quantization of interacting theory

5.3.1 An attempt

Example

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\varphi\right)\left(\partial^\mu\varphi\right) - \frac{1}{2}m^2\varphi^2 - \frac{\lambda}{4}\varphi^4$$

$$\left(\Box + m^2\right)\varphi(x) = -\lambda^3 \qquad \rightarrow \qquad \text{eq. of motion}$$

Hamiltonian formalism

$$\begin{cases} \pi = \frac{\partial\mathcal{L}}{\partial\partial_0\varphi} = \partial_0\varphi \\ \mathcal{H} = \pi(\partial_0\varphi) - \mathcal{L} = \mathcal{H}_\prime + \mathcal{H}_{\text{int}} \end{cases}$$

We can see that $\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}}$
Poisson brackets:
We can impose

$$\begin{cases} \{\varphi, \pi\} = \delta^3(x-y) \\ \{\varphi,\varphi\} = \{\pi,\pi\} = 0 \end{cases}$$

Canonical quantization

$$\begin{aligned}&[\varphi(\overline{x},t), \pi(\overline{y},t)] = i\delta^3(\overline{x}-\overline{y}) \\ &[\varphi(\overline{x},t), \varphi(\overline{y},t)] = 0 = [\pi(\overline{x},t), \pi(\overline{y},t)]\end{aligned}$$

Now we want to invert $\phi \qquad \rightarrow \qquad a, a^\dagger$. Problem: what is the solution

$$\boxed{\left(\Box + m^2\right)\varphi = -\lambda\varphi^3}$$

$$\varphi \sim \cancel{ae^{-ikx} + a^\dagger e^{ikx}} \qquad \text{No solution of eq of motion}$$

We are not able to obtain a basis of eigenvectors of H

$$\cancel{[a, a^\dagger] = \delta^3(k-p)} \qquad \cancel{[a, a] = 0 = [a^\dagger, a^\dagger]}$$

5.3.2 Problems quantizing an interacting field theory

- $\left(\Box + m^2\right)\varphi = -\lambda\varphi^3$. Not able to solve the eq. of motion exactly.
- No solution of the form $\varphi \sim ae^{-ikx} + a^\dagger e^{ikx}$
- $a^\dagger |0\rangle$ is not an eigenstate of H.

Because we do not know exactly, we use perturbation theory.

5.4 The *S*-matrix expansion

In this section we are interested in scattering processes. We start off with particles far apart from each other a long time before the scattering occurs. This is the initial state $|i\rangle$. In the scattering process the particles come close together, interact and fly apart again. After the scattering process the particles are again far apart and non-interacting.
The study of such scattering processes is greatly simplified by using the interaction picture. This is due to two simplifications:

1. The operators in the interaction picture satisfy the Heisenberg-like equations of motion, but involving the free Hamiltonian H_0 only, not the complete Hamiltonian H.
2. If the interaction Lagrangian density $\mathcal{L}_I$ does not contain derivatives, the fields canonically conjugate to the interacting fields and to the free fields are identical. The interacting fields then satisfy the same commutation relations as the free fields. (TODO ?)

A recap of the salient features of the interaction picture:
The system is described by a time-dependent state vector $|\Phi(t)\rangle$. This state vector satisfies the equation of motion

$$i\frac{\mathrm{d}}{\mathrm{d}t}|\Phi(t)\rangle = H_I(t)\,|\Phi(t)\rangle\,,$$

where

$$H_I(t) = e^{iH_0(t-t_0)} H_I^S e^{-iH_0(t-t_0)}$$

is the interaction Hamiltonian in the interaction picture and H_I^S is the interaction Hamiltonian in the Schrödinger picture. The free Hamiltonian H_0 commutes with itself and thus is the same in both pictures.
The time evolution of the state vector can also be described using the unitary operator $U(t_2, t_1)$ such that $|\Phi(t_2)\rangle = U(t_2, t_1)\,|\Phi(t_1)\rangle$. This operator is given by

$$U(t_2, t_1) = e^{iH_0 t_2} e^{-iH(t_2-t_1)} e^{-iH_0 t_1}.$$

So we describe the system state vector $|\Phi(t)\rangle$. At time $t = -\infty$ the system is in the initial state $|\Phi(-\infty)\rangle = |i\rangle$. Then the interaction occurs and at time $t = \infty$ we find the system in the final state $|\Phi(+\infty)\rangle$.
The initial and final states can be related by

$$|\Phi(+\infty)\rangle = U(+\infty, -\infty)\,|\Phi(-\infty)\rangle = U(+\infty, -\infty)\,|i\rangle$$

We call $U(+\infty, -\infty)$ the S-matrix; so $|\Phi(+\infty)\rangle = S\,|i\rangle$. The S-matrix is still unitary.
The final state $|\Phi(+\infty)\rangle$ is in general a superposition of different possible particle configurations. For any particular possible particle configuration $|f\rangle$ (TODO define!!! eigenstates of H_0), the probability of finding the system in the state $|f\rangle$ after the interaction is given by

$$\begin{aligned} |\langle f|\Phi(+\infty)\rangle|^2 &= |\langle f|S|i\rangle|^2 \\ &\equiv |S_{fi}|^2 . \end{aligned}$$

This is the transition probability. If we consider a complete orthonormal set of particles states (again TODO define), $|\Phi(+\infty)\rangle$ can be expanded as

$$|\Phi(+\infty)\rangle = \sum_f |f\rangle\,\langle f|\Phi(+\infty)\rangle = \sum_f |f\rangle\, S_{fi}.$$

The unitarity of the S-matrix implies

$$\sum_f |S_{fi}|^2 = 1.$$

which asserts that we will find the system in a final state $|f\rangle$, but does allow for the creation and destruction of particles.
We are clearly very interested in expressions for the S-matrix and its matrix elements S_{fi}. Such expressions can be obtained by solving the equation

$$i\frac{\mathrm{d}}{\mathrm{d}t}\,|\Phi(t)\rangle = H_I(t)\,|\Phi(t)\rangle$$

with the initial condition $|\Phi(-\infty)\rangle = |i\rangle$. This differential equation can be done iteratively:

$$\begin{aligned} |\Phi(t)\rangle &= |i\rangle + (-i)\int_{-\infty}^{t} \mathrm{d}t_1 H_I(t_1)\,|\Phi(t_1)\rangle \\ &= |i\rangle + (-i)\int_{-\infty}^{t} \mathrm{d}t_1 H_I(t_1)\,|i\rangle + (-i)^2\int_{-\infty}^{t}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2 H_I(t_1)H_I(t_2)\,|\Phi(t_2)\rangle \\ &= \left[\sum_{n=0}^{\infty}(-i)^n\int_{-\infty}^{t}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\ldots\int_{-\infty}^{t_{n-1}}\mathrm{d}t_n H_I(t_1)H_I(t_2)\ldots H_I(t_n)\right]|i\rangle \end{aligned}$$

Taking the limit $t \to +\infty$, we get an expression for the S-matrix:

$$\sum_{n=0}^{\infty}(-i)^n\int_{-\infty}^{+\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\ldots\int_{-\infty}^{t_{n-1}}\mathrm{d}t_n H_I(t_1)H_I(t_2)\ldots H_I(t_n)$$

This expression can be written using the time-ordered product. Then for any n

$$H_I(t_1)H_I(t_2)\ldots H_I(t_n) = T\left[H_I(t_1)H_I(t_2)\ldots H_I(t_n)\right]$$

since the integration bounds enforce $t_1 \geq t_2 \geq \ldots \geq t_n$.
Using the properties of the time-ordered product, the integrals can be written as going from $-\infty$ to $+\infty$ if each term is divided by $n!$. To do that we need to *assume* that $H_I(t)$ contains an even number of fermionic factors (as in QED), so that the reordering process introduces no extra factors of -1.
Take for example the term $n = 2$:

$$\begin{aligned}\int_{-\infty}^{+\infty} \mathrm{d}t_1 \int_{-\infty}^{+\infty} \mathrm{d}t_2 T\left[H_I(t_1)H_I(t_2)\right] &= \int_{-\infty}^{+\infty} \mathrm{d}t_1 \left(\int_{-\infty}^{t_1} \mathrm{d}t_2 T\left[H_I(t_1)H_I(t_2)\right] + \int_{t_1}^{+\infty} \mathrm{d}t_2 T\left[H_I(t_1)H_I(t_2)\right]\right) \\ &= \int_{-\infty}^{+\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2 T\left[H_I(t_1)H_I(t_2)\right] + \int_{-\infty}^{+\infty} \mathrm{d}t_2 \int_{-\infty}^{t_2} \mathrm{d}t_1 T\left[H_I(t_1)H_I(t_2)\right]\end{aligned}$$

where in the last step we have used (TODO correct theorem). The assumption that $H_I(t)$ contains an even number of fermionic factors means that

$$T\left[H_I(t_1)H_I(t_2)\right] = T\left[H_I(t_2)H_I(t_1)\right]$$

Thus we can rename the integration variables $t_1 \to t_2$ and $t_2 \to t_1$ in the second term to get

$$\int_{-\infty}^{+\infty} \mathrm{d}t_1 \int_{-\infty}^{+\infty} \mathrm{d}t_2 T\left[H_I(t_1)H_I(t_2)\right] = 2\int_{-\infty}^{+\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2$$

So the S-matrix can be written

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2 \ldots \int_{-\infty}^{\infty} \mathrm{d}t_n H_I(t_1)H_I(t_2)\ldots H_I(t_n).$$

Writing this in terms of the Hamiltonian density $\mathcal{H}_I$ gives an explicitly covariant result:

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} \mathrm{d}^4x_1 \int_{-\infty}^{\infty} \mathrm{d}^4x_2 \ldots \int_{-\infty}^{\infty} \mathrm{d}^4x_n \mathcal{H}_I(x_1)\mathcal{H}_I(x_2)\ldots\mathcal{H}_I(x_n) \equiv \sum_{n=0}^{\infty} S^{(n)}.$$

Finally we can recognise this series as the matrix exponential

$$S = T\left[\exp -i\int_{-\infty}^{+\infty} \mathrm{d}^4x \mathcal{H}_I(x)\right].$$

5.5 Wick's theorem

The S-matrix contains information about a very large number of processes, however most will not contribute to a particular transition $|i\rangle \to |f\rangle$. So finding a useful way to expand the expression for S_{fi} becomes very important. The following procedure, due to Dyson and Wick, does just that.
The idea is to rewrite the time-ordered product as a sum of normal ordered products. Since in a normal ordered product all absorption operators are to the right of the creation operators, such product do not cause emission and re-absorption.

5.5.1 Contractions

In order to formulate Wick's theorem, we need to introduce the convenient notation of the *contractions*.

For a product of two field operators, $A(x_1)$ and $B(x_2)$, the contraction is defined as

$$\overbrace{A(x_1)B(x_2)} \equiv \langle]|T\left[A(x_1)B(x_2)\right]|\,0\rangle 0$$

The only non-vanishing contractions are the Feynman propagators

$$\begin{aligned}
\overbrace{\varphi(x_1)\varphi}(x_2) &= i\Delta_F(x_1-x_2) && \text{Real scalar field}\\
\overbrace{\varphi(x_1)\varphi^\dagger}(x_2) = \overbrace{\varphi^\dagger(x_2)\varphi}(x_1) &= i\Delta_F(x_1-x_2) && \text{Complex scalar field}\\
\overbrace{\psi_\alpha(x_1)\overline{\psi}_\beta}(x_2) = -\overbrace{\overline{\psi}_\beta(x_2)\psi_\alpha}(x_1) &= iS_{F\alpha\beta}(x_1-x_2) && \text{Fermionic field}\\
\overbrace{A^\mu(x_1)A^\nu}(x_2) &= iD_F^{\mu\nu}(x_1-x_2) && \text{Vector field}
\end{aligned}$$

We can also define contractions in an arbitrary normal ordered product of field operators:

$$N\left(\overbrace{A\overbrace{BC}D\overbrace{EF\ldots JK}L\overbrace{M\ldots}}\right) = (-1)^p\overbrace{AK}\overbrace{BC}\overbrace{EL}\ldots N\left(DF\ldots JM\ldots\right)$$

where p is is the number of transpositions of fermionic operators. It makes sense to take the contractions out of the normal product because they are not operators, but scalars, vectors or spinors.

5.5.2 Statement

If $A, B, \ldots, Y, Z$ are field operators, Wick's theorem can be stated as follows:

$$\begin{aligned}
T\left(ABCD\ldots WXYZ\right) = {} & N\left(ABCD\ldots WXYZ\right)\\
& + N\left(\overbrace{AB}C\ldots YZ\right) + N\left(\overbrace{ABC}\ldots YZ\right) + \ldots + N\left(A\overbrace{BC}\ldots YZ\right) + \ldots\\
& + N\left(\overbrace{AB}\overbrace{CD}\ldots YZ\right) + N\left(\overbrace{ABCDE}\ldots YZ\right) + \ldots + N\left(\overbrace{ABCDE\ldots YZ}\right) + \ldots\\
& + \ldots
\end{aligned}$$

TODO: No equal times contractions!!

5.5.3 Proof

TODO (by induction)

5.6 The S-matrix expansion in QED

The method for calculating contributions to the matrix elements S_{fi} will be applied here to quantum electrodynamics. This is the theory Feynman originally developed it for. Other theories are analogous.
Feynman developed his diagrams and rules very intuitively. It was Dyson and Wick that put it on the formal footing it is presented with here.

5.6.1 Interactions in quantum electrodynamics

In quantum electrodynamics the interaction Hamiltonian is given by

$$\begin{aligned}\mathcal{H}_I(x) &= -\mathcal{L}_I(x) \\ &= -e\mathbb{N}\left[\overline{\psi}(x)\not{A}(x)\psi(x)\right] \\ &= -e\mathbb{N}\left[(\overline{\psi}^+ + \overline{\psi}^-)(\not{A}^+ + \not{A}^-)(\psi^+ + \psi^-)\right]_x\end{aligned}$$

The minus superscripts refer to the creation operators and the plus subscripts to the annihilation operators.

5.6.2 Feynman diagrams in configuration space

5.6.2.1 First order processes

We start by looking at the first order term $S^{(1)}$.

$$\begin{aligned}S^{(1)} &= -i\int \mathrm{d}^4x_1\mathcal{H}_I(x_1) \\ &= ie\int \mathrm{d}^4x_1\mathbb{N}(\overline{\psi}\not{A}\psi)_{x_1} \\ &= ie\int \mathrm{d}^4x_1\mathbb{N}\left[(\overline{\psi}^+ + \overline{\psi}^-)(\not{A}^+ + \not{A}^-)(\psi^+ + \psi^-)\right]_{x_1}\end{aligned}$$

This can be expanded into eight terms, each corresponding to a basic process in QED. For example, $ie\int \mathrm{d}^4x_1\mathbb{N}\left[\overline{\psi}^+\not{A}^-\psi^+\right]_{x_1}$ corresponds to the annihilation of an electron positron pair with the creation of a photon. The Feynman diagrams of all eight processes are given in figure TODO.
None of these processes are physically possible due to conservation of energy and momentum and the fact that $k^2 = 0$ for photons and $p^2 = m^2$ for fermions.
For any unphysical processes (and thus in particular for these processes) the matrix element vanishes

$$\langle f|S|i\rangle = 0.$$

In fact (TODO: why?) the stronger result

$$\langle f|S^{(n)}|i\rangle = 0$$

holds for any unphysical process at any order n. In particular we have seen that for all processes

$$\langle f|S^{(1)}|i\rangle = 0.$$

5.6.2.2 Second order processes

Applying Wick's theorem, we see that $S^{(2)}$ can be expanded into the following terms:

$$S^{(2)} = S^{(2)}_A + S^{(2)}_B + S^{(2)}_C + S^{(2)}_D + S^{(2)}_E + S^{(2)}_F$$

where

$$S_A^{(2)} = -\frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}\not{A}\psi)_{x_1}(\overline{\psi}\not{A}\psi)_{x_2}\right]$$

$$S_B^{(2)} = -\frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}\not{A}\overbrace{\psi)_{x_1}(\overline{\psi}}\not{A}\psi)_{x_2}\right] - \frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overbrace{\overline{\psi}\not{A}\psi)_{x_1}(\overline{\psi}\not{A}}\psi)_{x_2}\right]$$

$$S_C^{(2)} = -\frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}\gamma^\mu\overbrace{A_\mu\psi)_{x_1}(\overline{\psi}\gamma^\nu}A_\nu\psi)_{x_2}\right]$$

$$S_D^{(2)} = -\frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}\gamma^\mu\overbrace{A_\mu\psi)_{x_1}(\overline{\psi}\gamma^\nu}A_\nu\psi)_{x_2}\right] - \frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overbrace{\overline{\psi}\gamma^\mu A_\mu\psi)_{x_1}(\overline{\psi}\gamma^\nu A_\nu}\psi)_{x_2}\right]$$

$$S_E^{(2)} = -\frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}\not{A}\psi)_{x_1}(\overline{\psi}\not{A}\psi)_{x_2}\right]$$

$$S_F^{(2)} = -\frac{e^2}{2!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}\gamma^\mu A_\mu\psi)_{x_1}(\overline{\psi}\gamma^\nu A_\nu\psi)_{x_2}\right]$$

We consider these terms one by one.

- The term $S_A^{(2)}$ corresponds to two separate processes like the ones shown in figure TODO. These are unphysical and thus $S_A^{(2)}$ does not contribute to any real transitions.

- The two terms of $S_B^{(2)}$ are identical to each other, since vector operators at different locations commute, fermionic operators at different locations anti-commute, there are two of them and the spinor indices are self-contained. That last point means the gamma matrices in each group only operate on spinors in the group, so if the whole group is moved together we do not need to worry about the commutation of the gamma matrices. To be completey sure, the spinorial components should written out explicitly. So we can write

  $$-e^2\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}\not{A}\overbrace{\psi)_{x_1}(\overline{\psi}}\not{A}\psi)_{x_2}\right]$$

  Here there are two uncontracted photon operators and two uncontracted fermion operators that absorb or create external particles. The contraction corresponds to the fermion propagator.

  By selecting the creation or absorption parts of the operators, we get terms corresponding to different processes. As an example we will take a look at **Compton scattering**

  $$\gamma + e^- \to \gamma + e^-.$$

  The relevant part of $S_B^{(2)}$ for this process needs to contain an uncontracted ψ^+ to absorb the initial electron and $\overline{\psi}^-$ to create the final electron. All terms of $S_B^{(2)}$ that do not have that produce a final state that is orthogonal to $|f\rangle = |\gamma, e^-\rangle$. Consequently these terms do not contribute to S_{fi}.

  In our choice of coordinates the electron is absorbed at x_2, so $\psi^+(x_2)$, and recreated at x_1, so $\overline{\psi}^-(x_1)$ (because $\psi^+(x_2)$ and $\overline{\psi}^-(x_1)$ are uncontracted). This still leaves us a choice for the photon: either $\not{A}^+(x_1)$ or $\not{A}^+(x_2)$ can absorb the initial photon, meaning that correspondingly either $\not{A}^-(x_2)$ or $\not{A}^-(x_1)$ must create the final one. This choice means

there are two contributions:

$$S^{(2)}(\gamma e^- \to \gamma e^-) = -e \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 \overline{\psi}^-(x_1)\gamma^\mu i S_F(x_1 - x_2)\gamma^\nu A^-_\mu(x_1)A^+_\nu(x_2)\psi^+(x_2)$$
$$- e \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 \overline{\psi}^-(x_1)\gamma^\mu i S_F(x_1 - x_2)\gamma^\nu A^-_\mu(x_2)A^+_\nu(x_1)\psi^+(x_2)$$

The processes described by $S^{(2)}_B$ are Compton scattering by positrons ($\gamma + e^+ \to \gamma + e^+$) and two-photon **pair annihilation** ($e^+ + e^- \to \gamma + \gamma$) and **creation** ($\gamma + \gamma \to e^+ + e^-$) processes.

- The term $S^{(2)}_C$ contains four uncontracted fermion operators and thus describes fermion-fermion scattering: electron-electron, electron-positron and positron-positron.
  - As a first example we will consider **Møller scattering**

    $$e^- + e^- \to e^- + e^-$$

    The relevant part of $S^{(2)}_C$ is given by

    $$S^{(2)}(2e^- \to 2e^-) = \frac{-e^2}{2!} \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 \mathbb{N}\left[(\overline{\psi}^-\gamma^\mu\psi^+)_{x_1}(\overline{\psi}^-\gamma^\nu\psi^+)_{x_2}\right] i D_{F\,\mu\nu}(x_1 - x_2)$$

    We are interested in the transition

    $$|i\rangle = c^\dagger(2)c^\dagger(1)\,|0\rangle \to |f\rangle = c^\dagger(2')c^\dagger(1')\,|0\rangle$$

    Where the labels $1, 2, 1', 2'$ mean the electrons can have different spins and momenta. Applying this to $S^{(2)}_C$ yields 4 terms for $S^{(2)}_{fi}$, since either initial electron can be absorbed by either ψ^+ operator and either final electron can be emitted by either $\overline{\psi}_-$ operator.
    These are in fact two pairs of two identical contributions related by exchanging $x_1 \leftrightarrow x_2$. This allows us to drop the prefactor $1/2$ and consider only two terms, just like we did with the Compton scattering. In general we can omit the factor $1/n!$ if we consider only topologically different Feynman diagrams. This will be elucidated and proven later.
    TODO diagram.
    If we write the parts of the operators $\psi^+(x)$ and $\overline{\psi}^-(x)$ that are proportional to $c(j)$ and $c^\dagger(j)$ ($j = 1, 2, 1', 2'$) as

    $$\psi^+_j(x) = c(j)f_j(x) \qquad \overline{\psi}^-_j(x) = c^\dagger(j)g_j(x)$$

    then the relevant part of the $S^{(2)}$-matrix is given by

    $$S^{(2)}(e^-(1) + e^-(2) \to e^-(1') + e^-(2')) = S_a + S_b$$

    with

    $$S_a = -e^2 \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 N\left[(\overline{\psi}^-_{1'}\gamma^\mu\psi^+_1)_{x_1}(\overline{\psi}^-_{2'}\gamma^\nu\psi^+_2)_{x_2}\right] i D_{F\mu\nu}(x_1 - x_2)$$
    $$S_b = -e^2 \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 N\left[(\overline{\psi}^-_{1'}\gamma^\mu\psi^+_1)_{x_1}(\overline{\psi}^-_{2'}\gamma^\nu\psi^+_2)_{x_2}\right] i D_{F\mu\nu}(x_1 - x_2).$$

Notice that S_b is just S_a with the labels $1', 2'$ exchanged. So the transition amplitude is

$$\begin{aligned}\langle f|S^{(2)}(2e^- \to 2e^-)|i\rangle &= \langle f|S_a|i\rangle + \langle f|S_b|i\rangle \\ &= \langle 0|c(1')c(2')S_a c^\dagger(2)c^\dagger(1)|0\rangle + \langle 0|c(1')c(2')S_b c^\dagger(2)c^\dagger(1)|0\rangle \\ &= \langle 0|c(1')c(2')S_a c^\dagger(2)c^\dagger(1)|0\rangle - \langle 0|c(2')c(1')S_b c^\dagger(2)c^\dagger(1)|0\rangle\end{aligned}$$

This means that the transition amplitude is completely antisymmetric under exchange of electrons. (In non-relativistic quantum mechanics this follows from Pauli's exclusion principle.)

This argument generalises: whenever the initial state or final state contains several identical fermions, the transition amplitude $\langle f|S|i\rangle$ is completely antisymmetric.

– The fact that the operator $\psi(x)$ can absorb an electron or create a positron implies that transition amplitudes are antisymmetric with respect to intial electron and final positron states. (A similar argument applies to $\overline{\psi}(x)$ for initial positrons and final electrons.) In order to investigate this, we consider **Bhabha scattering**:

$$e^+ + e^- \to e^+ + e^-.$$

As with the Møller scattering there are four contributions: two identical pairs.

$$S^{(2)}(e^-(1) + e^-(2) \to e^-(1') + e^-(2')) = S_a + S_b$$

with

$$\begin{aligned}S_a &= -e^2 \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 N\left[(\overline{\psi}^-\gamma^\mu\psi^+)_{x_1}(\overline{\psi}^+\gamma^\nu\psi^-)_{x_2}\right] iD_{F\mu\nu}(x_1 - x_2) \\ S_b &= -e^2 \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 N\left[(\overline{\psi}^-\gamma^\mu\psi^-)_{x_1}(\overline{\psi}^+\gamma^\nu\psi^+)_{x_2}\right] iD_{F\mu\nu}(x_1 - x_2).\end{aligned}$$

TODO diagrams.

- The term $S_D^{(2)}$ has two uncontracted fermion fields and gives rise to two processes: **electron self-energy** and **positron self-energy**. TODO diagram. In this process the fermion interacts with itself.
- The term $S_E^{(2)}$ describes a similar effect for photons: **photon self-energy**, also known as **vacuum polerisation**. TODO diagram. This term does, however, introduce a novel feature: a *closed fermion loop*.

$$S^{(2)}(\gamma \to \gamma) = -e^2 \int \mathrm{d}^4x_1\, \mathrm{d}^4x_2 N\left[(\overbrace{\overline{\psi}\not{A}^-\psi)_{x_1}(\overline{\psi}\not{A}^+\psi}^{})_{x_2}\right].$$

Then the normal product can be rewritten as an explicit sum over the spinorial indices $\alpha, \beta, \gamma, \delta$ (remember we assume a sum over repeated indices):

$$\begin{aligned}N\left[(\overbrace{\overline{\psi}_\alpha \not{A}^-_{\alpha\beta}\psi_\beta)_{x_1}(\overline{\psi}_\gamma \not{A}^+_{\gamma\delta}}^{}\psi_\delta)_{x_2}\right] &= (-1)\overbrace{\psi_\delta(x_2)\overline{\psi}_\alpha}(x_1)\not{A}^-_{\alpha\beta}(x_1)\overbrace{\psi_\beta(x_1)\overline{\psi}_\gamma}(x_2)\not{A}^+_{\gamma\delta}(x_2) \\ &= (-1)\,\mathrm{Tr}\left[iS_F(x_2 - x_1)\not{A}^-(x_1)iS_F(x_1 - x_2)\not{A}^+(x_2)\right].\end{aligned}$$

The trace and factor (-1) are characteristic of closed fermion loops.

- The term $S_F^{(2)}$ couples to the vacuum state. For elementary applications such vacuum diagrams may be omitted. TODO diagram.

5.6.3 Feynman diagrams in momentum space

In the last section we calculated terms of the S-matrix operator that were relevant for particular transitions at a given order. In practice, one is usually interested in the corresponding matrix element

$$\langle f|S^{(n)}|i\rangle.$$

The states $|i\rangle$ and $|f\rangle$ are usually specified by particles of known momenta, polarisation and spin, i.e. momentum eigenstates. This means we calculate matrix elements by Fourier transforming the fields and picking out the correct the appropriate operators.
The Fourier transforms of the propagators are

- $\overbrace{\psi(x_1)\overline{\psi}}(x_2) = iS_F(x_1 - x_2) = \dfrac{1}{(2\pi)^4}\displaystyle\int \mathrm{d}^4 p i S_F(p) e^{ip(x_1-x_2)}$

$$S_F(p) \equiv \frac{\not{p} + m}{p^2 - m^2 + i\epsilon} = \frac{1}{\not{p} - m + i\epsilon}$$

- $\overbrace{A^\mu(x_1)A^\nu}(x_2) = iD_F^{\mu\nu}(x_1 - x_2) = \dfrac{1}{(2\pi)^4}\displaystyle\int \mathrm{d}^4 k i D_F^{\mu\nu}(k) e^{-ik(x_1-x_2)}$

$$D_F^{\mu\nu}(k) = \frac{-\eta^{\mu\nu}}{k^2 + i\epsilon}$$

The Fourier of the uncontracted fields $\psi, \overline{\psi}$ and A^μ are given by

$$\begin{aligned}
\psi(x) &= \psi^+(x) + \psi^-(x) \\
&= \sum_{r,\mathbf{p}} \left(\frac{m}{VE_\mathbf{p}}\right)^{1/2} \left[c_r(\mathbf{p}) u_r(\mathbf{p}) e^{-ipx} + d_r^\dagger(\mathbf{p}) v_r(\mathbf{p}) e^{ipx}\right] \\
\overline{\psi}(x) &= \overline{\psi}^+(x) + \overline{\psi}^-(x) \\
&= \sum_{r,\mathbf{p}} \left(\frac{m}{VE_\mathbf{p}}\right)^{1/2} \left[d_r(\mathbf{p}) \overline{v}_r(\mathbf{p}) e^{-ipx} + c_r^\dagger(\mathbf{p}) \overline{u}_r(\mathbf{p}) e^{ipx}\right] \\
A^\mu(x) &= A^{\mu+}(x) + A^{\mu-}(x) \\
&= \sum_{r,\mathbf{k}} \left(\frac{1}{2V\omega_\mathbf{k}}\right)^{1/2} \left[\epsilon_r^\mu(\mathbf{k}) a_r(\mathbf{k}) e^{-ikx} + \epsilon_r^\mu(\mathbf{k}) a_r^\dagger(\mathbf{k}) e^{ikx}\right]
\end{aligned}$$

This gives, for example, after suppression of the spin and polarisation labels:

$$\begin{aligned}
\psi^+(x)\left|e^-\mathbf{p}\right\rangle &= |0\rangle \left(\frac{m}{VE_\mathbf{p}}\right)^{1/2} u(\mathbf{p}) e^{-ipx} & \overline{\psi}^-(x)|0\rangle &= \sum \left|e^-\mathbf{p}\right\rangle \left(\frac{m}{VE_\mathbf{p}}\right)^{1/2} \overline{u}(\mathbf{p}) e^{ipx} \\
\overline{\psi}^+(x)\left|e^+\mathbf{p}\right\rangle &= |0\rangle \left(\frac{m}{VE_\mathbf{p}}\right)^{1/2} \overline{v}(\mathbf{p}) e^{-ipx} & \psi^-(x)|0\rangle &= \sum \left|e^+\mathbf{p}\right\rangle \left(\frac{m}{VE_\mathbf{p}}\right)^{1/2} v(\mathbf{p}) e^{ipx} \\
A_\mu^+(x)|\gamma\mathbf{k}\rangle &= |0\rangle \left(\frac{1}{2V\omega_\mathbf{k}}\right)^{1/2} \epsilon_\mu(\mathbf{k}) e^{-ikx} & A_\mu^-(x)|0\rangle &= \sum |\gamma\mathbf{k}\rangle \left(\frac{1}{2V\omega_\mathbf{k}}\right)^{1/2} \epsilon_\mu(\mathbf{k}) e^{ikx}
\end{aligned}$$

Using these equations it is straightforward to calculate S-matrix elements, as the following example will show.

5.6.3.1 A first order example

We consider the scattering of an electron with emission of a photon. The initial and final states are given by:

$$\begin{cases} |i\rangle = |e^-\mathbf{p}\rangle = c^\dagger(\mathbf{p})\,|0\rangle \\ |f\rangle = |e^-\mathbf{p'};\gamma\mathbf{k'}\rangle = c^\dagger(\mathbf{p'})a^\dagger(\mathbf{k'})\,|0\rangle \end{cases}$$

Writing the only relevant term of $S^{(1)}$, we obtain

$$\begin{aligned}
\langle f|S^{(1)}|i\rangle &= \langle e^-\mathbf{p'};\gamma\mathbf{k'}|ie\int \mathrm{d}^4x\overline{\psi}^-(x)\gamma^\mu A_\mu^-(x)\psi^+(x)|e^-\mathbf{p}\rangle \\
&= ie\int \mathrm{d}^4x\langle e^-\mathbf{p'};\gamma\mathbf{k'}|\overline{\psi}^-(x)\gamma^\mu A_\mu^-(x)|0\rangle\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-ipx} \\
&= ie\int \mathrm{d}^4x\left\langle e^-\mathbf{p'};\gamma\mathbf{k'}\right|\overline{\psi}^-(x)\gamma^\mu\left(\sum_{\mathbf{k}_1}|\gamma\mathbf{k}_1\rangle\left(\frac{1}{2V\omega_{\mathbf{k}_1}}\right)^{1/2}\epsilon_\mu(\mathbf{k}_1)e^{ik_1x}\right)\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-ipx} \\
&= ie\int \mathrm{d}^4x\sum_{\mathbf{k}_1}\left\langle e^-\mathbf{p'};\gamma\mathbf{k'}\right|\overline{\psi}^-(x)\left|\gamma\mathbf{k}_1\right\rangle\gamma^\mu\left(\frac{1}{2V\omega_{\mathbf{k}_1}}\right)^{1/2}\epsilon_\mu(\mathbf{k}_1)e^{ik_1x}\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-ipx} \\
&= ie\int \mathrm{d}^4x\sum_{\mathbf{k}_1}\left\langle e^-\mathbf{p'};\gamma\mathbf{k'}\right|\left(\sum_{\mathbf{p}_1}\left|e^-\mathbf{p}_1;\gamma\mathbf{k}_1\right\rangle\left(\frac{m}{VE_{\mathbf{p}_1}}\right)^{1/2}\overline{u}(\mathbf{p}_1)e^{ip_1x}\right)\gamma^\mu\left(\frac{1}{2V\omega_{\mathbf{k}_1}}\right)^{1/2}\epsilon_\mu(\mathbf{k}_1)e^{ik_1x}\left(\frac{m}{VE_\mathbf{p}}\right)^{1} \\
&= ie\int \mathrm{d}^4x\sum_{\mathbf{k}_1,\mathbf{p}_1}\langle e^-\mathbf{p'};\gamma\mathbf{k'}|e^-\mathbf{p}_1;\gamma\mathbf{k}_1\rangle\left(\frac{m}{VE_{\mathbf{p}_1}}\right)^{1/2}\overline{u}(\mathbf{p}_1)e^{ip_1x}\gamma^\mu\left(\frac{1}{2V\omega_{\mathbf{k}_1}}\right)^{1/2}\epsilon_\mu(\mathbf{k}_1)e^{ik_1x}\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-} \\
&= ie\int \mathrm{d}^4x\left[\left(\frac{m}{VE_{\mathbf{p'}}}\right)^{1/2}\overline{u}(\mathbf{p'})e^{ip'x}\right]\gamma^\mu\left[\left(\frac{1}{2V\omega_{\mathbf{k'}}}\right)^{1/2}\epsilon_\mu(\mathbf{k'})e^{ik'x}\right]\times\left[\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-ipx}\right]
\end{aligned}$$

The x-dependent part of this expression is

$$\int \mathrm{d}^4x e^{ix(p'+k'-p)} = (2\pi)^4\delta^4(p'+k'-p)$$

TODO: explain better + infinite volume and time
Our final result is

$$\langle f|S^{(1)}|i\rangle = \left[(2\pi)^4\delta^4(p'+k'-p)\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}\left(\frac{m}{VE_{\mathbf{p'}}}\right)^{1/2}\left(\frac{1}{2V\omega_{\mathbf{k'}}}\right)^{1/2}\right]\mathcal{M}$$

with

$$\mathcal{M} = ie\overline{u}(\mathbf{p'})\not{\epsilon}(\mathbf{k'})u(\mathbf{p}).$$

The quantity $\mathcal{M}$ is called the Feynman amplitude for the process represented by the Feynman diagram in figure TODO.
The delta function $\delta^4(p'+k'-p)$ ensures conservation of energy and momentum. For real particles ($p^2 = p'^2 = m^2, k'^2 = 0$) (TODO: define on shell somewhere) this can never be satisfied and thus the matrix element is zero.
In general the matrix element of any process to any order can be calculated by bracketing the relevant part of the S-matrix (as discussed before) in side the relevant initial and final momentum eigenstates.
The only new feature in higher order calculations is the appearance of propagators.

5.6.4 A second order example: Compton scattering

The transition is given by

$$\begin{cases} |i\rangle = |e^-\mathbf{p};\gamma\mathbf{k}\rangle = c^\dagger(\mathbf{p})a^\dagger(\mathbf{k})\,|0\rangle \\ |f\rangle = \left|e^-\mathbf{p}';\gamma\mathbf{k}'\right\rangle = c^\dagger(\mathbf{p}')a^\dagger(\mathbf{k}')\,|0\rangle \end{cases}$$

From our previous discussion we know that the relevant parts of the S-matrix are

$$S_a = -e\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\overline{\psi}^-(x_1)\gamma^\mu iS_F(x_1-x_2)\gamma^\nu A^-_\mu(x_1)A^+_\nu(x_2)\psi^+(x_2)$$

$$S_b = -e\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\overline{\psi}^-(x_1)\gamma^\mu iS_F(x_1-x_2)\gamma^\nu A^-_\mu(x_2)A^+_\nu(x_1)\psi^+(x_2)$$

Calculation of the matrix elements for S_a and S_b is straightforward.

$$\begin{aligned}
\langle f|S_a|i\rangle &= -e\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\left\langle e^-\mathbf{p}';\gamma\mathbf{k}'\right|\overline{\psi}^-(x_1)A^-_\mu(x_1)\gamma^\mu iS_F(x_1-x_2)\gamma^\nu A^+_\nu(x_2)\psi^+(x_2)\left|e^-\mathbf{p};\gamma\mathbf{k}\right\rangle \\
&= -e\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\left\langle e^-\mathbf{p}';\gamma\mathbf{k}'\right|\overline{\psi}^-(x_1)A^-_\mu(x_1)\gamma^\mu iS_F(x_1-x_2)\,|0\rangle \\
&\qquad\qquad \times\gamma^\nu\left(\frac{1}{2V\omega_\mathbf{k}}\right)^{1/2}\epsilon_\nu(\mathbf{k})e^{-ikx_2}\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-ipx_2} \\
&= -e\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\left\langle e^-\mathbf{p}';\gamma\mathbf{k}'\right|\overline{\psi}^-(x_1)A^-_\mu(x_1)\gamma^\mu\,|0\rangle\left(\frac{1}{(2\pi)^4}\int \mathrm{d}^4p iS_F(p)e^{ip(x_1-x_2)}\right) \\
&\qquad\qquad \times\gamma^\nu\left(\frac{1}{2V\omega_\mathbf{k}}\right)^{1/2}\epsilon_\nu(\mathbf{k})e^{-ikx_2}\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-ipx_2} \\
&= -e\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\left(\frac{m}{VE_{\mathbf{p}'}}\right)^{1/2}\overline{u}(\mathbf{p}')e^{ip'x}\left(\frac{1}{2V\omega_{\mathbf{k}'}}\right)^{1/2}\not{\epsilon}(\mathbf{k}')e^{ik'x_1}\left(\frac{1}{(2\pi)^4}\int \mathrm{d}^4p iS_F(p)e^{ip(x_1-x_2)}\right) \\
&\qquad\qquad \times\left(\frac{1}{2V\omega_\mathbf{k}}\right)^{1/2}\not{\epsilon}(\mathbf{k})e^{-ikx_2}\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}u(\mathbf{p})e^{-ipx_2}
\end{aligned}$$

The integration over x_1 and x_2 gives

$$\begin{aligned}
\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2e^{ix_1(p'+k'-q)}e^{ix_2(q-p-k)} &= (2\pi)^4\delta^{(4)}(p'+k'-q)(2\pi)^4\delta^{(4)}(q-p-k) \\
&= (2\pi)^4\delta^{(4)}(p'+k'-p-k)(2\pi)^4\delta^{(4)}(q-p-k)
\end{aligned}$$

which fixes the energy-momentum of the virtual intermediate electron. So the matrix element is given by

$$\langle f|S_a|i\rangle = \left[(2\pi)^4\delta^{(4)}(p'+k'-p-k)\left(\frac{m}{VE_\mathbf{p}}\right)^{1/2}\left(\frac{m}{VE_{\mathbf{p}'}}\right)^{1/2}\left(\frac{1}{2V\omega_\mathbf{k}}\right)^{1/2}\left(\frac{1}{2V\omega_{\mathbf{k}'}}\right)^{1/2}\right]\mathcal{M}_a$$

where

$$\mathcal{M}_a = -e^2\overline{u}(\mathbf{p}')\not{\epsilon}(\mathbf{k}')iS_F(q=p+k)\not{\epsilon}(\mathbf{k})u(\mathbf{p})$$

For the matrix element $\langle f|S_b|i\rangle$, the Feynman amplitude $\mathcal{M}_a$ is replaced with

$$\mathcal{M}_b = -e^2\overline{u}(\mathbf{p}')\not{\epsilon}(\mathbf{k})iS_F(q=p-k')\not{\epsilon}(\mathbf{k}')u(\mathbf{p}).$$

We see that the intermediate particle cannot be a real particle because there cannot be energy-momentum conservation for three particles at the vertex, so $q^2 \neq m^2$.
These calculations, while straightforward, are not short. They can be performed substantially quicker using Feynman rules.

5.7 Feynman diagrams and rules in QED

5.7.1 Matrix elements

If the momenta of the initial and final states are specified, the S-matrix element of the transition is given by

$$\langle f|S|i\rangle = \delta_{fi} + \left[(2\pi)^4\delta^4(P_f - P_i)\prod_{\text{ext.}}\left(\frac{m}{VE}\right)^{1/2}\prod_{\text{ext.}}\left(\frac{1}{2V\omega}\right)^{1/2}\right]\mathcal{M} \tag{5.1}$$

where $P_i \equiv \sum_i p_i$ and $P_f \equiv \sum_f p'_f$ are the total four-momenta in the initial and final states and the products are over all external fermions and photons respectively, E and ω being the energies of the individual fermions and photons respectively.
This expression also corresponds to the limit of an infinite time interval $T \to \infty$ and infinite volume $V \to \infty$. In the finite case, the $(2\pi)^4\delta^4(P_f - P_i)$ should be replaced by

$$\int_{-T/2}^{T/2} \mathrm{d}t \int_V \mathrm{d}^3\mathbf{x} e^{ix(P_f - P_i)} \equiv \delta_{TV}(P_f - P_i).$$

Taking its limit gives us the delta function as it should.

5.7.2 Drawing Feynman diagrams

The first step to determining the Feynman amplitude is constructing the the correct Feynman diagrams. A diagram is constructed for each distinct term contributing to the S-matrix.
Assume we want to calculate the matrix element for a particular transition where the initial and final momentum eigenstates are given.
To obtain the Feynman diagrams:

1. Write the initial particles on the left and the final particles on the right.
2. Attach the correct legs to each particle (TODO diagrams)
3. Connect the legs together using QED vertices (TODO).

A couple of remarks need to be made:

- The order of the diagram is the number of QED vertices used.
- Two diagrams are considered the same if one can be obtained from the other by moving around the vertices but keeping the external particles fixed.

5.7.2.1 Two-particle scattering and Mandelstam variables

There are three inequivalent types of second order diagrams that can describe two-particle scattering events (all others can be transformed into one of these by moving vertices). Often processes can be described with two such diagrams; we have already seen several examples (TODO refer).
The diagrams are called **s-channel** (space channel), **t-channel** (time channel) and **u-channel**.
TODO diagrams. + generic two-particle scattering

For any scattering process involving two particles, the Mandelstam variables s, t, u are defined as

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_3 + p_4)^2 \\ t &= (p_1 - p_3)^2 = (p_4 - p_2)^2 \\ u &= (p_1 - p_4)^2 = (p_3 - p_2)^2. \end{aligned}$$

The suggestive naming of these variables is due to the fact that they correspond to the four-momenta of the propagators in the respective diagram. Of course, all the Mandelstam variables are defined regardless of the relevant diagram.
The Mandelstam variables have some interesting properties:

- The sum of the Mandelstam variables obeys

  $$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2$$

  where m_i is the mass of particle i. This can be proven by adding

  $$\begin{cases} s = (p_1 + p_2)^2 = p_1^2 + p_2^2 + 2p_1 \cdot p_2 \\ t = (p_1 - p_3)^2 = p_1^2 + p_3^2 - 2p_1 \cdot p_3 \\ u = (p_1 - p_4)^2 = p_1^2 + P_4^2 - 2p_1 \cdot p_4 \end{cases}$$

  and using $p_i^2 = m_i^2$ we get

  $$\begin{aligned} s + t + u &= m_1^2 + m_2^2 + m_3^2 + m_4^2 + 2p_1^2 + 2p_1 \cdot p_2 - 2p_1 \cdot p_3 - 2p_1 \cdot p_4 \\ &= m_1^2 + m_2^2 + m_3^2 + m_4^2 + 2p_1 \cdot (p_1 + p_2 - p_3 - p_4) \\ &= m_1^2 + m_2^2 + m_3^2 + m_4^2 \end{aligned}$$

  where the last equality follows from conservation of momentum ($p_1 + p_2 - p_3 - p_4 = 0$).

- In the relativistic limit the rest mass can be neglected, so

  $$\begin{aligned} s &\approx 2p_1 \cdot p_2 \approx 2p_3 \cdot p_4 \\ t &\approx -2p_1 \cdot p_3 \approx -2p_2 \cdot p_4 \\ u &\approx -2p_1 \cdot p_4 \approx -2p_3 \cdot p_2 \end{aligned}$$

5.7.3 Determining the Feynman amplitude: Feynman rules

The Feynman amplitude is given by

$$\mathcal{M} = \sum_{n=1}^{\infty} \mathcal{M}^{(n)}$$

where the contribution $\mathcal{M}^{(n)}$ comes from the n^{th} order perturbation term $S^{(n)}$; it can be obtained by adding the contribution from all relevant, distinct Feynman diagrams and following the following Feynman rules:

1. Each element in the diagram has a certain algebraic factor associated to it. We multiply all those factors together to get the Feynman amplitude. We have to be a bit careful about the order we multiply things in, because not all the factors commute.

   The factors are, for each (TODO images)

- Vertex
$$ie\gamma^\mu$$
where μ is a label unique to that vertex;
- Internal photon line
$$iD_{F\,\mu\nu}(k) = i\frac{-\eta_{\mu\nu}}{k^2 + i\epsilon}$$
where μ and ν are the labels of the vertices and k is the momentum;
- Internal fermion line
$$iS_F(p) = i\frac{1}{\not{p} - m + i\epsilon}$$
where p is the momentum;
- External line, one of the following:
  - Initial electron:
$$u_r(\mathbf{p})$$
  - Final electron:
$$\overline{u}_r(\mathbf{p})$$
  - Initial positron:
$$\overline{v}_r(\mathbf{p})$$
  - Final positron:
$$v_r(\mathbf{p})$$
  - Initial photon:
$$\epsilon_{r\,\mu}(\mathbf{k})$$
  - Final photons:
$$\epsilon^*_{r\,\mu}(\mathbf{k})$$

2. The spinor factors along each fermion line, i.e. the γ-matrices S_F-functions and spinors associated with vertices fermion propagators and external fermion lines, are ordered such that, reading from **right to left** they occur in the same sequence as following the fermion line in the direction of its arrows.

3. If a fermionic line starts and ends with an external line, the contributing factor is of the form transpose of spinor times γ-matrices times spinor. This is just a complex number and thus can be moved around the multiplication without changing anything.

4. If a fermionic line does not start and end with an external line, it must be a loop. In this case take the trace and multiply by (-1). Thus we have again obtained a number that commutes with everything.

5. The photonic factors are components of four-vectors or tensors. Being components, they are just numbers and can also be moved about with impunity. In fact we will typically contract them with the γ-matrices.

6. Energy and momentum are conserved at every vertex. If there were a four-momentum q that were not fixed by this, then we carry out the integration
$$(2\pi)^{-4}\int \mathrm{d}^4 q.$$
Such an integration is necessary for each closed loop.

7. Multiply by a phase factor of (-1) if an odd number of interchanges is required to write the fermion operators in the correct normal form.

TODO verify. (including phase 2π, MS p.120) TODO: Do we have all contributions? + factor $1/n!$

5.8 Experimental measurements

5.8.1 Cross section

We have so far spent a lot of time calculating S_{fi}. The is not an experimentally measurable quantity. Instead we need to calculate experimentally verifiable quantities (cross-sections) from it.
We consider a scattering process in which two particles collide (any process that requires more is much less likely) and produces N final particles.

$$p_1 + p_2 \to \sum_{f=1}^{N} p'_f$$

As before, initial and final particles are assumed to be in definite polarisation states; and we can write

$$\langle f|S|i\rangle = \delta_{fi} + (2\pi)^4\delta^4\left(\sum p'_f - \sum p_i\right)\prod_i\left(\frac{1}{2VE_i}\right)^{1/2}\prod_f\left(\frac{1}{2VE'_f}\right)^{1/2}\prod_l(2m_l)^{1/2}\mathcal{M}$$

where the index l runs over all external leptons.
In deriving an expression for the cross section, it is helpful to take the time interval T finite. Then the transition probability per unit time can be defined as

$$w = \frac{|\langle f|S|i\rangle|^2}{T}$$

Square of delta (easier in finite case) TODO????:

$$\left[\delta_{TV}(P_f - P_i)\right]^2 = TV(2\pi)^4\delta^4(P_f - P_i)$$

To obtain the transition rate to a group of final states with momenta in the intervals $(\mathbf{p'}_f, \mathbf{p'}_f + \mathrm{d}\mathbf{p'}_f)$, with of course $f = 1, \ldots, N$, we must multiply w by the number of those states which is

$$\prod_f \frac{V\,\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}.$$

Flux is

$$\Phi = \frac{V}{v_{\mathrm{rel}}}.$$

Differential cross-section

$$\begin{aligned}\mathrm{d}\sigma &= w\frac{V}{v_{\mathrm{rel}}}\prod_f\frac{V\,\mathrm{d}^3\mathbf{p'}}{(2\pi)^3}\\ &= (2\pi)^4\delta^4\left(\sum p'_f - \sum p_i\right)\frac{1}{4E_1E_2v_{\mathrm{rel}}}\left(\prod_l(2m_l)\right)\left(\prod_f\frac{\mathrm{d}^3\mathbf{p'}_f}{(2\pi)^3 2E'_f}\right)|\mathcal{M}|^2\end{aligned}$$

This equation holds for any Lorentz frame in which the colliding particles are moving collinearly. We wish to rewrite $E_1 E_2 v_{\text{rel}}$. From special relativity we know that

$$\begin{cases} p_1^2 = E_1^2 - \mathbf{p_1}^2 = m_1^2 \\ p_2^2 = E_2^2 - \mathbf{p_2}^2 = m_2^2 \end{cases}$$

Because the colliding particles are moving collinearly, we know $\mathbf{p_1} \cdot \mathbf{p_2} = |\mathbf{p_1}||\mathbf{p_2}|$. We compare the quantities

$$\begin{aligned} p_1^2 p_2^2 &= E_1^2 E_2^2 + \mathbf{p_1}^2 \mathbf{p_2}^2 - E_1^2 \mathbf{p_2}^2 - E_2^2 \mathbf{p_1}^2 \\ &= m_1^2 m_2^2 = (m_1 m_2)^2 \end{aligned}$$

and

$$\begin{aligned} (p_1 p_2)^2 &= (E_1 E_2 - \mathbf{p_1} \cdot \mathbf{p_2})^2 \\ &= (E_1 E_2)^2 + \mathbf{p_1}^2 \mathbf{p_2}^2 - 2 E_1 E_2 |\mathbf{p_1}||\mathbf{p_2}|. \end{aligned}$$

Subtracting we get

$$\begin{aligned} (p_1 p_2)^2 - (m_1 m_2)^2 &= E_1^2 \mathbf{p_2}^2 + E_2^2 \mathbf{p_1}^2 - 2 E_1 E_2 |\mathbf{p_1}||\mathbf{p_2}| \\ &= (\mathbf{p}_1 E_2 - \mathbf{p_2} E_1)^2 = \left[E_1 E_2 \left(\frac{\mathbf{p_1}}{E_1} - \frac{\mathbf{p_2}}{E_2} \right) \right]^2 \\ &= (E_1 E_2 v_{\text{rel}})^2 \end{aligned}$$

Thus

$$E_1 E_2 v_{\text{rel}} = [(p_1 p_2)^2 - m_1^2 m_2^2]^{1/2}.$$

We now consider two important and useful frames:

1. In the centre-of-mass frame $\mathbf{p_1} = -\mathbf{p_2}$. Thus

$$\begin{aligned} v_{\text{rel}} &= \left| \frac{\mathbf{p_1}}{E_1} - \frac{\mathbf{p_2}}{E_2} \right| = \left| \frac{\mathbf{p_1}}{E_1} + \frac{\mathbf{p_1}}{E_2} \right| \\ &= \left| \mathbf{p_1} \frac{E_1 + E_2}{E_1 E_2} \right| = |\mathbf{p_1}| \frac{E_1 + E_2}{E_1 E_2} \end{aligned}$$

2. In the laboratory frame $\mathbf{p_2} = 0$. Thus

$$v_{\text{rel}} = \frac{|\mathbf{p_1}|}{E_1}$$

Finally we apply the formula to a process that leads to a two body final state

5.8.2 Decay rate

$$p \to \sum_{f=1}^{N} p'_f$$

5.8.3 Spin and polarisation sums

In most experimental setups colliding beams are unpolarised and final polarisations are not measured. To account for that we average $\left|\overline{\mathcal{M}}\right|^2$ over all initial polarisation states and we sum over all final polarisation states.

$$\left|\overline{\mathcal{M}}\right|^2 = \left(\prod_i \frac{1}{2j_i+1}\right) \sum_{s_i} \sum_{s_f} \left|\mathcal{M}_{fi}\right|^2$$

We consider a Feynman amplitude of the form

$$\mathcal{M}_{fi} = \overline{u}_s(\mathbf{p}')\Gamma u_r(\mathbf{p})$$

The complex conjugate can be calculated as

$$\begin{aligned}\mathcal{M}_{fi}^* = \left(\mathcal{M}_{fi}\right)^\dagger = u_r^\dagger \Gamma^\dagger \overline{u}_s^\dagger = u_r^\dagger \Gamma^\dagger (u_s^\dagger \gamma_0)^\dagger = u_r^\dagger \Gamma^\dagger \gamma_0 u_s = u_r^\dagger \gamma_0 \left(\gamma_0 \Gamma^\dagger \gamma_0\right) u_s \\ = \overline{u}_r \left(\gamma_0 \Gamma^\dagger \gamma_0\right) u_s \equiv \overline{u}_r \tilde{\Gamma} u_s\end{aligned}$$

$$\mathcal{M}^* = (\text{reversed process}) \times (-1)^{\#\text{vertices} \,+\, \#\text{propagators}}$$

where the reversed process is obtained by interchanging initial and final states.

5.8.3.1 External fermions

$$\left|\overline{\mathcal{M}}\right|^2 = \frac{1}{2} \sum_{s_i, s_f} \mathcal{M}_{fi}^* \mathcal{M}_{fi}$$

Using

$$\begin{cases} (\not{p} - m) u_r(p) = 0 & \overline{u}_r(p)(\not{p} - m) = 0 \\ (\not{p} + m) v_r(p) = 0 & \overline{v}_r(p)(\not{p} + m) = 0 \end{cases}$$

and the completeness relation

$$\sum_{s=1}^{2} \left[u_{s\alpha}(p)\overline{u}_{s\beta}(p) - v_{s\alpha}(p)\overline{v}_{s\beta}(p)\right] = \delta_{\alpha\beta}$$

where α and β are spinorial indices, we get

$$\sum_s u_s(p)\overline{u}_s(p) = \frac{\not{p} + m}{2m} \qquad \sum_s v_s(p)\overline{u}_s(p) = \frac{\not{p} - m}{2m}$$

$$\left|\overline{\mathcal{M}}\right|^2 = \frac{1}{2} \operatorname{Tr}\left(\Gamma \frac{\not{p}_i + m}{2m} \tilde{\Gamma} \frac{\not{p}_f + m}{2m}\right)$$

where

$$\tilde{\Gamma} = \gamma^0 \Gamma^\dagger \gamma^0$$

and minus mass if antiparticle

5.8.3.2 External photons

Completeness relation

$$\sum_{\text{pol}} \epsilon_\mu(k)\epsilon^*_\nu(k) = -\eta_{\mu\nu}$$

$$|\overline{\mathcal{M}}|^2 = -\mathcal{M}^\mu \mathcal{M}^*_\mu$$

Ward identity Gauge transformation

$$\epsilon^\mu(k) \to \epsilon'^\mu(k) = \epsilon^\mu(k) + \alpha k^\mu$$

$$\mathcal{M} \to \mathcal{M}' = \mathcal{M} + \alpha k^\mu \mathcal{M}_\mu$$

Thus

$$k^\mu \mathcal{M}_\mu = 0$$

Crossing symmetries

Chapter 6

Quantum electrodynamics

6.1 Other leptons

So far we have only considered the interaction of electrons and positrons with the electromagnetic field. This theory can easily be adapted to describe the interactions of all charged leptons with the electromagnetic field.
The other charged leptons are

- Muon μ^- and anti-muon μ^+
- Tauon τ^- and anti-tauon τ^+

All leptons have spin $\frac{1}{2}$ and charge $\pm e$. In fact within experimental accuracy (which is very high), all leptons have the same properties, except mass:

$$m_\mu = 105.7\,\mathrm{MeV} \qquad m_\tau = 1776.84(17)\,\mathrm{MeV}.$$

This is known as $e-\mu-\tau$ universality. In particular this means each lepton can be described with a Dirac spinor field: $\psi_l(x)$, where $l = e, \mu, \tau$.
The new Lagrangian necessary to incorporate these new particles is given by

$$\begin{cases} \mathcal{L}_0 = \sum_l \overline{\psi}_l(x)\left(i\gamma^\mu\partial_\mu - m_l\right)\psi_l(x) \\ \mathcal{H}_I(x) = -\mathcal{L}_I(x) = -e\sum_l N\left[\overline{\psi}_l(x)\not{A}(x)\psi_l(x)\right]. \end{cases}$$

Each term of the interaction Hamiltonian only contains one type of lepton. Consequently quantum electrodynamics does not allow leptons to change type. The electron, muon and tauon numbers are conserved

$$\begin{cases} N(e) = N(e^-) - N(e^+) \\ N(\mu) = N(\mu^-) - N(\mu^+) \\ N(\tau) = N(\tau^-) - N(\tau^+) \end{cases}$$

The Feynman rules are unchanged by the addition of these other leptons, except now there are three distinct types of fermion line that each stay of the same type along its whole length.

6.2 Processes in lowest order

6.2.1 Muon pair production in (e^-e^+) collision

Process

$$e^+(p_1, r_1) + e^-(p_2, r_2) \to \mu^+(p_1', s_1) + \mu^-(p_2', s_2)$$

6.2.2 Electron-proton elastic scattering

6.2.3 Bhabha scattering

6.2.4 Compton scattering

6.3 Scattering by an external field

In Coulomb gauge

$$A_e^\mu(x) = \left(\frac{Ze}{4\pi|\mathbf{x}|}, 0, 0, 0\right)$$

$$A_e^\mu(q) = \left(\frac{Ze}{|\mathbf{q}|^2}, 0, 0, 0\right)$$

6.4 Radiative corrections

Chapter 7

Renormalization

7.1 Propagator self-energy

7.1.1 Photon self energy

7.1.2 Electron self-energy

7.2 Renormalisability

$$K \equiv 4l - f_i - 2b_i$$

Chapter 8

Regularisation

Chapter 9

Gauge theories

Chapter 10

Path integral formulation

Chapter 11

Axiomatic quantum field theory

Chapter 12

Lattice quantum field theory

Part XIII

Particle physics

Chapter 1

Standard model

1.1 Electroweak interaction

IVB:

$$J_\alpha(x) = \sum_l \bar{\psi}_l(x)\gamma^\alpha(1-\gamma^5)\psi_{\nu_l}(x) = J_V^\alpha(x) - J_A^\alpha(x) = 2\sum_l \bar{\psi}_l^L \gamma^\alpha \psi_{\nu_l}^L$$

$$K_{IVB} = 4l - f_i$$

1.2 Quantum chromodynamics

1.3 Higgs mechanism

1.4 Yukawa

$$\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y.$$

Chapter 2

Accelerators

2.1 Electrostatic

2.2 Cyclotron

2.3 Synchrotron

2.4 Linear-beam

Chapter 3

Detectors

Chapter 4

Open problems

Part XIV

Quantum many particle systems

Chapter 1

Introduction

Some of the tools developed for quantum field theory, in particular second quantisation and Green's functions, greatly simplify the discussion of many identical interacting particles.
The second quantised operators automatically in corporate the Fermi or Bose statistics at each step.
Green's functions allow the use of perturbation theory and Feynman diagrams.
We will often assume $T = 0$. For electrons this is not a very strong assumption as the Fermi temperature for metals is close to room temperature.
TODO spin ¿¡ coordinates

Chapter 2

The Schrödinger equation in second quantisation

We start with the time-dependent Schrödinger equation, which is given by

$$i\hbar\frac{\partial}{\partial t}\Psi(x_1,\ldots,x_N,t) = H\Psi(x_1,\ldots,x_N,t)$$

together with an appropriate set of boundary conditions. The quantity x_k denotes the coordinates of the k^{th} particle. It includes the spatial coordinate $\mathbf{x}_k$ and any discrete variables such as the z component of spin for a system of fermions or the z component of isotopic spin for a system of nucleons. (TODO ????)
Mostly we will only consider time-independent interactions between two particles. The Hamiltonian is then given by:

$$H = \sum_{k=1}^{N} T(x_k) + \frac{1}{2}\sum_{\substack{k,l=1\\k\neq l}}^{N} V(x_k,x_l)$$

We now choose a complete set of time-independent single-particle wave functions that incorporate the boundary conditions and are eigenvalues of the Hamiltonian, $\psi_E(x)$ where E is used to label the functions, in other words it is a quantum number. It is useful to assume the quantum numbers are ordered: $E \in 1,2,\ldots$.
We now wish to expand Ψ in terms of ψ_{E_k}. If Ψ were a function only of x_1, we would be able to write

$$\Psi = \sum_{E_1'} a_{E_1'}\psi_{E_1'}(x_1).$$

where each $a_{E_1'}$ is just a constant. Now Ψ is not only a function, so in general $a_{E_1'}$ depends on $x_2,\ldots,x_N$. This is a function in one fewer variable that Ψ and can be expanded in the same way. Repeating this process, Ψ can be fully expanded as

$$\Psi(x_1,\ldots,x_N,t) = \sum_{E_1',\ldots,E_N'} C(E_1',\ldots,E_N';t)\psi_{E_1'}(x_1)\ldots\psi_{E_N'}(x_N). \tag{2.1}$$

We now write the Schrödinger equation with this expression for Ψ, multiply both side with

$$\prod_{k=1}^{N}\psi_{E_k}^{\dagger}(x_k)$$

where $E_1 \dots E_N$ is a *fixed* set of quantum numbers and integrate over all space.
On the left-hand side this has the effect of projecting out $C(E_1, \dots, E_N, t)$.

$$\begin{aligned}
\int \mathrm{d}x \left(\prod_{k=1}^{N} \psi^\dagger_{E_k}(x_k) \right) \cdot i\hbar \frac{\partial}{\partial t} \Psi &= \int \mathrm{d}x i\hbar \frac{\partial}{\partial t} \left(\prod_{k=1}^{N} \psi^\dagger_{E_k}(x_k) \right) \Psi \\
&= \int \mathrm{d}x i\hbar \frac{\partial}{\partial t} \left(\prod_{k=1}^{N} \psi^\dagger_{E_k}(x_k) \right) \sum_{E'_1, \dots, E'_N} C(E'_1, \dots, E'_N, t) \left(\prod_{k=1}^{N} \psi_{E'_k}(x_k) \right) \\
&= i\hbar \frac{\partial}{\partial t} \sum_{E'_1, \dots, E'_N} C(E'_1, \dots, E'_N, t) \int \mathrm{d}x \left(\prod_{k=1}^{N} \psi^\dagger_{E_k}(x_k) \right) \left(\prod_{k=1}^{N} \psi_{E'_k}(x_k) \right) \\
&= i\hbar \frac{\partial}{\partial t} C(E_1, \dots, E_N, t)
\end{aligned}$$

where $\mathrm{d}x = \mathrm{d}x_1 \dots \mathrm{d}x_N$ and the last step follows because

$$\int \mathrm{d}x_k \psi^\dagger_{E_k}(x_k) \psi_{E'_k}(x_k) = \langle E_k | E'_k \rangle = \delta_{E_k, E'_k}.$$

The right-hand side is a little more complicated. We start by isolating the kinetic energy term.

$$\begin{aligned}
\int \mathrm{d}x \left(\prod_{k=1}^{N} \psi^\dagger_{E_k}(x_k) \right) \sum_{k=1}^{N} T(x_k) \Psi &= \sum_{k=1}^{N} \int \mathrm{d}x \left(\prod_{i=1}^{N} \psi^\dagger_{E_i}(x_i) \right) T(x_k) \sum_{E'_1, \dots, E'_N} C(E'_1, \dots, E'_N, t) \left(\prod_{i=1}^{N} \psi_{E'_i}(x_i) \right) \\
&= \sum_{k=1}^{N} \sum_{E'_k} \int \mathrm{d}x_k \psi^\dagger_{E_k}(x_k) T(x_k) \psi_{E'_k}(x_k) \times C(E_1, \dots, E_{k-1}, E'_k, E_{k+1}, \dots E_N, t) \\
&= \sum_{k=1}^{N} \sum_{W} \langle E_k | T | W \rangle \times C(E_1, \dots, E_{k-1}, W, E_{k+1}, \dots, E_N, t)
\end{aligned}$$

This follows because $T(x_k)$ commutes with all $\psi(x_l)$ with $k \neq l$.
Following a similar procedure for the potential energy term, we get

$$\begin{aligned}
&\frac{1}{2} \sum_{\substack{k,l=1 \\ k \neq l}}^{N} \sum_{W} \sum_{W'} \int \int \mathrm{d}x_k \, \mathrm{d}x_l \psi^\dagger_{E_k}(x_k) \psi^\dagger_{E_l}(x_l) V(x_k, x_l) \psi_W(x_k) \psi_{W'}(x_l) \\
&\qquad \times C(E_1, \dots, E_{k-1}, W, E_{k+1}, \dots, E_{l-1}, W', E_{l+1}, \dots, E_N, t) \\
&= \frac{1}{2} \sum_{\substack{k,l=1 \\ k \neq l}}^{N} \sum_{W} \sum_{W'} \langle E_k E_l | V | W W' \rangle \times C(E_1, \dots, E_{k-1}, W, E_{k+1}, \dots, E_{l-1}, W', E_{l+1}, \dots, E_N, t)
\end{aligned}$$

Now for reference we put all these parts together into one equation

$$\begin{aligned}
i\hbar \frac{\partial}{\partial t} C(E_1, \dots, E_N, t) &= \sum_{k=1}^{N} \sum_{W} \langle E_k | T | W \rangle \times C(E_1, \dots, E_{k-1}, W, E_{k+1}, \dots, E_N, t) \\
&+ \frac{1}{2} \sum_{\substack{k,l=1 \\ k \neq l}}^{N} \sum_{W} \sum_{W'} \langle E_k E_l | V | W W' \rangle \times C(E_1, \dots, E_{k-1}, W, E_{k+1}, \dots, E_{l-1}, W', E_{l+1}, \dots, E_N, t)
\end{aligned} \tag{2.2}$$

Now we incorporate the statistics:

$$\Psi(\dots x_i \dots x_j \dots; t) = \pm\Psi(\dots x_j \dots x_i \dots; t)$$

Because the ψs commute in the expansion 2.1, we see that the change of sign must come from the coefficients:

$$C(\dots E_i \dots E_j \dots; t) = \pm C(\dots E_j \dots E_i \dots; t)$$

2.1 Bosons

For bosons the coefficients are symmetric under exchange of quantum numbers. This mean we can group all of the same states together. For example

$$C(1,2,1,3,2,4,\dots;t) = C(\underbrace{1,1,1,\dots}_{n_1},\underbrace{2,2,2,\dots}_{n_2},\dots;t)$$

where we have introduced the occupation numbers $n_1, n_2, \dots$. These occupation numbers uniquely determine the coefficients, so we can write

$$\bar{C}(n_1, n_2, \dots, n_\infty; t) \equiv C(\underbrace{1,1,1,\dots}_{n_1},\underbrace{2,2,2,\dots}_{n_2},\dots;t)$$

Now we require Ψ to be normalised to unity. We assume the single particle wave functions ψ_{E_k} to already be properly normalised. Looking back at equation 2.1, we see that this yields the condition

$$\sum_{E_1,\dots,E_N} |C(E_1,\dots,E_N,t)|^2 = 1.$$

For an expression in terms of occupation numbers, we split the sum into two factors: A sum over all possible values of all occupation numbers and a sum over all possible values of all quantum numbers E_k, given the occupation numbers fixed by the first sum:

$$\sum_{n_1,\dots n_\infty} \sum_{\substack{E_1,\dots,E_N \\ (n_1,\dots n_\infty)}} |\bar{C}(n_1,\dots,n_\infty;t)|^2 = \sum_{n_1,\dots n_\infty} |\bar{C}(n_1,\dots,n_\infty;t)|^2 \sum_{\substack{E_1,\dots,E_N \\ (n_1,\dots n_\infty)}} 1 = 1.$$

In our notation expressions in brackets under summations are conditions that the things we are summing over must fulfill.
The second sum is the equivalent of putting N object into boxes with n_1 objects in the first box, n_2 in the second etc.

$$\sum_{\substack{E_1,\dots,E_N \\ (n_1,\dots n_\infty)}} 1 = \frac{N!}{n_1! n_2! \dots n_\infty!}$$

The normalisation condition thus becomes

$$\sum_{n_1,\dots n_\infty} |\bar{C}(n_1,\dots,n_\infty;t)|^2 \frac{N!}{n_1! n_2! \dots n_\infty!} = 1.$$

This is more simply expressed if we define another coefficient

$$f(n_1,\dots,n_\infty;t) \equiv \left(\frac{N!}{n_1! n_2! \dots n_\infty!}\right)^{1/2} \bar{C}(n_1,\dots,n_\infty;t)$$

rendering the normalisation condition

$$\sum_{n_1,\ldots n_\infty} |f(n_1,\ldots,n_\infty;t)|^2 = 1.$$

The original wave function can be rewritten in therms of these coefficients:

$$\Psi(x_1,\ldots,x_N;t) = \sum_{n_1,\ldots,n_\infty\left(\sum_i n_i = N\right)} f(n_1,\ldots,n_\infty;t)\Phi_{n_1 n_2 \ldots n_\infty}(x_1,\ldots,x_N) \tag{2.3}$$

where we have defined

$$\Phi_{n_1 n_2 \ldots n_\infty}(x_1,\ldots,x_N) \equiv \left(\frac{n_1! n_2! \ldots n_\infty!}{N!}\right)^{1/2} \sum_{\substack{E_1,\ldots,E_N \\ (n_1,\ldots n_\infty)}} \psi_{E_1}(x_1)\ldots\psi_{E_N}(x_N)$$

In other words the functions $\Phi_{n_1 n_2 \ldots n_\infty}$ form a complete set. They are also completely symmetrised and orthonormal.
We can now return to the analysis of equation 2.2. Now, crucially, we can do the following substitution:

$$\sum_{k=1}^{N} \quad \rightarrow \quad \sum_E n_E. \tag{2.4}$$

This follows because for every k that E_k takes the same value, the same term is contributed to the sum.
The substitution is straightforward for the kinetic energy term. In the potential energy term, the condition $k \neq l$ complicates matters somewhat. The solution is to perform the substitution

$$\sum_{\substack{k,l=1 \\ k\neq l}}^{N} \quad \rightarrow \quad \sum_E \sum_{E'} n_E(n_{E'} - \delta_{EE'}). \tag{2.5}$$

The resulting equations are horribly long, but are simplified massively by using the creation and destruction operators $b_k^\dagger, b_k$ from the second quantisation.

2.1.1 Creation and destruction operators

We remind ourselves that the time-independent creation and destruction operators $b_k^\dagger$ and b_k satisfy

$$\begin{cases} [b_k, b_{k'}^\dagger] = \delta_{kk'} \\ [b_k, b_{k'}] = [b_k^\dagger, b_{k'}^\dagger] = 0 \end{cases}.$$

Consequently

$$\begin{cases} b_k |n_k\rangle = \sqrt{n_k}\,|n_k - 1\rangle \\ b_k^\dagger |n_k\rangle = \sqrt{n_k + 1}\,|n_k + 1\rangle \end{cases}.$$

We can rewrite the equation 2.3 in abstract sate vector notation

$$|\Psi(t)\rangle = \sum_{n_1,\ldots,n_\infty} f(n_1,\ldots,n_\infty;t)\,|n_1 n_2 \ldots n_\infty\rangle\,.$$

The number operator $(b_k^\dagger b_k)$ for different modes commute, so (TODO why)

$$|n_1 n_2 \dots n_\infty\rangle = |n_1\rangle\,|n_2\rangle \dots |n_\infty\rangle\,.$$

Now finally we take equation 2.2 and we apply the following manipulations (which we will not make explicit because of the length of the equations):

1. Rewrite in terms of $\bar{C}$:
$$\begin{cases} C(E_1 \dots E_N; t) = \bar{C}(n_1, \dots, n_\infty; t) \\ C(E_1, \dots, E_{k-1}, W, E_{k+1}, \dots, E_N; t) = \bar{C}(n_1, \dots, n_{E_k} - 1, \dots, n_W + 1, \dots, n_\infty; t) \end{cases}$$

2. Apply substitutions 2.4 and 2.5. Now we are no longer summing over indices k of E_k, but over values of E, E', W, W'. We rename these i, j, k, l for ease.

3. Rewrite in terms of the coefficients f. In order to do that we split up the sum
$$\sum_{i,j,k,l} = \sum_{i\neq j\neq k\neq l} + \sum_{i=j\neq k\neq l} + \dots$$

4. Plug the result into the equation
$$i\hbar\frac{\partial}{\partial t}\,|\Psi(t)\rangle = \sum_{n_1,\dots,n_\infty} i\hbar\frac{\partial}{\partial t} f(n_1, \dots, n_\infty; t)\,|n_1 n_2 \dots n_\infty\rangle$$
Finally we relabel the occupation numbers and using
$$\sqrt{n_i' + 1}\sqrt{n_j'}\,\left|n_1' \dots n_i' + 1 \dots n_j' - 1 \dots n_\infty'\right\rangle = b_i^\dagger b_j\,|n_1' n_2' \dots n_\infty'\rangle$$
we get
$$i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle = \hat{H}\,|\psi(t)\rangle$$
with
$$\boxed{\hat{H} = \sum_{i,j} b_i^\dagger b_j \langle i|T|j\rangle + \frac{1}{2}\sum_{i,j,k,l} b_i^\dagger b_j^\dagger \langle ij|V|kl\rangle b_l b_k}\,.$$

This resulting equation in second quantisation is completely equivalent to the originl one. The matrix elements $\langle i|T|j\rangle$ and $\langle ij|V|kl\rangle$ are just complex numbers.
For bosons the order $b_l b_k$ does not matter, we write it in this order however because for Fermions the expression will be formally the same and there the creation and annihilation operators do not commute.

2.2 Fermions

For fermions the wavefunction is antisymmetric:

$$C(\dots E_i \dots E_j \dots; t) = -C(\dots E_j \dots E_i \dots; t)$$

Consequently the occupation numbers are either one or zero. We can analogously define $\bar{C}(n_1, n_2, \dots, n_\infty; t)$, $f(n_1, n_2, \dots, n_\infty; t)$ and $\Phi_{n_1 n_2 \dots n_\infty}$, except now $\bar{C}$ can pick up a minus sign and the wave functions Φ are given by the Slater determinant

$$\Phi_{n_1 n_2 \dots n_\infty}(x_1 \dots x_N) = \sqrt{\frac{\prod_i n_i!}{N!}}\begin{vmatrix} \varphi_{E_1^0}(x_1) & \dots & \varphi_{E_1^0}(x_N) \\ \vdots & & \vdots \\ \varphi_{E_N^0}(x_1) & \dots & \varphi_{E_N^0}(x_N) \end{vmatrix}$$

2.2.1 Creation and destruction operators

The operators for fermions satisfy the anti-commutator relations

$$\begin{cases} \left\{a_i, a_j^\dagger\right\} = \delta_{ij} \\ \{a_i, a_j\} = \left\{a_i^\dagger, a_j^\dagger\right\} = 0 \end{cases}$$

$$\hat{H} = \sum_{ij} a_i^\dagger a_j \langle i|T|j\rangle + \frac{1}{2}\sum_{ijkl} a_i^\dagger a_j^\dagger \langle ij|V|kl\rangle a_l a_k$$

Consequently

$$\begin{cases} a^\dagger |0\rangle = |1\rangle & a^\dagger |1\rangle = 0 \\ a|1\rangle = |0\rangle & a|0\rangle = 0 \end{cases}$$

We can write the state vectors using creation operators

$$|n_1 n_2 \dots n_\infty\rangle = (a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2}\dots(a_\infty^\dagger)^{n_\infty}|0\rangle$$

. Because of the anti-commutator relations, we need to keep track of the sign. For example

$$a_j |n_1 n_2 \dots n_\infty\rangle = (-1)^{n_1+n_2+\dots+n_{j-1}}(a_1^\dagger)^{n_1}\dots(a_j a_j^\dagger)\dots(a_\infty^\dagger)^{n_\infty}|0\rangle$$

so

$$a_j^\dagger |n_1 n_2 \dots n_\infty\rangle = (-1)^{n_1+n_2+\dots+n_{j-1}}|n_1 \dots (n_j+1)\dots n_\infty\rangle$$

(if $n_j = 0$, otherwise we get 0). Similarly

$$a_j |n_1 \dots n_j \dots n_\infty\rangle = (-1)^{n_1+n_2+\dots+n_{j-1}}|n_1 \dots (n_j-1)\dots n_\infty\rangle$$

(if $n_j = 1$). We call $n_1 + n_2 + \dots + n_{j-1} \equiv S_j$.
The number operator introduces no extra phases

$$a_j^\dagger a_j |\dots n_j \dots\rangle = n_j |\dots n_j \dots\rangle$$

The phase signs introduced by the operators are exactly the ones introduced by the reordering necessary to define $\bar{C}$.
So finally we get a very similar result for the Hamiltonian in second quantization for fermions:

$$\hat{H} = \sum_{i,j} a_i^\dagger a_j \langle i|T|j\rangle + \frac{1}{2}\sum_{i,j,k,l} a_i^\dagger a_j^\dagger \langle ij|V|kl\rangle a_l a_k$$

Now the order of the operators is important ($a_l a_k = -a_k a_l$). The order also ensures $\hat{H}$ is Hermitian.

2.3 Fields

In first quantization:

$$\mathbf{p} = \sum^N ???$$

Where i is a particle index. In the second quantization this becomes

$$\hat{\mathbf{p}} = \sum_{i,j} a_i^\dagger a_j \langle i|T|j\rangle ????????$$

Where i, j are state indices.
For fermions:

$$\mathbf{p} = \hbar\mathbf{k} \qquad S_z = \pm 1 \qquad |\mathbf{S}| = \tfrac{1}{2}$$

$$\varphi_{E_i,\alpha}(\mathbf{x}) = \begin{cases} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{\sqrt{V}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{\sqrt{V}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{cases} = \varphi_{\mathbf{k},\alpha}(\mathbf{x})$$

Quantum field theory so fields. We use $c^\dagger, c$ to denote the a and b operators.

$$\hat{\psi}_\alpha(x) = \sum_i \varphi_{E_i,\alpha}(\mathbf{x}) c_{i,\alpha}$$

With i spin index (the book does not use them)

$$\hat{\psi}_\alpha^\dagger(x) = \sum_i \varphi_{E_i,\alpha}^\dagger(\mathbf{x}) c_{i,\alpha}^\dagger$$

These operators create or annihilate particles in a particular point in space so they are field operators in the second quantization. So we shift our focus from a give state to a give point in space.

$$\begin{aligned} \sum_{i,j} a_i^\dagger a_j \langle i|T|j\rangle &= \sum_{\substack{i,j \\ \alpha,\beta}} a_{i,\alpha}^\dagger \int \mathrm{d}^3x \varphi_{E_i,\alpha}^\dagger(\mathbf{x}) \left(\frac{-\hbar^2\nabla^2}{2m}\right) \varphi_{E_j,\beta}(\mathbf{x}) a_{j,\beta} \\ &= \sum_{\alpha,\beta} \int \mathrm{d}^3x \left(\sum_i a_{i,\alpha}^\dagger \varphi^\dagger\right) ....(???) \\ &= \sum_{\alpha,\beta} \int \mathrm{d}^3x \hat{\psi}_\alpha^\dagger(\mathbf{x}) \left(\frac{-\hbar^2\nabla^2}{2m}\right) \hat{\psi}_\beta(\mathbf{x}) \end{aligned}$$

We write

$$\begin{aligned} \left[\hat{\psi}_\alpha(\mathbf{x}), \hat{\psi}_\beta^\dagger(\mathbf{x})\right]_\pm &= \sum_j \sum_l \varphi_{E_j,\alpha}(\mathbf{x}) \varphi_{E_l,\beta}^\dagger(\mathbf{x}) \left[c_j, c_l^\dagger\right]_\pm \\ &= ??? \\ &= \delta(\mathbf{x} - \mathbf{x'}) \delta_{\alpha,\beta} ???? \end{aligned}$$

Where minus is for bosons, plus for fermions
So

$$\hat{H} = \int \mathrm{d}^3x \hat{\psi}_\alpha^\dagger(\mathbf{x}) T(\mathbf{x}) \hat{\psi}_\alpha(\mathbf{x}) + \frac{1}{2} \int \mathrm{d}^3x \int \mathrm{d}^3x' \hat{\psi}_\alpha^\dagger(\mathbf{x}) \hat{\psi}_\beta^\dagger(\mathbf{x'}) V(\mathbf{x}, \mathbf{x'}) \hat{\psi}_\beta(\mathbf{x'}) \hat{\psi}_\alpha(\mathbf{x})$$

In first quantization:

$$\int \mathrm{d}x \psi^\dagger(x) V(x) \psi(x)$$

Where the ψs are wavefunctions and V is an operator
In second quantization:

$$\int \mathrm{d}x \hat{\psi}^\dagger(x) V(x) \hat{\psi}(x)$$

Now the $\hat{\psi}$s are operators and Vx is a complex number.
One body operator
First quantization (with i particle index):

$$J = \sum_{i=1}^{N} J(x_i)$$

Second quantization (with i, j state indices):

$$\hat{J} = \sum_{i,j} \langle i|J|j\rangle c_i^\dagger c_j = \int \mathrm{d}^3 x \hat{\psi}^\dagger(\mathbf{x}) J(x) \hat{\psi}(\mathbf{x})$$

Number density operator:

$$n(\mathbf{x}) = \sum_{i=1}^{N} \delta(\mathbf{x} - \mathbf{x_i})$$

So

$$\int \mathrm{d}^3 x' \hat{\psi}(\mathbf{x'}) \delta(\mathbf{x} - \mathbf{x'}) \hat{\psi}(\mathbf{x'}) = \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}(\mathbf{x})$$

$$= \sum \varphi ??$$

Total number operator:

$$\hat{N} \equiv \int \mathrm{d}^3 x \hat{n}(\mathbf{x}) = \sum c_i^\dagger c_i = \sum_i \hat{n}_i = \int \mathrm{d}^3 x \hat{\psi}(\mathbf{x}) \hat{\psi}(\mathbf{x})$$

$$\left[\hat{N}, \hat{H}\right] = 0 \qquad \left[\hat{\mathbf{P}}, \hat{H}\right] = 0 = \left[\hat{\mathbf{J}}, \hat{H}\right]$$

2.4 Example: the jellium model

Now a real application of this technique to a physical system. The system is the degenerate electron gas ("Jellium model"). This is a very simple model in the sense that you can think of your system as a cubic box with side L ($V = L^3$). The only constraint on the electrons is that they must be inside the box. We are mainly interested in bulk properties, so we will be working in the thermodynamic limit ($V \to \infty, N \to \infty$ keeping $n = \frac{N}{V}$ finite). In order to confine the electrons we apply a uniform positive background (gel in the box) so the whole thing is charge neutral.
It is a surprisingly good model for various situation (such as electrons in metals, white dwarf stars, neutron stars). There is a pressure due to the Pauli exclusion principle
For this problem a reasonable set of single particle wave functions incorporating the boundary conditions is given by the plane wave states

$$\psi_{\mathbf{k}\lambda} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{\sqrt{V}} \eta_\lambda$$

where η_λ are the two spin functions for spin-up and spin-down along a chosen z axis.

$$\eta_\uparrow = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \eta_\downarrow = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The periodic boundary conditions determine the allowed wavenumbers as

$$k_i = \frac{2\pi n_i}{L} \qquad i = x, y, z \qquad n_i = 0, \pm 1, \pm 2, \ldots$$

The Hamiltonian can be written as the sum of three terms,

$$\hat{H} = \hat{H}_{el} + \hat{H}_b + \hat{H}_{el-b}$$

where the first term arises from the interactions between electrons, the second is the energy of the positive background and the third is the interaction energy between the electrons and the positive background.
We will write down expressions for all the terms. The terms diverge individually, but together give a meaningful result. We include an exponential factor in order to force all the integrals converge. Later we will take the limit of $\mu \to 0$ (after we have taken the thermodynamic limit) so that the exponential factors do not influence the physics.

$$\begin{aligned} H_b &= \frac{1}{2} \int_V \mathrm{d}\mathbf{x} \int_V \mathrm{d}\mathbf{x'} \frac{\rho(\mathbf{x})\rho(\mathbf{x'})}{|\mathbf{x} - \mathbf{x'}|} e^{-\mu|\mathbf{x}-\mathbf{x'}|} \\ &= \frac{e^2}{2} \int_V \mathrm{d}\mathbf{x} \int_V \mathrm{d}\mathbf{x'} \frac{n_b(\mathbf{x}) n_b(\mathbf{x'})}{|\mathbf{x} - \mathbf{x'}|} e^{-\mu|\mathbf{x}-\mathbf{x'}|} \\ &= \frac{e^2 n^2}{2} \int_V \mathrm{d}\mathbf{x} \int_V \mathrm{d}\mathbf{y} \frac{1}{|y|} e^{-\mu|y|} \end{aligned}$$

Here $n = N/V$ is the number density of the electrons (the density ρ of the positive background must equal $e \cdot n$ in order for the system to be neutral). Evaluating these integrals (where we already assume $L \to \infty$ for the integral over $\mathbf{y}$ and use polar coordinates), we get

$$H_b = \frac{1}{2} e^2 \frac{N^2}{V} \frac{4\pi}{\mu^2}$$

Then for interaction between the electrons and the background, we get

$$\begin{aligned} H_{el-b} &= -e^2 \frac{N}{V} \sum_{i=1}^{N} \int_V \mathrm{d}^3\mathbf{x} \frac{e^{-\mu(|\mathbf{x}-\mathbf{r}_i|)}}{|\mathbf{x} - \mathbf{r}_i|} \\ &= -e^2 \frac{N}{V} \sum_{i=1}^{N} \int_V \mathrm{d}^3\mathbf{z} \frac{e^{-\mu|\mathbf{z}|}}{|\mathbf{z}|} \\ &= -e^2 \frac{N^2}{V} \frac{4\pi}{\mu^2} \end{aligned}$$

So far we have

$$H = -\frac{1}{2} e^2 \frac{N^2}{S} \frac{2\pi}{\mu} + H_{el}$$

We now consider the Hamiltonian in second quantization. Given that H_b and H_{el-b} are just numbers, they stay the same.

$$\begin{aligned}\hat{H}_{el} &= \hat{T} + \hat{V} \\ &= \sum_{ij} a_i^\dagger \langle i|T|j\rangle a_j + \frac{1}{2}\sum_{ijkl} a_i^\dagger a_j^\dagger \langle ij|V|V\rangle kl a_l a_k\end{aligned}$$

First we tackle the matrix element in the kinetic energy term.

$$\begin{aligned}\langle i|T|j\rangle &= \langle \mathbf{k}\lambda|T|\mathbf{k}'\lambda'\rangle \\ &= \langle \mathbf{k}\lambda|\frac{\hbar^2\nabla^2}{2m}|\mathbf{k}'\lambda'\rangle \\ &= \frac{\hbar^2 k'^2}{2m}\langle \mathbf{k}\lambda|\mathbf{k}'\lambda'\rangle \\ &= \frac{\hbar^2 k^2}{2m}\delta_{\lambda\lambda'}\delta_{\mathbf{k}\mathbf{k}'}\end{aligned}$$

So the kinetic energy operator is

$$\hat{T} = \sum_{\mathbf{k}\lambda} \frac{\hbar^2 k^2}{2m} a^\dagger_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda}$$

Next comes the matrix element in the potential energy term.

$$\begin{aligned}\langle ij|V|kl\rangle &= \langle \mathbf{k'}\lambda', \mathbf{p'}\mu'|V|\mathbf{k}\lambda, \mathbf{p}\mu\rangle \\ &= \int \mathrm{d}\mathbf{x}\int \mathrm{d}\mathbf{x'}\psi\,\dagger_{\mathbf{k'},\lambda'}(\mathbf{x})\psi^\dagger_{\mathbf{p'},\mu'}(\mathbf{x'})V(\mathbf{x}-\mathbf{x'})\psi_{\mathbf{k},\lambda}(\mathbf{x})\psi_{\mathbf{p},\mu}(\mathbf{x'}) \\ &= \frac{e^2}{V^2}\int\int \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{x'} e^{-i\mathbf{k'}\cdot\mathbf{x}}\eta_{\lambda'}(1)^\dagger e^{-i\mathbf{p'}\cdot\mathbf{x'}}\eta_{\mu'}(2)^\dagger \frac{e^{-\mu|\mathbf{x}-\mathbf{x'}|}}{|\mathbf{x}-\mathbf{x'}|} e^{i\mathbf{k}\cdot\mathbf{x}}\eta_\lambda(1)e^{i\mathbf{p}\cdot\mathbf{x'}}\eta_\mu(2) \\ &= \frac{e^2}{V}\delta_{\lambda\lambda'}\delta_{\mu\mu'}\int \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{x'} e^{i\mathbf{x}(\mathbf{k}-\mathbf{k'})}e^{i\mathbf{x'}(\mathbf{p}-\mathbf{p'})}\frac{e^{\mu|\mathbf{x}-\mathbf{x'}|}}{|\mathbf{x}-\mathbf{x'}|} \\ &= \frac{e^2}{V}\delta_{\lambda\lambda'}\delta_{\mu\mu'}\int \mathrm{d}\mathbf{y}\left(\mathrm{d}\mathbf{x'}e^{i\mathbf{x'}(\mathbf{k}-\mathbf{k'})}e^{i\mathbf{x'}(\mathbf{p}-\mathbf{p'})}\right)e^{i\mathbf{y}(\mathbf{k}-\mathbf{k'})}\frac{e^{\mu|\mathbf{y}|}}{|\mathbf{y}|} \\ &= \frac{e^2}{V}\delta_{\lambda\lambda'}\delta_{\mu\mu'}\int \mathrm{d}\mathbf{y} e^{i\mathbf{y}(\mathbf{k}-\mathbf{k'})}\frac{e^{\mu\mathbf{y}}}{\mathbf{y}}\delta_{\mathbf{k}+\mathbf{p},\mathbf{k'}+\mathbf{p'}}\end{aligned}$$

Where $\mathbf{y} = \mathbf{x} - \mathbf{x}'$ and we have used that

$$\int \mathrm{d}^3 x e^{i(\mathbf{k_2}-\mathbf{k_1})\cdot\mathbf{x}} = \delta_{\mathbf{k_1}\mathbf{k_2}}$$

Solving this integral, we get

$$\langle ij|V|kl\rangle = \frac{e^2}{V}\delta_{\lambda\lambda'}\delta_{\mu\mu'}\delta_{\mathbf{k}+\mathbf{p},\mathbf{k'}+\mathbf{p'}}\frac{4\pi}{(\mathbf{k}-\mathbf{k'})^2+\mu^2}$$

This means we get the following expression for $\hat{V}$:

$$\begin{aligned}\hat{V} &= \frac{e^2}{S}\sum_{\substack{\mathbf{k'},\lambda',\mathbf{p'},\mu'\\ \mathbf{k},\lambda,\mathbf{p},\mu}} \frac{2\pi}{q+\mu}\delta_{\mathbf{k}-\mathbf{k'},\mathbf{p'}-\mathbf{p}}\delta_{\lambda\lambda'}\delta_{\mu\mu'}\cdot a^\dagger_{\mathbf{k'},\lambda'}a^\dagger_{\mathbf{p'},\mu'}a_{\mathbf{p},\mu}a_{\mathbf{k},\lambda}\\ &= \frac{e^2}{S}\sum_{\substack{\mathbf{k},\mathbf{p},\mathbf{q}\\ \lambda,\mu}}\frac{2\pi}{q+\mu}a^\dagger_{\mathbf{k}+\mathbf{q},\lambda}a^\dagger_{\mathbf{p}-\mathbf{q},\mu}a_{\mathbf{p},\mu}a_{\mathbf{k},\lambda}\end{aligned}$$

Where we have defined $\mathbf{q} = \mathbf{k'} - \mathbf{k} = \mathbf{q} - \mathbf{q'}$
We split this sum into a part with $q = 0$ and a part with $q \neq 0$.

$$\hat{V} = \frac{e^2}{2S}\sum_{\substack{\mathbf{kpq}\\ \lambda\mu}}\frac{2\pi}{\mu}a^\dagger_{\mathbf{k},\lambda}a^\dagger_{\mathbf{p},\mu}a_{\mathbf{p},\mu}a_{\mathbf{k},\lambda} + \frac{e^2}{2S}\sum_{\substack{\mathbf{kp}\\ \mathbf{q}\neq 0\\ \lambda\mu}}\frac{2\pi}{q+\mu}a^\dagger_{\mathbf{k}+\mathbf{q},\lambda}a^\dagger_{\mathbf{p}-\mathbf{q},\mu}a_{\mathbf{p},\mu}a_{\mathbf{k},\lambda}$$

We have

$$[\hat{n}_{\mathbf{p}\mu}, a_{\mathbf{k},\lambda}] = -\delta_{\mathbf{kp}}\delta_{\mu\lambda}a_{\mathbf{k}\lambda}$$

so

$$a^\dagger_{\mathbf{p}\mu}a_{\mathbf{p}\mu}a_{\mathbf{k}\lambda} = a_{\mathbf{k}\lambda}a^\dagger_{\mathbf{p}\mu}a_{\mathbf{p}\mu} - \delta_{\mathbf{kp}}\delta_{\mu\lambda}a_{\mathbf{k}\lambda}$$

If we now consider the part with $\mathbf{q} = 0$

$$\begin{aligned}\hat{V}_{\mathbf{q}=0} &= \frac{e^2}{2}\sum_{\substack{\mathbf{kp}\\ \lambda\mu}}\frac{2\pi}{\mu S}a^\dagger_{\mathbf{k},\lambda}a^\dagger_{\mathbf{p},\mu}a_{\mathbf{p},\mu}a_{\mathbf{k},\lambda}\\ &= \frac{e^2}{2}\sum_{\substack{\mathbf{kp}\\ \lambda\mu}}\frac{2\pi}{\mu S}a^\dagger_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda}\left[a^\dagger_{\mathbf{p}\mu}a_{\mathbf{p}\mu} - \delta_{\mathbf{kp}}\delta_{\mu\lambda}\right]\end{aligned}$$

Rewriting in terms of the number operator we get

$$\begin{aligned}\hat{V}_{\mathbf{q}=0} &= \frac{e^2}{2}\frac{2\pi}{\mu S}\sum_{\mathbf{k}\lambda}(\hat{n}_{\mathbf{k}\lambda})\left(\sum_{\mathbf{p}\mu}\hat{n}_{\mathbf{p}\mu}\right) - \frac{e^2}{2S}\frac{2\pi}{\mu}\left(\sum_{\mathbf{k}\lambda}\hat{n}_{\mathbf{k}\lambda}\right)\\ &= \frac{e^2}{2S}\frac{2\pi}{\mu}\left(N^2 - N\right)\\ &= \frac{e^2}{2}\frac{2\pi}{\mu}\frac{N^2}{S} - \frac{e^2}{2}\frac{2\pi}{\mu}\frac{N}{S}\end{aligned}$$

When we take our limit the second term goes to zero in the first step. The first term however remains and diverges in the second step. Luckily it exactly cancels the other diverging contributions to the Hamiltonian ($H_b + H_{el-b} = -\frac{1}{2}e^2\frac{N^2}{S}\frac{2\pi}{\mu}$). All divergences have now been dealt with, so we can always take $\mu \to 0$.

$$\hat{H} = \sum_{\mathbf{k}\lambda}\frac{\hbar^2k^2}{2m}a^\dagger_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda} + \frac{2\pi e^2}{2S}{\sum_{\substack{\mathbf{kpq}\\ \lambda\mu\\ \mathbf{q}\neq\mathbf{0}}}}'\frac{1}{q}a^\dagger_{\mathbf{k}+\mathbf{q},\lambda}a^\dagger_{\mathbf{p}-\mathbf{q},\mu}a_{\mathbf{p},\mu}a_{\mathbf{k},\lambda}$$

We write the apostrophe to remind ourselves that $\mathbf{q} = 0$ is not considered.
Compared to the three dimensional case, the factor of 2π is still different and we divide by q instead of q^2.
We now define some lengths, first r_0 in terms of the surface area per particle.

$$S = \pi r_0^2 N$$

Then we have the Bohr radius

$$a_0 = \frac{\hbar^2}{me^2}$$

Finally we define a dimensionless quantity

$$r_s \equiv \frac{r_0}{a_0}$$

We now use perturbation theory, so we write the Hamiltonian as $\hat{H} = \hat{H}_0 + \hat{H}_1$. In the high density limit ($r_s \to 0$) we can view $\hat{H}_1$ as a small perturbation.

$$\hat{H}_0 = \sum_{\mathbf{k}\lambda} \frac{\hbar^2 \bar{k}^2}{2m} a^\dagger_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} = \sum_{\mathbf{k}\lambda} \frac{\hbar^2 \bar{k}^2}{2m} \hat{n}_{\mathbf{k}\lambda}$$

In the high density limit we have a Fermi sphere. ($|\mathbf{k}| \leq k_F$)
We can determine the Fermi wavenumber k_F by computing the expectation value of the number operator in the ground state $|F\rangle$.

$$\begin{aligned} N = \langle F|\hat{N}|F\rangle &= \sum_{\mathbf{k}\lambda} \langle F|\hat{n}_{\mathbf{k}\lambda}|F\rangle \\ &= \sum_{\mathbf{k}\lambda} \theta(k_F - k) \\ &= (\Delta k)^{-1} \sum_\lambda \int \mathrm{d}^2 k \theta(k_F - k) \\ &= S(2\pi)^{-2} \cdot 2 \int \mathrm{d}^2 k \theta(k_F - k) \\ &= \frac{S}{2\pi} k_F^2 \end{aligned}$$

So we get

$$k_F = \sqrt{\frac{2\pi N}{S}} = \frac{\sqrt{2}}{r_0}$$

This expression is significantly different from the three dimensional case, due to a difference in Δk and the fact that we used the surface of a circle, not the volume of a sphere.

We can now calculate E_0.

$$\begin{aligned}
E_0 = \langle F|\hat{H}_0|F\rangle &= \sum_{\mathbf{k},\lambda} \frac{\hbar^2 k^2}{2m}\langle F|\hat{n}_{\mathbf{k}\lambda}|F\rangle \\
&= 2\sum_{|\mathbf{k}|\leq k_F} \frac{\hbar^2 k^2}{2m} \\
&= \frac{\hbar^2}{m} S(2\pi)^{-2}\int_0^{k_F} \mathrm{d}^2 k k^2 \\
&= \frac{\hbar^2}{m} S(2\pi)^{-1}\int_0^{k_F} \mathrm{d}k k^3 \\
&= \frac{\hbar^2}{m} N\frac{k_F^2}{4} \\
&= \frac{\hbar^2}{m} N\frac{1}{2r_0^2} \\
&= N\frac{e^2}{2a_0 r_s^2}
\end{aligned}$$

And we also calculate E_1, recalling that

$$\hat{H}_1 = \frac{e^2}{2S}{\sum_{\substack{\mathbf{kpq}\\ \lambda\mu}}}' \frac{2\pi}{q} a^\dagger_{\mathbf{k}+\mathbf{q},\lambda} a^\dagger_{\mathbf{p}-\mathbf{q},\mu} a_{\mathbf{p},\mu} a_{\mathbf{k},\lambda}$$

We calculate

$$E_1 = \langle F|\hat{H}_1|F\rangle = \frac{e^2}{2S}{\sum_{\substack{\mathbf{kpq}\\ \lambda\mu}}}' \frac{2\pi}{q}\langle F|a^\dagger_{\mathbf{k}+\mathbf{q},\lambda} a^\dagger_{\mathbf{p}-\mathbf{q},\mu} a_{\mathbf{p},\mu} a_{\mathbf{k},\lambda}|F\rangle$$

As in the three dimensional case only the exchange term contributes. So the matrix element becomes

$$\begin{aligned}
\delta_{\mathbf{k}+\mathbf{q},\mathbf{p}}\delta_{\lambda,\mu}\langle F|a^\dagger_{\mathbf{k}+\mathbf{q},\lambda} a^\dagger_{\mathbf{k},\lambda} a_{\mathbf{k}+\mathbf{q},\lambda} a_{\mathbf{k},\lambda}|F\rangle &= -\delta_{\mathbf{k}+\mathbf{q},\mathbf{p}}\delta_{\lambda,\mu}\langle F|\hat{n}_{\mathbf{k}+\mathbf{q},\lambda}\hat{n}_{\mathbf{k},\lambda}|F\rangle \\
&= -\delta_{\mathbf{k}+\mathbf{q},\mathbf{p}}\delta_{\lambda,\mu}\theta(k_F - |\mathbf{k}+\mathbf{q}|)\theta(k_F - k)
\end{aligned}$$

So the first order potential energy is given by the following expression:

$$\begin{aligned}
E_1 &= -\frac{e^2}{2S}{\sum_{\substack{\mathbf{kq}\\ \lambda}}}' \frac{2\pi}{q}\theta(k_F - |\mathbf{k}+\mathbf{q}|)\theta(k_F - k) \\
&= -\frac{e^2}{2}\frac{2\pi S}{(2\pi)^4} 2\int \mathrm{d}^2 k\, \mathrm{d}^2 q q^{-2}\theta(k_F - |\mathbf{k}+\mathbf{q}|)\theta(k_F - k) \\
&= -\frac{e^2 S}{(2\pi)^3}\int \mathrm{d}^2 P\, \mathrm{d}^2 q q^{-2}\theta(k_F - |\mathbf{P}+\tfrac{\mathbf{q}}{2}|)\theta(k_F - |\mathbf{P}-\tfrac{\mathbf{q}}{2}|)
\end{aligned}$$

Where $\mathbf{P} = \mathbf{k} + \frac{1}{2}\mathbf{q}$. In order to evaluate this integral, we first calculate an expression for a circle segment of a circle of radius R of height h:

$$R^2 \arccos\left(\frac{R-h}{R}\right) - (R-h)\sqrt{2Rh - h^2}$$

The integral over $\mathbf{P}$ yields a surface equals to two such segments with height $k_F - \frac{q}{2}$ from circles of radius k_F. Thus we have

$$\begin{aligned}\int \mathrm{d}^2 P \theta(k_F - |\mathbf{P} + \tfrac{\mathbf{q}}{2}|)\theta(k_F - |\mathbf{P} - \tfrac{\mathbf{q}}{2}|) &= 2\left(k_F^2 \arccos\left(\frac{q}{2k_F}\right) - \frac{q}{2}\sqrt{k_F^2 - \frac{q^2}{4}}\right)\theta(2k_F - q) \\ &= 2k_F^2\left(\arccos(x) - x\sqrt{1-x^2}\right)\theta(1-x)\end{aligned}$$

With $x = \frac{q}{2k_F}$
Therefor we get for the first order potential energy

$$\begin{aligned}E_1 &= -\frac{e^2 S}{(2\pi)^3} 2k_F^2 2k_F \int_0^1 2\pi\left(\arccos(x) - x\sqrt{1-x^2}\right) \\ &= -\frac{e^2 S}{(2\pi)^2} 4k_F^3\left(1 - \frac{1}{3}\right) \\ &= -\frac{e^2 S}{(2\pi)^2}\frac{8k_F^3}{3} \\ &= -\frac{e^2}{2\pi} N \frac{8}{3} k_F \\ &= -\frac{e^2}{2\pi} N \frac{8\sqrt{2}}{3 a_0 r_s}\end{aligned}$$

This expression is again completely different from the three dimensional case, mostly due to the different geometry.
So the ground state energy per particle in the high-density limit is given approximately as

$$\frac{E}{N} = \frac{e^2}{2a_0}\left[\frac{1}{r_s^2} - \frac{8\sqrt{2}}{3\pi}\frac{1}{r_s} + \dots\right]$$

The minimum is at

$$r_s = \frac{3\pi}{4\sqrt{2}} \approx 1.666$$

and the minimum energy per particle is

$$\frac{E}{N} = -0.36\frac{e^2}{2a_0}$$

This means that compared to the three dimensional case, r_s is smaller (so the density is higher) and the binding energy is higher.

Recap of the Jellium model

$$\frac{E}{N} \underset{r_S \to 0}{=} \left(\frac{e^2}{2a_0}\right)\left[\underbrace{\frac{2.21}{r_S^2}}_{\text{kinetic}} - \underbrace{\frac{0.916}{r_S}}_{\text{exchange}} + \underbrace{\dots}_{\text{correlation energy}}\right]$$

The name correlation energy was introduced by Wigner. It was called stupidity energy by Feynman.
We will try in this course to calculate more terms in this perturbative expansion.

Putting all the contributions together we get the following figure: (see notes)
Variational principle

$$\langle F|\hat{H}|F\rangle \geq \langle 0|\hat{H}|0\rangle$$

(with F the Fermi state and 0 the exact state)
But is our approximation actually any good? We consider bulk sodium metal. Cohesive energy

$$\begin{cases} \frac{E}{N} = -1.13eV \\ r_S = 3.96 \end{cases}$$

So our crude model is not too bad.
A first question is why is the real energy higher? In the real case we don't have a uniform background, but localised charges. So electrons have less space to move around. This means a larger momentum (from the uncertainty principle) and thus higher kinetic energy and a positive contribution to the total energy.
For real metals r_S is between 2 and 6.
Low density limit ($r_S \rightarrow \infty$). Wigner (1938). Solid. Lattice of electrons that only move a bit around equilibrium in positive background. In this case we can apply classical electrostatic approaches. (Madelung, Ewald method). The result is

$$\frac{E}{N} = \frac{e^2}{2a_0}\left[\frac{-1.79}{r_s} + \frac{2.66}{r_s^{3/2}} + \dots\right]$$

Real metals are between the high and low density limits.

$$E_{\text{corr}} = 0.0622 \ln r_S + const. + \mathcal{O}(r_S \ln r_S)$$

Read chapter 2 (useful for later) We move on to chapter 3

Chapter 3

Green's functions in the ground-state formalism

The single-particle Green's function is defined by

$$iG_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \frac{\langle]|T\left[\hat{\psi}_{H\alpha}(\mathbf{x},t)\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}',t')\right|\Psi_0\rangle\Psi_0}{\langle\Psi_0|\Psi_0\rangle}$$

where $|\Psi_0\rangle$ is the Heisenberg ground state of the interacting system satisfying

$$\hat{H}\,|\Psi_0\rangle = E\,|\Psi_0\rangle\,,$$

$\hat{\psi}_{H\alpha}(\mathbf{x},t)$ is a Heisenberg operator, i.e. with time-dependence given by

$$\hat{\psi}_{H\alpha}(\mathbf{x},t) = e^{i\hat{H}t/\hbar}\hat{\psi}_{\alpha}(\mathbf{x})e^{-i\hat{H}t/\hbar}$$

and α and β label the components of the field operator.
If $\hat{H}$ is time independent, then G depends only on the time difference $t-t'$:

$$iG_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \begin{cases} e^{iE(t-t')/\hbar}\frac{\langle]|T\left[\hat{\psi}_{H\alpha}(\mathbf{x})e^{-i\hat{H}(t-t')/\hbar}\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}')\right|\Psi_0\rangle\Psi_0}{\langle\Psi_0|\Psi_0\rangle} & t>t' \\ \pm e^{-iE(t-t')/\hbar}\frac{\langle]|T\left[\hat{\psi}_{H\alpha}(\mathbf{x})e^{i\hat{H}(t-t')/\hbar}\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}')\right|\Psi_0\rangle\Psi_0}{\langle\Psi_0|\Psi_0\rangle} & t'>t \end{cases}$$

Green's functions are studied because the Feynman rules for finding their n^{th} order contributions in perturbation theory are relatively simple but they still contain properties of great interest

- expectation value of each single particle operator in the ground state of the system
- ground-state total energy
- elementary excitations.

3.1 Relation to observables

3.1.1 Expectation value of single-particle operators

In general a second-quantized operator is given by

$$\hat{J} = \int \mathrm{d}^3x \sum_{\alpha\beta} \hat{\psi}^{\dagger}_{\beta}(\mathbf{x})J_{\beta\alpha}(\mathbf{x})\hat{\psi}_{\alpha}(\mathbf{x}).$$

This suggests the following expression for the second-quantized density $\hat{\mathcal{J}}(\mathbf{x})$:

$$\hat{\mathcal{J}}(\mathbf{x}) = \sum_{\alpha\beta} \hat{\psi}^\dagger_\beta(\mathbf{x}) J_{\beta\alpha}(\mathbf{x}) \hat{\psi}_\alpha(\mathbf{x}).$$

The ground-state expectation value or the operator density is given by

$$\begin{aligned}
\left\langle \hat{\mathcal{J}}(\mathbf{x}) \right\rangle &\equiv \frac{\langle \Psi_0 | \hat{\mathcal{J}}(\mathbf{x}) | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \\
&= \lim_{\mathbf{x'} \to \mathbf{x}} \sum_{\alpha\beta} J_{\beta\alpha}(\mathbf{x}) \frac{\langle \Psi_0 | \hat{\psi}^\dagger_\beta(\mathbf{x'}) \hat{\psi}_\alpha(\mathbf{x}) | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \\
&= \pm i \lim_{t' \to t^+} \lim_{\mathbf{x'} \to \mathbf{x}} \sum_{\alpha\beta} J_{\beta\alpha}(\mathbf{x}) G_{\alpha\beta}(\mathbf{x}, t; \mathbf{x'}, t') \\
&= \pm i \lim_{t' \to t^+} \lim_{\mathbf{x'} \to \mathbf{x}} \operatorname{Tr}[J(\mathbf{x}) G(\mathbf{x}, t; \mathbf{x'}, t')]
\end{aligned}$$

Some examples:

- Number density:
$$\langle \hat{n}(\mathbf{x}) \rangle = \pm i \operatorname{Tr} G(\mathbf{x}, t; \mathbf{x}, t^+)$$
- Total kinetic energy:
$$\left\langle \hat{T} \right\rangle = \pm i \int \mathrm{d}^3 x \lim_{\mathbf{x'} \to \mathbf{x}} \left[-\frac{\hbar^2 \nabla^2}{2m} \operatorname{Tr} G(\mathbf{x}, t; \mathbf{x'}, t^+) \right]$$

3.1.2 Ground-state energy

We now have an expression for the kinetic energy. We would also like an expression for the potential energy so we can add both contributions to obtain the ground-state energy. The potential is not a single particle operator, so we need to find a new trick: we will use the Schrödinger equation to extract the potential energy.
Assuming $\langle \Psi_0 | \Psi_0 \rangle$ is normalised, we wish to compute the following:

$$\left\langle \hat{V} \right\rangle = \frac{1}{2} \sum_{\substack{\alpha\alpha' \\ \beta\beta'}} \int \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x'} \langle \Psi_0 | \hat{\psi}^\dagger_\alpha(\mathbf{x}) \hat{\psi}^\dagger_\beta(\mathbf{x}) V_{\substack{\alpha\alpha' \\ \beta\beta'}}(\mathbf{x}, \mathbf{x'}) \hat{\psi}_{\beta'}(\mathbf{x'}) \hat{\psi}_{\alpha'}(\mathbf{x'}) | \Psi_0 \rangle$$

For any Heisenberg operator, the following holds:

$$i\hbar \frac{\partial}{\partial t} \hat{\psi}_{H\alpha}(\mathbf{x}, t) = e^{i\hat{H}t/\hbar} \left[\hat{\psi}_{H\alpha}(\mathbf{x}), \hat{H} \right] e^{-i\hat{H}t/\hbar} = \left[\hat{\psi}_\alpha(\mathbf{x}, t), \hat{H} \right]$$

Using the expression $\hat{T} = \sum_\beta \int \mathrm{d}\mathbf{x} \hat{\psi}^\dagger_\alpha(\mathbf{x}) T(\mathbf{x}) \hat{\psi}_\alpha(\mathbf{x})$ for the kinetic energy in second quantization, we see that

$$\left[\hat{\psi}_\alpha(\mathbf{x}), \hat{T} \right] = \sum_\beta \int \mathrm{d}\mathbf{y} \lim_{\mathbf{y'} \to \mathbf{y}} \left(\frac{-\hbar^2 \nabla^2}{2m} \right) \left[\hat{\psi}_\alpha(\mathbf{x}), \hat{\psi}^\dagger_\beta(\mathbf{y'}) \hat{\psi}_\beta(\mathbf{y}) \right]$$

Next we appreciate the identity

$$\begin{aligned}[A,BC] &= ABC - BCA = ABC - BAC + BAC - BCA \\ &= \begin{cases} [A,B]\,C - B\,[C,A] \\ \{A,B\}\,C - B\,\{C,A\}\,. \end{cases}\end{aligned}$$

This allows us to write the previous expression in terms of either commutators **or** anti-commutators. Thus the following will be correct for both bosons and fermions.
So

$$\begin{aligned}\left[\hat{\psi}_\alpha(\mathbf{x}), \hat{T}\right] &= \sum_\beta \int \mathrm{d}\mathbf{y} \lim_{\mathbf{y'}\to\mathbf{y}} \left(\frac{-\hbar^2\nabla^2}{2m}\right) \delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{y'})\hat{\psi}_\beta(\mathbf{y}) \\ &= -\frac{\hbar^2\nabla^2 x}{2m}\hat{\psi}_\alpha(\mathbf{x}) \\ &= T(\mathbf{x})\hat{\psi}_\alpha(\mathbf{x})\end{aligned}$$

For the potential energy we quite simply need to apply the identity twice: we split the product of four field operators in the right part of the commutator into two and two, and then into (anti-)commutators of single field operators we can compute.

$$\begin{aligned}\left[\hat{\psi}_\alpha(\mathbf{x}), \hat{V}\right] &= \frac{1}{2}\sum_{\substack{\beta\beta' \\ \gamma\gamma'}} \int \mathrm{d}\mathbf{y} \int \mathrm{d}\mathbf{y'} \left[\hat{\psi}_\alpha(\mathbf{x}), \hat{\psi}^\dagger_\beta(\mathbf{y})\hat{\psi}^\dagger_\gamma(\mathbf{y'}) \underset{\gamma\gamma'}{V_{\beta\beta'}}(\mathbf{y},\mathbf{y'})\hat{\psi}_{\gamma'}(\mathbf{y'})\hat{\psi}_{\beta'}(\mathbf{y})\right] \\ &= -\frac{1}{2}\sum_{\beta\beta'\gamma'} \int \mathrm{d}\mathbf{y}\hat{\psi}^\dagger_\beta(\mathbf{z}) \underset{\alpha\gamma'}{V_{\beta\beta'}}(\mathbf{y},\mathbf{x})\hat{\psi}_{\gamma'}(\mathbf{x})\hat{\psi}_{\beta'}(\mathbf{y}) + \frac{1}{2}\sum_{\beta'\gamma\gamma'} \int \mathrm{d}\mathbf{y'}\hat{\psi}^\dagger_\gamma(\mathbf{z'}) \underset{\gamma\gamma'}{V_{\alpha\beta'}}(\mathbf{x},\mathbf{y'})\hat{\psi}_{\gamma'}(\mathbf{y'})\hat{\psi}_{\beta'}(\mathbf{x})\end{aligned}$$

In the first term we can change the dummy variables

$$\begin{cases} \beta \to \gamma \\ \beta' \to \gamma' \\ \gamma' \to \beta' \\ \mathbf{y} \to \mathbf{y'} \end{cases}$$

and using the anticommutivity of the fields $\hat{\psi}$ as well as the symmetry of the potential

$$\underset{\beta\beta'}{V_{\alpha\alpha'}}(\mathbf{x},\mathbf{x'}) = \underset{\alpha\alpha'}{V_{\beta\beta'}}(\mathbf{x'},\mathbf{x})$$

we obtain

$$\left[\hat{\psi}_\alpha(\mathbf{x}), \hat{V}\right] = \sum_{\beta'\gamma\gamma'} \int \mathrm{d}\mathbf{y'}\hat{\psi}^\dagger_\gamma(\mathbf{z'}) \underset{\gamma\gamma'}{V_{\alpha\beta'}}(\mathbf{x},\mathbf{y'})\hat{\psi}_{\gamma'}(\mathbf{y'})\hat{\psi}_{\beta'}(\mathbf{x}).$$

Filling all this into the equation for the time evolution of a Heisenberg operator, we get

$$i\hbar\frac{\partial}{\partial t}\hat{\psi}_{H\alpha}(\mathbf{x},t) = \left[\hat{\psi}_{H\alpha}(\mathbf{x},t), \hat{H}\right] = -\frac{\hbar^2\nabla^2_x}{2m}\hat{\psi}_{H\alpha}(\mathbf{x},t) + \sum_{\beta'\gamma\gamma'} \int \mathrm{d}\mathbf{y'}\hat{\psi}^\dagger_{H\gamma}(\mathbf{y'},t) \underset{\gamma\gamma'}{V_{\alpha\beta'}}(\mathbf{x},\mathbf{z'})\hat{\psi}_{H\gamma'}(\mathbf{y'},t)\hat{\psi}_{H\beta'}(\mathbf{x},t).$$

Rearranging slightly:

$$\left[i\hbar\frac{\partial}{\partial t} - T(\mathbf{x})\right]\hat{\psi}_{H\alpha}(\mathbf{x},t) = \sum_{\beta'\gamma\gamma'}\int \mathrm{d}\mathbf{y'}\hat{\psi}^{\dagger}_{H\gamma}(\mathbf{y'},t)V_{\substack{\alpha\beta'\\\gamma\gamma'}}(\mathbf{x},\mathbf{z'})\hat{\psi}_{H\gamma'}(\mathbf{y'},t)\hat{\psi}_{H\beta'}(\mathbf{x},t).$$

To simplify we multiply by $\hat{\psi}_{H\alpha}(\mathbf{x'},t')$ on the left and take the ground-state expectation value.

$$\left[i\hbar\frac{\partial}{\partial t} - T(\mathbf{x})\right]\langle\Psi_0|\hat{\psi}^{\dagger}_{H\alpha}(\mathbf{x'},t')\hat{\psi}_{H\alpha}(\mathbf{x},t)|\Psi_0\rangle =$$
$$\sum_{\beta'\gamma\gamma'}\int \mathrm{d}\mathbf{y'}\langle\Psi_0|\hat{\psi}^{\dagger}_{H\alpha}(\mathbf{x'},t')\hat{\psi}^{\dagger}_{H\gamma}(\mathbf{y'},t)V_{\substack{\alpha\beta'\\\gamma\gamma'}}(\mathbf{x},\mathbf{z'})\hat{\psi}_{H\gamma'}(\mathbf{y'},t)\hat{\psi}_{H\beta'}(\mathbf{x},t)|\Psi_0\rangle.$$

In the limit $\mathbf{x'}\to\mathbf{x}, t'\to t^{+}$, the left side is equal to

$$\pm i\lim_{t'\to t^{+}}\lim_{\mathbf{x'}\to\mathbf{x}}\left[i\hbar\frac{\partial}{\partial t} - T(\mathbf{x})\right]G_{\alpha\alpha}(\mathbf{x},t;\mathbf{x'},t').$$

Summing over α and integrating over $\mathbf{x}$, the right hand side becomes double the ground-state expectation value for the potential energy, $\left\langle\hat{V}\right\rangle$.

$$\left\langle\hat{V}\right\rangle = \pm i\frac{1}{2}\int \mathrm{d}\mathbf{x}\lim_{t'\to t^{+}}\lim_{\mathbf{x'}\to\mathbf{x}}\sum_{\alpha}\left[i\hbar\frac{\partial}{\partial t} - T(\mathbf{x})\right]G_{\alpha\alpha}(\mathbf{x},t;\mathbf{x'},t')$$

Finally we have an expression for the ground-state energy solely in terms of the single particle Green's function:

$$\begin{aligned}E &= \left\langle\hat{T}+\hat{V}\right\rangle = \left\langle\hat{H}\right\rangle\\ &= \pm i\frac{1}{2}\int \mathrm{d}\mathbf{x}\lim_{\substack{\mathbf{x'}\to\mathbf{x}\\t'\to t^{+}}}\left[i\hbar\frac{\partial}{\partial t} + T(\mathbf{x})\right]\mathrm{Tr}\,G(\mathbf{x},t;\mathbf{x'},t')\end{aligned}$$

3.1.2.1 Green's functions in momentum space

We assume the system is homogeneous and in a large box of volume V. ($\left[\hat{\mathbf{p}},\hat{H}\right] = 0; G_{\alpha\beta}(\mathbf{x}-\mathbf{x}',t-t')$
We can then write the single particle Green's function as

$$G_{\alpha\beta}(\mathbf{x}t,\mathbf{x}'t') \equiv \sum_{\mathbf{k}}\frac{1}{V}\int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{2\pi}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}e^{-i\omega(t-t')}G_{\alpha\beta}(\mathbf{k},\omega)$$

In the limit $V\to\infty$, the sum becomes an integral

$$G_{\alpha\beta}(\mathbf{x}t,\mathbf{x}'t') = (2\pi)^{-4}\int \mathrm{d}\mathbf{k}\int_{-\infty}^{+\infty}\mathrm{d}\omega e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}e^{-i\omega(t-t')}G_{\alpha\beta}(\mathbf{k},\omega)$$

We get the following expression for the total energy:

$$\begin{aligned}E &= \pm\frac{1}{2}i\int \mathrm{d}\mathbf{x}\lim_{\substack{\mathbf{x'}\to\mathbf{x}\\t'\to t^{+}}}\frac{1}{(2\pi)^4}\int \mathrm{d}\mathbf{k}\int_{-\infty}^{+\infty}\mathrm{d}\omega\left(\hbar\omega + \frac{\hbar^2k^2}{2m}\right)e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}e^{-i\omega(t-t')}\,\mathrm{Tr}\,G[\mathbf{k},\omega]\\ &= \pm\frac{1}{2}i\frac{V}{(2\pi)^4}\lim_{\eta\to 0^{+}}\int \mathrm{d}\mathbf{k}\int_{-\infty}^{+\infty}\mathrm{d}\omega e^{i\omega\eta}\left(\frac{\hbar^2k^2}{2m} + \hbar\omega\right)\mathrm{Tr}\,G(\mathbf{k},\omega)\end{aligned}$$

Where we have defined $\eta \equiv t' - t$.
The total number of particles is given by

$$N = \int \mathrm{d}\mathbf{x}\, \langle \hat{n}(\mathbf{x}) \rangle = \pm i \frac{V}{(2\pi)^4} \lim_{\eta \to 0^+} \int \mathrm{d}\mathbf{k} \int_{-\infty}^{\infty} \mathrm{d}\omega e^{i\omega\eta} \operatorname{Tr} G(\mathbf{k}, \omega)$$

3.2 Calculating the Green's function for free, noninteracting fermions

We now attempt to calculate the single-particle Green's function for free, noninteracting fermion, which we call G^0. In this case $\hat{H} = \hat{H}_0 = \hat{T}$ and $\hat{H}_0 |F\rangle = E_0 |F\rangle$ if $k < k_F$, i.e. the fermion is inside the Fermi sphere.
It is convenient to perform a transformation to holes and particles. In the definition of the field

$$\hat{\psi}(\mathbf{x}) = \sum_{\mathbf{k},\lambda} \psi_{\mathbf{k}\lambda}(\mathbf{x}) c_{\mathbf{k}\lambda}$$

we redefine the fermion operator $c_{\mathbf{k}\lambda}$ as

$$c_{\mathbf{k}\lambda} = \begin{cases} a_{\mathbf{k}\lambda} & (|\mathbf{k}| > k_F) \quad \text{particles} \\ b^\dagger_{-\mathbf{k}\lambda} & (|\mathbf{k}| \leq k_F) \quad \text{holes} \end{cases}$$

The anti-commutation rules are preserved:

$$\{a_{\mathbf{k}}, a^\dagger_{\mathbf{k}'}\} = \{b_{\mathbf{k}}, b^\dagger_{\mathbf{k}'}\} = \delta_{\mathbf{k}\mathbf{k}'}$$

$$\{a_{\mathbf{k}}, b^\dagger_{\mathbf{k}'}\} = 0$$

This means it's canonical transformation the physics is preserved.

$$\begin{aligned} \hat{H}_0 &= \sum_{\mathbf{k}\lambda} \epsilon_{\mathbf{k}} c^\dagger_{\mathbf{k}\lambda} c^\dagger_{\mathbf{k}\lambda} \\ &= \sum_{\substack{|\mathbf{k}| \leq k_F \\ \lambda}} \epsilon_{\mathbf{k}} b_{\mathbf{k}\lambda} b^\dagger_{\mathbf{k}\lambda} + \sum_{\substack{|\mathbf{k}| > k_F \\ \lambda}} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} \\ &= \sum_{\substack{|\mathbf{k}| \leq k_F \\ \lambda}} \epsilon_{\mathbf{k}} - \sum_{\substack{|\mathbf{k}| \leq k_F \\ \lambda}} b^\dagger_{\mathbf{k}\lambda} b_{\mathbf{k}\lambda} + \sum_{\substack{|\mathbf{k}| > k_F \\ \lambda}} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} \\ &= \text{filled Fermi sea} + \text{holes} + \text{particles} \end{aligned}$$

Where $\epsilon_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}$ and remembering that $b_{\mathbf{k}\lambda} b^\dagger_{\mathbf{k}\lambda} = 1 - b^\dagger_{\mathbf{k}\lambda} b_{\mathbf{k}\lambda}$.
If there are no particles or holes, the energy is that of the Fermi sea. Creating holes lowers the energy, whereas creating a particle raises the energy. If the total number of fermions is fixed, particles and holes occur in pairs. Because $k_{\text{particle}} > k_F \geq k_{\text{hole}}$ each particle-hole pair has has a net positive energy. So the filled Fermi sea represents the ground state.
In order to calculate the Green's function, we need the operator fields in terms of these operators. In the Schrödinger picture this is straightforward

$$\hat{\psi}_{S\alpha}(\mathbf{x}) = \sum_{|\mathbf{k}| > k_F} \psi_{\mathbf{k}\alpha}(\mathbf{x}) a_{\mathbf{k}\alpha} + \sum_{|\mathbf{k}| > k_F} \psi_{\mathbf{k}\alpha}(\mathbf{x}) b^\dagger_{-\mathbf{k}\alpha}$$

In order to get the field in the Heisenberg picture

$$\hat{\psi}_{H\alpha}(\mathbf{x},t) = e^{i\hat{H}t/\hbar}\hat{\psi}_{\alpha}(\mathbf{x})e^{-i\hat{H}t/\hbar}$$

with $\hat{\psi}_{\alpha}(\mathbf{x}) = \sum_{\mathbf{k}} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{\sqrt{V}}\eta_{\alpha}c_{\mathbf{k}\alpha}$, we need to evaluate

$$\begin{aligned} e^{i\hat{H}_0t/\hbar}c_{\mathbf{k}\alpha}e^{-i\hat{H}_0t/\hbar} &= \sum_{n=0}^{\infty}\frac{1}{n!}\left[\frac{i}{\hbar}\hat{H}_0t,\left[\frac{i}{\hbar}\hat{H}_0t,\ldots\left[\frac{i\hat{H}_0t}{\hbar},c_{\mathbf{k}\alpha}\right]\ldots\right]\right] \\ &= \sum_{n=0}^{\infty}\frac{(it/\hbar)^n}{n!}\left[\hat{H}_0,\left[\hat{H}_0,\ldots\left[\hat{H}_0,c_{\mathbf{k}\alpha}\right]\ldots\right]\right] \end{aligned}$$

where $\hat{H}_0$ is the same as $\hat{H}$ because we are working with a non-interacting system.
The innermost commutator is

$$\begin{aligned} \left[\hat{H}_0,c_{\mathbf{k}\alpha}\right] &= -\left[c_{\mathbf{k}\alpha},\hat{H}_0\right] = -\sum_{\mathbf{k}'\alpha'}\epsilon_{\mathbf{k}'}\left[c_{\mathbf{k}\alpha},c^{\dagger}_{\mathbf{k}'\alpha'}c_{\mathbf{k}'\alpha'}\right] \\ &= -\epsilon_{\mathbf{k}}c_{\mathbf{k}\alpha} \end{aligned}$$

Where we have used our cool identity. As a result we get

$$\begin{aligned} e^{i\hat{H}_0t/\hbar}c_{\mathbf{k}\alpha}e^{-i\hat{H}_0t/\hbar} &= \sum_{n=0}^{\infty}\frac{(it/\hbar)^n(-\epsilon_{\mathbf{k}})^n}{n!}c_{\mathbf{k}\alpha} \\ &= e^{-\frac{i}{\hbar}\epsilon_{\mathbf{k}}t}c_{\mathbf{k}\alpha} \end{aligned}$$

We also introduce $\omega_{\mathbf{k}}$ such that $\hbar\omega_{\mathbf{k}} = \epsilon_{\mathbf{k}}$and then get

$$\hat{\psi}_{H\alpha}(\mathbf{x}) = \sum_{|\mathbf{k}|>k_F}\psi_{\mathbf{k}\alpha}(\mathbf{x})e^{-i\omega_{\mathbf{k}}t}a_{\mathbf{k}\alpha} + \sum_{|\mathbf{k}|>k_F}\psi_{\mathbf{k}\alpha}(\mathbf{x})e^{-i\omega_{\mathbf{k}}t}b^{\dagger}_{-\mathbf{k}\alpha}$$

Now we are ready to tackle the Green's function proper, which by definition is given by

$$iG^0_{\alpha\beta}(\mathbf{x}t,\mathbf{x}'t') = \langle]|T[\hat{\psi}_{H\alpha}(\mathbf{x}t)\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}',t')|\Phi_0\rangle\Phi_0$$

The particle and hole destructor operators both annihilate the ground state:

$$b_{\mathbf{k}\alpha}\left|\Phi_0\right\rangle = 0 \qquad a_{\mathbf{k}\alpha}\left|\Phi_0\right\rangle = 0$$

So many terms do not give a contribution:

$$iG^0_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \begin{cases} \frac{1}{V}\sum_{\substack{|\mathbf{k}|>k_F \\ |\mathbf{k}'|>k_F}} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_kt)}e^{-i(\mathbf{k}'\cdot\mathbf{x}'-\omega'_kt')}\eta_{\alpha}\eta^{\dagger}_{\beta}\langle\Phi_0|a_{\mathbf{k}\alpha}a^{\dagger}_{\mathbf{k}\beta}|\Phi_0\rangle & (t>t') \\ -\frac{1}{V}\sum_{\substack{|\mathbf{k}|\leq k_F \\ |\mathbf{k}'|\leq k_F}} e^{i(\mathbf{k}'\cdot\mathbf{x}'+\omega'_kt')}e^{-i(\mathbf{k}\cdot\mathbf{x}+\omega_kt)}\eta^{\dagger}_{\beta}\eta_{\alpha}\langle\Phi_0|b_{\mathbf{k}'\beta}b^{\dagger}_{\mathbf{k}\alpha}|\Phi_0\rangle & (t<t') \end{cases}$$

Changing variables, we get

$$iG^0_{\alpha\beta}(\mathbf{x}t,\mathbf{x}'t') = \frac{1}{V}\delta_{\alpha\beta}\sum_{\mathbf{k}}e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')}e^{-i\omega_{\mathbf{k}}(t-t')}\left[\theta(t-t')\theta(k-k_F)-\theta(t'-t)\theta(k_F-k)\right]$$

Letting the volume go to infinity, the summation becomes an integral:

$$iG^0_{\alpha\beta}(\mathbf{x}t,\mathbf{x}'t') = \frac{1}{(2\pi)^3}\delta_{\alpha\beta}\int \mathrm{d}\mathbf{k} e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')}e^{-i\omega_{\mathbf{k}}(t-t')}\left[\theta(t-t')\theta(k-k_F)-\theta(t'-t)\theta(k_F-k)\right]$$

using the integral representation for the step function

$$\theta(t-t') = -\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi i}\frac{e^{-i\omega(t-t')}}{\omega+i\eta}$$

the Green's function becomes

$$G^0_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t) = \frac{1}{(2\pi)^4}\int \mathrm{d}\mathbf{k}\int_{-\infty}^{+\infty}\mathrm{d}\omega e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}e^{-i\omega(t-t')}\delta_{\alpha\beta}\left[\frac{\theta(k-k_F)}{\omega-\omega_k+i\eta}+\frac{\theta(k_F-k)}{\omega-\omega_k-i\eta}\right].$$

Comparing with

$$G_{\alpha\beta}(\mathbf{x}-\mathbf{x}',t-t') = \frac{1}{(2\pi)^4}\int \mathrm{d}\mathbf{k}\int_{-\infty}^{+\infty}\mathrm{d}\omega e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}e^{-i\omega(t-t')}G_{\alpha\beta}(\mathbf{k},\omega)$$

We find the Green's function for non interaction fermions in momentum space:

$$G^0_{\alpha\beta}(\mathbf{k},\omega) = \delta_{\alpha\beta}\left[\frac{\theta(k-k_F)}{\omega-\omega_k+i\eta}+\frac{\theta(k_F-k)}{\omega-\omega_k-i\eta}\right]$$

Where the first term refers to particles and the second to holes. $\omega_k = \frac{\epsilon_k}{\hbar} = \frac{\hbar k^2}{2m}$

3.3 The Lehmann representation

We now go back to the general interacting case. Lehmann representation is modification of Green functions in order to emphasize certain aspects, in particular elementary excitations. We do the calculations for fermions, because the possibility of Bose condensation at $T = 0$ produces additional complications.
Again we start from the definition of the Green's function

$$iG_{\alpha\beta}(\mathbf{x}t,\mathbf{x}'t') = \langle]|T\left[\hat{\psi}_{H\alpha}(\mathbf{x}t)\hat{\psi}^\dagger_{H\beta}(\mathbf{x}'t')\right|\Psi_0\rangle\Psi_0.$$

We insert a complete set of Heisenberg states $\{|\Psi_n\rangle\}$ (with $\sum_n |\Psi_n\rangle\langle\Psi_n| = \mathbb{1}$ and $\hat{H}|\Psi_n\rangle = E_n|\Psi_n\rangle$) in this definition (expanding the time-ordered product for fermionic fields):

$$iG_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \sum_n\Big[\theta(t-t')\langle\Psi_0|\hat{\psi}_{H\alpha}(\mathbf{x}t)|\Psi_n\rangle\langle\Psi_n|\hat{\psi}^\dagger_{H\beta}(\mathbf{x}'t')|\Psi_0\rangle - \theta(t'-t)\langle\Psi_0|\hat{\psi}^\dagger_{H\beta}(\mathbf{x}'t')|\Psi_n\rangle\langle\Psi_n|\hat{\psi}_{H\alpha}(\mathbf{x}t)|\Psi_0\rangle\Big]$$

The time-dependence of the matrix elements can be made explicit:

$$iG_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \sum_n\Big[\theta(t-t')e^{-i(E_n-E)(t-t')/\hbar}\langle\Psi_0|\hat{\psi}_\alpha(\mathbf{x})|\Psi_n\rangle\langle\Psi_n|v|\Psi_0\rangle - \theta(t'-t)e^{i(E_n-E)(t-t')/\hbar}\langle\Psi_0|\hat{\psi}^\dagger_\beta(\mathbf{x}')|\Psi_n\rangle\langle\Psi_n|\hat{\psi}_\alpha(\mathbf{x})|\Psi_0\rangle\Big]$$

The relevant states $|\Psi_n\rangle$ contain $N \pm 1$ particles if the state Ψ_0 contains N particles.

First assume that $\hat{H}$ does not depend on time, the with $\tau = t - t'$ we can write

$$G_{\alpha\beta}(\mathbf{x}, \mathbf{x}', \tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}\omega e^{-i\omega\tau} \tilde{G}_{\alpha\beta}(\mathbf{x}, \mathbf{x}', \omega)$$

where

$$\tilde{G}_{\alpha\beta}(\mathbf{x}, \mathbf{x}', \omega) = \sum_n \left[\frac{\langle \Psi_0 | \hat{\psi}_\alpha(\mathbf{x}) | \Psi_n \rangle \langle \Psi_n | \hat{\psi}^\dagger_\beta(\mathbf{x}') | \Psi_0 \rangle}{\omega - (E_n - E_0)/\hbar + i\eta} + \frac{\langle \Psi_0 | \hat{\psi}^\dagger_\beta(\mathbf{x}') | \Psi_n \rangle \langle \Psi_n | \hat{\psi}_\alpha(\mathbf{x}) | \Psi_0 \rangle}{\omega + (E_n - E_0)/\hbar - i\eta} \right]$$

This is a meromorphic function with poles at

$$\omega_p = \pm(E_n - E_0)/\hbar \mp i\eta$$

Now we take a look at the first denominator:

$$\begin{aligned} \omega &= \frac{(E_n - E_0)}{\hbar} - i\eta = \frac{E_n(N+1)}{\hbar} - \frac{E_0(N)}{\hbar} - i\eta \\ &= \frac{E_n(N+1) - E_0(N+1)}{\hbar} + \frac{E_0(N+1) - E_0(N)}{\hbar} - i\eta \\ &= \frac{\epsilon_n(N+1)}{\hbar} + \frac{\mu}{\hbar} - i\eta \end{aligned}$$

Where μ is the chemical potential. It measures the gain in energy if we add one more particles to our system. $\mu(N+1) = \mu(N) + O(1/N)$. The excitation energy ϵ_n is non negative.
The second denominator gives a similar result, yielding

$$\tilde{G}_{\alpha\beta}(\mathbf{x}, \mathbf{x}', \omega) = \sum_n \left[\frac{\langle \Psi_0 | \hat{\psi}_\alpha(\mathbf{x}) | \Psi_n \rangle \langle \Psi_n | \hat{\psi}^\dagger_\beta(\mathbf{x}') | \Psi_0 \rangle}{\omega - \epsilon_n(N+1)/\hbar - \mu/\hbar + i\eta} + \frac{\langle \Psi_0 | \hat{\psi}^\dagger_\beta(\mathbf{x}') | \Psi_n \rangle \langle \Psi_n | \hat{\psi}_\alpha(\mathbf{x}) | \Psi_0 \rangle}{\omega + \epsilon_n(N-1)/\hbar - \mu/\hbar - i\eta} \right]$$

At this point we assume translational invariance, so the momentum operator commutes with $\hat{H}$. It is natural to use the plane wave basis for such a system. The momentum operator for such a system is given by

$$\hat{\mathbf{P}} \equiv \sum_\alpha \int \mathrm{d}\mathbf{x} \hat{\psi}^\dagger_\alpha(\mathbf{x})(-i\hbar \boldsymbol{\nabla}) \hat{\psi}_\alpha(\mathbf{x}) = \sum_{\mathbf{k}\lambda} \hbar\mathbf{k} c^\dagger_{\mathbf{k}\lambda} c_{\mathbf{k}\lambda}$$

The expression

$$\left[\hat{\psi}_\alpha(\mathbf{x}), \hat{\mathbf{p}} \right] = -i\hbar \nabla \psi_\alpha(\mathbf{x})$$

can be written in integral form

$$\hat{\psi}_\alpha(\mathbf{x}) = e^{-i\hat{\mathbf{p}}\cdot\mathbf{x}/\hbar} \hat{\psi}_\alpha(0) e^{i\hat{\mathbf{p}}/\hbar}$$

So for Heisenberg fields we can write

$$\hat{\psi}_{H\alpha}(\mathbf{x}, t) = e^{i(\hat{H}t - \hat{\mathbf{p}}\cdot\mathbf{x})/\hbar} \hat{\psi}_\alpha(0) e^{-i(\hat{H}t - \hat{\mathbf{p}}\cdot\mathbf{x})/\hbar}$$

Lehmann representation

$$G_{\alpha\beta}(\mathbf{k}, \omega) = \hbar V \sum_n \left[\frac{\langle \Psi_0 | \hat{\psi}_\alpha(0) | n\mathbf{k} \rangle \langle n\mathbf{k} | \hat{\psi}^\dagger_\beta(0) | \Psi_0 \rangle}{\hbar\omega - (\mu + \epsilon_n^{N+1}(\mathbf{k})) + i\eta} + \frac{\langle \Psi_0 | \hat{\psi}^\dagger_\beta(0) | n, -\mathbf{k} \rangle \langle n, -\mathbf{k} | \hat{\psi}_\alpha(0) | \Psi_0 \rangle}{\hbar\omega - (\mu - \epsilon_n^{N-1}(-\mathbf{k})) - i\eta} \right]$$

System with spin 1/2 particles with coulomb potential is invariant under rotations and reflections

$$\left[\hat{H}, \mathcal{P}\right] = 0$$

Basis set:

$$\{\mathbb{1}, \boldsymbol{\sigma}\} \qquad \sigma_x = ...$$

So we can decompose

$$G_{\alpha\beta}(\mathbf{k}, \omega) = G(\mathbf{k}, \omega)\mathbb{1} + b(\boldsymbol{\sigma} \cdot \mathbf{k})$$

This should depend only on the modulus of k (no preferred direction). Because $b(\boldsymbol{\sigma} \cdot \mathbf{k})$ is a pseudo-scalar, if there is reflection symmetry $b = 0$. So we get

$$G_{\alpha\beta}(\mathbf{k}, \omega) = \delta_{\alpha\beta} G(|\mathbf{k}|, \omega)$$

$$\sum_\alpha G_{\alpha\alpha} = \operatorname{Tr} G = (2s+1)G \qquad \to \qquad G = \frac{G_{\alpha\alpha}}{2s+1}$$

The only non-vanishing component is when $\beta = \alpha$

$$G_{\alpha\alpha}(\mathbf{k}, \omega) = \hbar V \sum_n \left[\frac{\langle \psi_0 | \hat{\psi}_\alpha(0) | n\mathbf{k} \rangle \langle n\mathbf{k} | \hat{\psi}^\dagger_\alpha(0) | \psi_0 \rangle}{\hbar\omega - (\mu + \epsilon_n^{N+1}(\mathbf{k})) + i\eta} + \frac{\langle \psi_0 | \hat{\psi}^\dagger_\alpha(0) | n, -\mathbf{k} \rangle \langle n, -\mathbf{k} | \hat{\psi}_\alpha(0) | \psi_0 \rangle}{\hbar\omega - (\mu - \epsilon_n^{N-1}(-\mathbf{k})) - i\eta} \right]$$

So

$$G(\mathbf{k}, \omega) = \frac{\hbar V}{2s+1} \sum_n \left[\frac{\left| \langle \psi_0 | \hat{\psi}_\alpha(0) | n\mathbf{k} \rangle \right|^2}{\hbar\omega - (\mu + \epsilon_n(\mathbf{k})) + i\eta} + \frac{\left| \langle \psi_0 | \hat{\psi}^\dagger_\alpha(0) | n, -\mathbf{k} \rangle \right|^2}{\hbar\omega - (\mu - \epsilon_n(-\mathbf{k})) - i\eta} \right]$$

We take $V \to \infty$
We write $\mathrm{d}n$ the number of states between ϵ and $\epsilon + \mathrm{d}\epsilon$. $(\epsilon < \epsilon_n(\mathbf{k}) < \epsilon + \mathrm{d}\epsilon)$

$$\mathrm{d}n = \left(\frac{\mathrm{d}n(\epsilon)}{\mathrm{d}\epsilon} \right) \mathrm{d}\epsilon$$

$$A(\mathbf{k}, \omega) \equiv \lim_{V\to\infty} \frac{\hbar V}{2s+1} \left| \langle \psi_0 | \hat{\psi}_\alpha(0) | n\mathbf{k} \rangle \right|^2 \frac{\mathrm{d}n}{\mathrm{d}\epsilon} \geq 0$$

$$B(\mathbf{k}, \omega) \equiv \lim_{V\to\infty} \frac{\hbar V}{2s+1} \left| \langle \psi_0 | \hat{\psi}^\dagger_\alpha(0) | n, -\mathbf{k} \rangle \right|^2 \frac{\mathrm{d}n}{\mathrm{d}\epsilon} \geq 0$$

$(\omega = \frac{\epsilon}{\hbar})$

$$\sum_n \to \int \mathrm{d}n = \int \mathrm{d}\epsilon \frac{\mathrm{d}n}{\mathrm{d}\epsilon} = \hbar \int \mathrm{d}\omega \frac{\mathrm{d}n}{\mathrm{d}\epsilon}$$

So

$$\frac{V}{2s+1} \sum_n ?????(applyingtheabove)$$

Spectral representation:

$$G(\mathbf{k}, \omega) = \int_0^\infty \mathrm{d}\omega' \left[\frac{A(\mathbf{k}, \omega')}{\omega - \hbar^{-1}\mu - \omega' + i\eta} + \frac{B(\mathbf{k}, \omega')}{\omega - \hbar^{-1}\mu + \omega' - i\eta} \right]$$

$$\int_0^\infty A(\mathbf{k}, \omega)\, \mathrm{d}\omega + \int_0^\infty B(\mathbf{k}, \omega)\, \mathrm{d}\omega = 1$$

ω not analytic in either top or bottom half complex plane.
Retarded or advanced GF

$$iG^R(\mathbf{x}t, \mathbf{x}', t') = \langle\psi_0| \left\{\hat{\psi}_{H\alpha}(\mathbf{x}, t), \hat{\psi}^\dagger_{H\beta}(\mathbf{x}'t')\right\} |\psi_0\rangle\theta(t - t')$$

$$iG^A(...) = ???$$

After similar procedure for Green function

$$G_{\alpha\beta}(\mathbf{k}, \omega)^{R,A} = \hbar V \sum_n \left[\frac{\langle\psi_0|\hat{\psi}_\alpha(0)|n\mathbf{k}\rangle\langle n\mathbf{k}|v|\psi_0\rangle}{\hbar\omega - (\mu + \epsilon_n(\mathbf{k})) \pm i\eta} + \frac{\langle\psi_0|\hat{\psi}^\dagger_\beta(0)|n, -\mathbf{k}\rangle\langle n, -\mathbf{k}|\hat{\psi}_\alpha(0)|\psi_0\rangle}{\hbar\omega - (\mu - \epsilon_n(-\mathbf{k})) \pm i\eta}\right]$$

Where plus and minus refer to advanced and retarded.
This expression is very similar, but with a crucial difference:

- G^R only has poles below the real axis and thus is analytical for $\mathfrak{Im}\,\omega > 0$.
- G^A only has poles above the real axis and thus is analytical for $\mathfrak{Im}\,\omega < 0$.

If ω is real and $\omega > \frac{\mu}{\hbar}$, then $G^R_{\alpha\beta}(\mathbf{k}, \omega) = G_{\alpha\beta}(\mathbf{k}, \omega)$.
If ω is real and $\omega < \frac{\mu}{\hbar}$, then $G^A_{\alpha\beta}(\mathbf{k}, \omega) = G_{\alpha\beta}(\mathbf{k}, \omega)$.
We can also obtain a spectral representation for these 2 functions

$$G_{\alpha\beta}(\mathbf{k}, \omega)^{R,A} = \int_0^\infty \mathrm{d}\omega' \left[\frac{A(\mathbf{k}, \omega')}{\omega - \hbar^{-1}\mu - \omega' \pm i\eta} + \frac{B(\mathbf{k}, \omega')}{\omega - \hbar^{-1}\mu + \omega' \pm i\eta}\right]$$

We write

$$\frac{1}{\omega \pm i\eta} = \sigma\frac{1}{\omega} \mp i\pi\delta(\omega)$$

where σ is the principle part. (Also missed formal definition)
So

$$G_{\alpha\beta}(\mathbf{k}, \omega)^{R,A} = \int_0^\infty \mathrm{d}\omega' \frac{A(\mathbf{k}, \omega')}{\omega - \hbar^{-1}\mu - \omega'} \mp i\pi A(...)???$$

Now we consider the imaginary part of this expression

$$\mathfrak{Im}\, G^{R,A}(\mathbf{k}, \omega') = \mp\pi A(\mathbf{k}, \omega' - \hbar^{-1}\mu) \mp B(\mathbf{k}, -\omega' + \hbar^{-1}\mu) \mp B(\mathbf{k}, -\omega' + \hbar^{-1}\mu)$$

??????????????????
Dispersion relations or Kramer - Kronig relations

$$\mathfrak{Re}\, G^{R,A}(\mathbf{k}, \omega) = \mp\mathcal{P}\int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega'}{\pi} \frac{\mathfrak{Im}\, G^{R,A}(\mathbf{k}, \omega')}{\omega - \omega'}$$

??? In general (for any system)

$$G_{\alpha\beta}(\mathbf{k}, \omega) \sim \frac{1}{\omega} \qquad \omega \to \infty$$

"Adiabatic switching on"

$$\hat{H} = \hat{H}_0 + \hat{V} = \hat{H}_0 + \hat{H}_1$$
$$\hat{H}(t) = \hat{H}_0 + e^{-\epsilon(t)}\hat{H}_1$$

$$\begin{cases} \hat{H}(0) = \hat{H}_0 + \hat{H}_0 \\ \hat{H}(t = ^{+\infty}_{-\infty}) \to \hat{H}_0 \end{cases}$$

Missed previous lecture
Lecture 09/04
20 minutes late.

$$\begin{aligned} iG_{\alpha\beta}(\mathbf{x},t,\mathbf{x}',t') &= \langle]|T\left[\hat{\psi}_{H\alpha}(\mathbf{x},t)\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}',t')\right| \psi_0\rangle\psi_0 \\ &= \langle\psi_0|\hat{\psi}_{H\alpha}(\mathbf{x},t)\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}',t')|\psi_0\rangle\theta(t-t') - \langle\psi_0|\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}',t')\hat{\psi}_{H\alpha}(\mathbf{x},t)|\psi_0\rangle\theta(t'-t) \end{aligned}$$

Green function: the physical interpretation

$$\langle\psi_0|\hat{\psi}(\mathbf{x},t)\hat{\psi}^{\dagger}(\mathbf{x}',t')|\psi_0\rangle \qquad t > t'$$

We have system. At time t' we create particle at location $\mathbf{x}'$. It evolves and then gets destroyed at time t and place $\mathbf{x}$.

$$\langle\psi_0|\hat{\psi}^{\dagger}(\mathbf{x}',t')\hat{\psi}(\mathbf{x},t)|\psi_0\rangle \qquad t' > t$$

We have system. Particle gets destroyed at time t and place $\mathbf{x}$. Hole evolves and gets filled at time t' and $\mathbf{x}'$.
This justifies the name propagator. (Describes additional particle or pole).

$$\hat{U}(t,t_0) = e^{i\hat{H}_0 t/\hbar}e^{-i\hat{H}(t-t_0)/\hbar}e^{-i\hat{H}_0 t_0/\hbar}$$

$$\hat{\psi}_I(t) = e^{i\hat{H}_0 t/\hbar}\hat{\psi}_S e^{-i\hat{H}_0 t/\hbar}$$

Go from interaction picture to Heisenberg picture

$$\hat{U}(0,t)\hat{\psi}_I(t)\hat{U}(t,0) = e^{i\hat{H}t/\hbar}\hat{\psi}_S e^{-i\hat{H}t/\hbar} = \hat{\psi}_H(t)$$

For $t > t'$:

$$\begin{aligned} iG_{\alpha\beta}(\mathbf{x},t,\mathbf{x}',t') &= \langle\psi_0|\hat{\psi}_{H\alpha}(\mathbf{x},t)\hat{\psi}^{\dagger}_{H\beta}(\mathbf{x}',t')|\psi_0\rangle \\ &= \langle\phi_0|\hat{U}(\infty,0)\hat{U}(0,t)\hat{\psi}_{I_\alpha}(\mathbf{x},t)\hat{U}(t,0)\hat{U}(0,t')\hat{\psi}_{I_\beta}(\mathbf{x}',t')\hat{U}(t',0)\hat{U}(0,-\infty)|\phi_0\rangle \\ &= \langle\phi_0|\hat{U}(\infty,t)\hat{\psi}_{I_\alpha}(\mathbf{x},t)\hat{U}(t,t')\hat{\psi}_{I_\beta}(\mathbf{x}',t')\hat{U}(t',-\infty)|\phi_0\rangle \end{aligned}$$

So we have evolution from $-\infty$ to t', t' to t and t to ∞.

$$\frac{\langle\psi_0|\hat{O}_H(t)|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} = \frac{\langle\phi_0|\hat{U}_\epsilon(\infty,t)\hat{O}_I(t)\hat{U}_\epsilon(t,-\infty)|\phi_0\rangle}{\left\langle\hat{S}\phi_0\right\rangle}$$

With

$$\hat{S} \equiv \hat{U}_\epsilon(\infty,0)\hat{U}_\epsilon(0,-\infty) = \hat{U}_\epsilon(\infty,-\infty)$$

Interaction picture.
From now on we adopt the conventions

$$\hat{\psi}_I \;\to\; \hat{\psi} \qquad x \equiv (\mathbf{x},t_x) = (\mathbf{x},x_0)$$

$$U(x_1, x_2) = V(\mathbf{x}_1, \mathbf{x}_2)\delta(t_1 - t_2)$$

(we assume the interaction is instantaneous, in a fully relativistic approach we may not assume this).

$$iG_{\alpha\beta}(x,y) = \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar}\right)^n \frac{1}{n!} \int_{-\infty}^{+\infty} \mathrm{d}t_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}t_n \frac{\langle]|T\left[\hat{H}_1(t_1)\ldots\hat{H}_n(t_n)\hat{\psi}_\alpha(x)\hat{\psi}^\dagger_\beta(y)\right|\phi_0\rangle\phi_0}{\langle\phi_0|\hat{S}|\phi_0\rangle}$$

We now look at numerator with $n = 0 \ + \ n = 1$.

$$i\tilde{G}_{\alpha\beta}(x,y) = \langle]|T\left[\hat{\psi}_\alpha(x)\hat{\psi}^\dagger_\beta(y)\right|\phi_0\rangle\phi_0 + \left(-\frac{i}{\hbar}\right)\sum_{\substack{\lambda\lambda'\\ \mu\mu'}} \frac{1}{2}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_1' U(x_1,x_1')\langle]|T\left[\hat{\psi}^\dagger_\lambda(x_1)\hat{\psi}^\dagger_\mu(x_1')\hat{\psi}_{\mu'}(x_1')\hat{\psi}_{\lambda'}(x_1)\hat{\psi}_\alpha(x)\hat{\psi}^\dagger_\beta(y\right.$$

Goal:

$$\langle]|T\left[\hat{\psi}^\dagger\ldots\hat{\psi}\hat{\psi}_\alpha(x)\hat{\psi}^\dagger_\beta(y)\right|\phi_0\rangle\phi_0$$

(non-interacting G.S)
Not easy process, but theorem simplifies it.
Particle-hole formalism

$$\hat{\psi}_S(\mathbf{x}) = \sum_{\substack{\mathbf{k}\lambda\\ |\mathbf{k}|>k_F}} \psi_{\mathbf{k}\lambda}(\mathbf{x})a_{\mathbf{k}\lambda} + \sum_{\substack{\mathbf{k}\lambda\\ |\mathbf{k}|\geq k_F}} \psi_{\mathbf{k}\lambda}(\mathbf{x})^\dagger_{-\mathbf{k}\lambda}$$

$$\hat{\psi}_I(\mathbf{x},t) = e^{i\hat{H}_0 t/\hbar}\hat{\psi}_S(\mathbf{x})e^{-i\hat{H}_0 t/\hbar}$$

???? (something something Fermi sphere)

$$\hat{\psi}(x) = \hat{\psi}^{(+)}(x) + \hat{\psi}^{(-)}(x)$$
$$\hat{\psi}^\dagger(x) = \hat{\psi}^{(+)\dagger}(x) + \hat{\psi}^{(-)\dagger}(x)$$

Destruction parts:

$$\hat{\psi}^{(+)}(x)\,|\phi_0\rangle = 0 \qquad \hat{\psi}^{(-)\dagger}(x)\,|\phi_0\rangle = 0$$

Normal order operator: put the destruction operator on the right.

$$N\left[\hat{\psi}^{(+)}\cdot\hat{\psi}^{(-)}\right] = -\hat{\psi}^{(-)}\cdot\hat{\psi}^{(+)}$$

(minus for Fermions)
Missed some lectures
Lecture 17/04 15 mins late
Feynman rules in momentum space: internal vertex. Just see notes

Part XV

Nuclear physics

becquerel uranium in sunshine, so fosforessence? No, rainy day.
ionic atomic radius

Chapter 1

Describing the nucleus

1.1 Nuclear radius

1.2 Mass and abundance of nuclides

1.3 Nuclear binding energy

1.4 Nuclear spin and moments

angular momentum, parity, electromagnetic moments, spin, hyperfine structure

Chapter 2

Nuclear models

Chapter 3

Radioactive decay

3.1 General features

3.2 Detection

3.3 Alpha decay

3.4 Beta decay

3.5 Gamma decay

3.5.1 Interactions with matter

3.5.1.1 Major processes

Compton scatteringing

Photoelectric effect Klein-nishina

Pair production

3.5.1.2 Minor processes

Rayleigh scattering

Thomson scattering

Resonance Raman scattering

EXAFS

Chapter 4

Nuclear reactions

Part XVI

Solid state physics

Chapter 1

Introduction

Condensed matter physics is the branch of physics that studies systems with a large number of constituents and where the constituents are close enough together (i.e. condensed enough) that the forces between them are strong.
Solid state physics studies the properties of rigid condensed matter, aka solids. In particular we will be discussing crystalline solids in this chapter. Amorphous solids and glasses require a very different treatment and will be discussed in the chapter on soft matter physics.

Chapter 2

Crystal structure

Crystals are three-dimensional periodic arrays of atoms. This may include atoms of different types. In this section we will develop the language and mathematics necessary to describe these periodic arrays. This is the language of crystallography.
The essential characteristic of crystals is that they are periodic. Hopefully it is somewhat intuitively obvious what that means, but we would also like to also express it in a more rigorous, mathematical way.

2.1 Periodicity and crystal structure

Periodicity means that if we add a fixed quantity an integer number of times, the resulting operation is a symmetry. In other words there is a vector $\mathbf{a}$ such that the transformation $\mathbf{r} \to \mathbf{r'}$ with

$$\mathbf{r'} = \mathbf{r} + n_1\mathbf{a}$$

is a symmetry for any integer n_1. This is obviously a form of translational symmetry. There may also be other symmetries that have the same form. Let us assume for example that there is a different vector $\mathbf{b}$ that has the same property as the vector $\mathbf{a}$. Applying the two symmetry operations will still give a symmetry operation

$$\mathbf{r} \to \mathbf{r'} = \mathbf{r} + n_1\mathbf{a} + n_2\mathbf{b}.$$

Conversely if we set n_1 or n_2 to 0, we get the original symmetry groups.
Now it may be that $\mathbf{b}$ is a multiple of $\mathbf{a}$, i.e. $\mathbf{b} = \alpha\mathbf{a}$. First we consider the case where α is a rational number. This means we can write α as $\frac{m}{n}$ where m and n are integers that have no common divisors other than one. In this case translating over an integer multiple of $\frac{\mathbf{a}}{n}$ is also a symmetry (this fact is not entirely trivial and follows from Bézout's identity). In other words, there is a vector $\mathbf{a'}$ such that

$$\begin{cases} \mathbf{a} = n\mathbf{a'} \\ \mathbf{b} = m\mathbf{a'} \end{cases}$$

and all translations of the form

$$\mathbf{r'} = \mathbf{r} + n_1\mathbf{a'}$$

are symmetries.
If α is *not* a rational number, then we can find arbitrarily small translations that are still symmetries (this follows from the fact that α can be approximated by $\frac{m}{n}$ with n arbitrarily

large while the greatest common divisor of m and n is still one). This would mean that there is continuous translational symmetry, i.e. the transformation

$$\mathbf{r} \to \mathbf{r'} = \mathbf{r} + \lambda \mathbf{a}$$

is a symmetry for any real λ. This should not fit with your idea of periodicity and thus we require α to be rational.
So we can assume $\mathbf{a}$ and $\mathbf{b}$ are linearly independent (otherwise we can just use $\mathbf{a'}$ instead). We can now compose this with a third symmetry. Following exactly the same reasoning we get

$$\mathbf{r} \to \mathbf{r'} = \mathbf{r} + n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}$$

with $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ linearly independent. In three dimensional space we can only ever find three linearly independent vectors, so we need look no further. Any translational symmetry (that is not continuous) can be written in this form. These operations are called lattice translation operations. The vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are called the fundamental translation vectors; they define a lattice which is a regular periodic arrangement of points in space. The magnitude of the translation vectors, a, b and c are called the lattice constants. The lattice and translation vectors are said to be primitive if every translational symmetry can be obtained by choosing the correct n_1, n_2 and n_3. We have shown that it is alway possible to find primitive translation vectors.
The full crystal structure can now be described by specifying a unit, called a basis that is repeated at every lattice point, which is what gives us periodicity. In summary

The crystal structure is specified by a lattice and a basis.

TODO image
It may be remarked at this point that a crystal structure may be described by different lattices and bases. In particular if the lattice is not primitive, the basis must necessarily be larger. The basis associated with a primitive lattice may be called a primitive basis. No basis contains fewer atoms than a primitive basis.

2.2 Lattice cells

We can also divide the crystal into functionally identical regions of space called cells. A unit cell will fill all space under the action of suitable lattice translation operations. A primitive cell is a unit cell of minimum volume. The basis contained in a primitive cell is a primitive basis. There are two important types of primitive cells:

1. The parallelepiped defined by primitive axes $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ is a primitive cell. It has a volume

$$V_c = |\mathbf{a} \times \mathbf{b} \cdot \mathbf{c}|.$$

2. The Wigner-Seitz primitive cell is defined by the procedure shown in figure TODO.

Positions of points in a unit cell are specified in terms of atomic coordinates u, v, w. The position vector is then given by

$$\mathbf{r} = u\mathbf{a} + v\mathbf{b} + w\mathbf{c}.$$

2.3 Fundamental types of lattices

Lattices may also exhibit other types of symmetry such as rotational and reflectional symmetry, not just translational. A lattice point group is the collection of symmetry operations which, when applied about a lattice point, leave the lattice invariant.
We can sort the lattices into types based on their point group symmetries. We each fundamental type of lattice is called a Bravais lattice. It is important to note that these are the symmetries of the *lattice*, not necessarily of the crystal structure. The crystal structure may have fewer symmetries if the basis does not have the symmetries the lattice has.

2.3.1 Bravais lattices in two dimensions

In this case there are two fundamental translation vectors, $\mathbf{a}$ and $\mathbf{b}$. Any of the infinite possible lattices are determined by the lengths a and b of the vectors along with the angle φ between them.
TODO

2.3.2 Bravais lattices in three dimensions

TODO, conventional unit cell.

2.4 Position and orientation of planes in crystals: Miller indies

The position and orientation of planes in crystals are typically specified in terms of Miller indices. The Miller indices (hkl) denote the family of planes that intercept the three points $\mathbf{a}/h, \mathbf{b}/k, \mathbf{c}/l$, or some integer multiple thereof. If one of the indices is zero, it means that the planes are parallel to that axis (the intercept is "at infinity"). A negative index is indicated by putting a bar above the index.
TODO: indices integers?? Meaning of (200). GCD = 1. Through lattice points.
The set of planes that are equivalent to (hkl) due to the symmetry of the bravais lattice is denote $\{hkl\}$. Again we must be careful to specify whether we are working with the lattice or the crystal structure. In the crystal the family $\{hkl\}$ will be smaller.
Directions can be denoted $[hkl]$. This means the direction of the vector

$$h\mathbf{a} + k\mathbf{b} + l\mathbf{c}.$$

2.5 Examples of simple crystal structures

TODO

Chapter 3

Crystal diffraction and the reciprocal lattice

Crystal diffraction is one of the main ways to investigate crystals. When X-ray diffraction was introduced in 1912 it decisively proved that crystals were comprised of periodically repeating units.

3.1 Bragg law

When an atom is exposed to electromagnetic radiation, the atomic electrons may scatter part or all of the radiation elastically, at the frequency of the incident radiation. At optical wavelengths the superposition of all the waves scattered elastically by the atoms in the crystal result in ordinary optical refraction as you would expect. If the wavelength of the incident radiation is comparable or smaller than the lattice constants, we might find one or more diffracted beams in directions quite different from the incident direction.
Lawrence Bragg presented an explanation of this phenomenon based on a very simple model. The model is so simplified in fact that it is only really credible because its result also follows from Laue's derivation. The model is as follows

1. Suppose the incident waves are reflected from parallel planes of atoms in the crystal. Any set of parallel planes will do, provided each plane passes through at least three non-collinear lattice points.

2. Suppose the reflection is specular (i.e. mirror-like). So we are briefly assuming geometrical optics for the reflection, but we must switch back to viewing the beam as wave-like immediately after.

3. Each plane only reflects a small fraction of the incoming beam.

4. The diffracted beams are only found when the reflections from parallel planes interfere constructively. If the planes are a distance d apart, the difference in distance traveled by rays reflected off adjacent planes is $2d\sin\theta$. See figure TODO. In order to have constructive interference, this distance must be an integer multiple of the wavelength.

Thus we get the **Bragg law**

$$\boxed{2d\sin\theta = n\lambda}$$

where λ is the wavelength of the incident beam.
Bragg reflection can only occur for wavelengths $\lambda \leq 2d$, which is why it is not observed with visible light.

3.2 Experimental diffraction methods

The Bragg law gives the locations of the peaks of diffraction in function of the wavelength λ and angle θ of the incoming beam. In order to find those peaks experimentally we must scan a continuous range of either λ or θ, usually the latter.
The incident beam may be composed of x-rays, neutrons or, less commonly, electrons.

3.2.1 Laue method

TODO

3.2.2 Rotating crystal method

TODO

3.2.3 Powder method

TODO

3.3 Laue derivation

TODO: make the explanation less terrible
We now give a more careful treatment of crystal diffraction due to Laue. We now take the incident beam to be a plane wave that is not greatly disturbed by the crystal, neither by the refractive index of the crystal nor by the loss of energy through scattering. We assume the crystal structure is made up of identical point scattering centers at every lattice point. Each atom scatters a part of the wave elastically. We then sum all of these contributions to find a condition for when they interfere constructively.
The incident beam is a plane wave and may be expressed as

$$F(\mathbf{x}) = F_0 \exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)]$$

with wavevector $\mathbf{k}$ and angular frequency ω. We assume the origin is within the crystal. We are calculating the diffraction condition in a stable state, so we may do the calculation at any time; for ease we consider the instant of time $t = 0$. Say there is an atom in the crystal at a point $\boldsymbol{\rho}$. Then the incident beam has an amplitude

$$F(\boldsymbol{\rho}) = F_0 \exp(i\mathbf{k} \cdot \boldsymbol{\rho})$$

at location $\boldsymbol{\rho}$ and time $t = 0$.
Now what does this scattered wave look like at a location $\mathbf{R}$ outside the crystal? The total phase factor is given by

$$\exp(i\mathbf{k} \cdot \boldsymbol{\rho}) \exp(ikr) = \exp(i\mathbf{k} \cdot \boldsymbol{\rho} + ikr)$$

where $\mathbf{r}$ is the displacement from $\boldsymbol{\rho}$ to $\mathbf{R}$ (so $\mathbf{R} = \boldsymbol{\rho} + \mathbf{r}$). Here we only need the magnitude because the scattered wave expands spherically and thus $\mathbf{k}$ and $\mathbf{r}$ point in the same direction (see figure TODO).
The amplitude of the wave is may be taken to be proportional to $\frac{1}{r}$ and to the electron concentration in the volume element. Due to our assumptions about the crystal structure, we can assume the electron concentration to be the same for every atom.
We can write

$$r^2 = (\mathbf{R} - \boldsymbol{\rho})^2 = R^2 + \rho^2 - 2\rho R\cos(\boldsymbol{\rho}, \mathbf{R})$$

where $\cos(\boldsymbol{\rho}, \mathbf{R})$ is the cosine of the angle between the vectors $\boldsymbol{\rho}$ and $\mathbf{R}$. If R is much larger than the dimensions of the crystal, $\rho/R \ll 1$ and we can approximate r as

$$r \approx R[1 - (2\rho/R)\cos(\boldsymbol{\rho}, \mathbf{R})]^{1/2} \approx R - \rho\cos(\boldsymbol{\rho}, \mathbf{R})$$

Thus the phase factor becomes

$$\exp[i\mathbf{k}\cdot\boldsymbol{\rho} + ikR - ik\rho\cos(\boldsymbol{\rho}, \mathbf{R})].$$

The factor $\exp(ikR)$ is constant and thus does not change for waves scattered from different $\boldsymbol{\rho}$. The only part of the amplitude that depends on $\boldsymbol{\rho}$ is the factor $1/r$. For R large enough we may take $1/r \approx 1/R$. Thus the amplitude is independent of $\boldsymbol{\rho}$.
The scattered wave has a wave vector $\mathbf{k'}$ with the same direction as $\mathbf{r}$, which is approximately the direction of $\mathbf{R}$ and because the scattering is elastic, $k' = k$. The phase factor can now be written

$$\exp[i\mathbf{k}\cdot\boldsymbol{\rho} - ik\rho\cos(\boldsymbol{\rho}, \mathbf{R})] = \exp[i\boldsymbol{\rho}\cdot(\mathbf{k} - \mathbf{k'})] = \exp[-i\boldsymbol{\rho}\cdot\Delta\mathbf{k}]$$

where we define $\Delta\mathbf{k} \equiv \mathbf{k'} - \mathbf{k}$.
We have assumed the scattering centers to be at the lattice points, so

$$\boldsymbol{\rho} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}$$

Now the total scattered radiation amplitude seen at $\mathbf{R}$ is proportional to

$$\mathcal{A} \equiv \sum_{\boldsymbol{\rho}} \exp(-i\rho\cdot\mathbf{k}) = \left(\sum_{n_1} \exp[-in_1(\mathbf{a}\cdot\Delta\mathbf{k})]\right)\left(\sum_{n_2} \exp[-in_2(\mathbf{b}\cdot\Delta\mathbf{k})]\right)\left(\sum_{n_3} \exp[-in_3(\mathbf{c}\cdot\Delta\mathbf{k})]\right)$$

The intensity is proportional to the square of the amplitude and thus proportional to

$$|\mathcal{A}|^2 = \left|\sum_{n_1} \exp[-in_1(\mathbf{a}\cdot\Delta\mathbf{k})]\right|^2 \left|\sum_{n_2} \exp[-in_2(\mathbf{b}\cdot\Delta\mathbf{k})]\right|^2 \left|\sum_{n_3} \exp[-in_3(\mathbf{c}\cdot\Delta\mathbf{k})]\right|^2$$

We now take a closer look at the first sum in this expression. We assume the crystal has a dimension Ma in the direction $\mathbf{a}$, where M is an integer. Thus the sum becomes

$$\sum_{n}^{M-1} \exp[-in(\mathbf{a}\cdot\Delta\mathbf{k})]$$

Using the series

$$\sum_{n=0}^{M-1} x^n = \sum_{n=0}^{\infty} x^n - \sum_{n=M}^{\infty} x^n = \frac{1}{1-x} - \frac{x^M}{1-x}$$

with $x = \exp[-i(\mathbf{a}\cdot\Delta\mathbf{k})]$, we get

$$\begin{aligned}\sum_n^{M-1} \exp[-in(\mathbf{a}\cdot\Delta\mathbf{k})] &= \frac{1-\exp[-iM(\mathbf{a}\cdot\Delta\mathbf{k})]}{1-\exp[-i(\mathbf{a}\cdot\Delta\mathbf{k})]} \\ &= \frac{\exp[-\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})]}{\exp[-i\frac{1}{2}(\mathbf{a}\cdot\Delta\mathbf{k})]} \cdot \frac{\exp[\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})] - \exp[-\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})]}{\exp[\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})] - \exp[-\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})]}\end{aligned}$$

Multiplying this by its complex conjugate, we get

$$\begin{aligned}\left|\sum_n \exp[-in(\mathbf{a}\cdot\Delta\mathbf{k})]\right|^2 &= \frac{\exp[\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})] - \exp[-\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})]}{\exp[\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})] - \exp[-\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})]} \cdot \frac{\exp[-\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})] - \exp[\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})]}{\exp[-\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})] - \exp[\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})]} \\ &= \frac{-1}{-1}\left(\frac{\cancel{\cos[\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})]} + i\sin[\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})] - \cancel{\cos[\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})]} + i\sin[\frac{1}{2}iM(\mathbf{a}\cdot\Delta\mathbf{k})]}{\cancel{\cos[\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})]} + i\sin[\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})] - \cancel{\cos[\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})]} + i\sin[\frac{1}{2}i(\mathbf{a}\cdot\Delta\mathbf{k})]}\right)^2 \\ &= \frac{\sin^2\frac{1}{2}M(\mathbf{a}\cdot\Delta\mathbf{k})}{\sin^2\frac{1}{2}(\mathbf{a}\cdot\Delta\mathbf{k})}.\end{aligned}$$

This function is steeply peaked with maxima whenever

$$\mathbf{a}\Delta\cdot\mathbf{k} = 2\pi q$$

with q an integer.

3.3.1 Diffraction conditions

Repeating the analysis for the directions $\mathbf{b}$ and $\mathbf{c}$, we obtain the **Laue equations**

$$\boxed{\mathbf{a}\Delta\cdot\mathbf{k} = 2\pi q \qquad \mathbf{b}\Delta\cdot\mathbf{k} = 2\pi r \qquad \mathbf{c}\Delta\cdot\mathbf{k} = 2\pi s}$$

where q, r, s are integers. We will later show that these equations are equivalent to the Bragg law.

3.3.2 Width of the maximum

To get an idea of the width of the peak, we note that the expression

$$\sin^2\frac{1}{2}M(2\pi q + x)$$

has a zero at $x = \frac{2\pi}{M}$. This is the closest zero to the peak. We conclude that the width of the peak is proportional to $1/M$.

3.4 The reciprocal lattice

It turns out the possible values of $\Delta\mathbf{k}$ also form a periodic lattice, namely the reciprocal lattice:

$$\Delta\mathbf{k} = q\mathbf{A} + r\mathbf{B} + s\mathbf{C}$$

where q, r, s are integers and the fundamental vectors of the reciprocal lattice are given by

$$\boxed{\mathbf{A} = 2\pi\frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}} \qquad \mathbf{B} = 2\pi\frac{\mathbf{c}\times\mathbf{a}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}} \qquad \mathbf{C} = 2\pi\frac{\mathbf{a}\times\mathbf{b}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}}$$

It is obvious that any $\Delta\mathbf{k}$ of this form is a solution to the Laue equations (just fill it in and see that they are satisfied). The denominators make sense because the triple product is cyclic. Conversely, any vector Δk can be written as a linear combination of $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$:

$$\Delta\mathbf{k} = \alpha\mathbf{A} + \beta\mathbf{B} + \gamma\mathbf{C}$$

we now need to show that α, β and γ are integers. We illustrate this for α. Using the fact that $\mathbf{a} \cdot \mathbf{A} = 2\pi$ and $\mathbf{a} \cdot \mathbf{B} = \mathbf{a} \cdot \mathbf{C} = 0$, we write

$$\mathbf{a} \cdot \Delta\mathbf{k} = \alpha\mathbf{a} \cdot \mathbf{A} = \alpha 2\pi.$$

From the Laue equations we see that α must be equal to the integer q and is thus an integer.

3.4.1 Interpretation and properties of the reciprocal lattice

We write a generic lattice point as

$$\boldsymbol{\rho} = m\mathbf{a} + n\mathbf{b} + p\mathbf{c} \qquad m, n, p \quad \text{integers}$$

and a generic reciprocal lattice point as

$$\mathbf{G} = h\mathbf{A} + k\mathbf{B} + l\mathbf{C} \qquad h, k, l \quad \text{integers}$$

We note the following properties of the reciprocal lattice:

- Vectors $\boldsymbol{\rho}$ of the lattice have the dimensions of $[L]$. Vectors $\mathbf{G}$ of the reciprocal lattice have the dimensions of $[L^{-1}]$.
- For any lattice points $\boldsymbol{\rho}$ and $\mathbf{G}$ the scalar product is

  $$\begin{aligned}\mathbf{G} \cdot \boldsymbol{\rho} &= (h\mathbf{A} + k\mathbf{B} + l\mathbf{C}) \cdot (m\mathbf{a} + n\mathbf{b} + p\mathbf{c}) \\ &= 2\pi(hm + kn + lp) = 2\pi \times (\text{integer}).\end{aligned}$$

  Hence also

  $$\exp(i\mathbf{G} \cdot \boldsymbol{\rho}) = 1.$$

- A diffraction pattern is a map of the reciprocal lattice of the crystal, in contrast to a microscope image which is a map of the real crystal structure. (TODO: clarify!!)
- The reciprocal lattice is a lattice in **Fourier space**. If we have a quantity (such as for example electron density) that is invariant under all crystal lattice translations, then the only values of $\mathbf{k}$ that contribute to the Fourier series are the lattice points $\mathbf{G}$.

  To prove this, say we have a periodic quantity $n(\mathbf{r})$ such that $n(\mathbf{r}) = n(\mathbf{r} + \boldsymbol{\rho})$ for any crystal lattice vector $\boldsymbol{\rho}$. We write the Fourier series

  $$n(\mathbf{r}) = \sum_{\mathbf{k}} n_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{\mathbf{k}} n_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}+\boldsymbol{\rho})} = \sum_{\mathbf{k}} n_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\boldsymbol{\rho})}$$

  So if we want the contribution associated with a wavevector $\mathbf{k}$ to be non-vanishing, we need $e^{i\mathbf{k}\cdot\boldsymbol{\rho})} = 1$ for all lattice point vectors $\boldsymbol{\rho}$. We have shown this to be the case for reciprocal lattice vectors $\mathbf{G}$. In fact it is only the case for reciprocal lattice vectors, because if we set $\boldsymbol{\rho} = \mathbf{a}$, $\boldsymbol{\rho} = \mathbf{b}$ and $\boldsymbol{\rho} = \mathbf{c}$ we get the Laue equations. Thus we can write the Fourier series as

  $$n(\mathbf{r}) = \sum_{\mathbf{G}} n_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}$$

- Every reciprocal lattice vector is normal to a lattice plane of the crystal lattice. In particular the vector
$$\mathbf{G}(hkl) \equiv h\mathbf{A} + k\mathbf{B} + l\mathbf{C}$$
is normal to the plane with Miller indices (hkl). In fact this fact can be used as an alternative definition of Miller indices. TODO proof.
- TODO spacing of planes of crystal lattice plus equivalence of Bragg law.

3.4.2 Ewald construction

TODO + fig 23 Kittel.

3.4.3 Brillouin zones

The primitive cell formed in the reciprocal lattice according to the Wigner-Seitz method is called the first Brillouin zone. This will be essential in the theory of electronic energy bands.

3.4.4 Examples

3.4.4.1 Reciprocal lattice to an bcc lattice.

TODO

3.4.4.2 Reciprocal lattice to an fcc lattice.

TODO

3.5 Geometrical structure factor

3.5.1 Of the bcc lattice

3.5.2 Of the fcc lattice

3.6 Atomic scattering or form factor

3.7 Relating the crystal structure to the diffraction pattern

Chapter 4

Crystal binding

4.1 Crystals of inert gasses

4.2 Ionic crystals

4.3 Covalent crystals

4.4 Metal crystals

4.5 Hydrogen-bonded crystals

4.6 Atomic radii

TODO: nuclear physics?

Chapter 5

Lattice vibrations

Chapter 6

Thermal properties of insulators

Chapter 7

Free electron Fermi gas

Why does this work so well? https://physicstoday.scitation.org/doi/10.1063/1.2995618

Chapter 8

Energy bands

8.1 Ansatz: the nearly free electron model

We essentially treat the electrons as being confined to the solid and only weakly perturbed by the periodic potential of the ion cores.

8.2 Wave equation of electron in a periodic potential

8.3 Approximate solution near a zone boundary

8.4 Number of states in a band

8.5 Construction of Fermi surfaces

8.6 Electrons, holes, and open orbits

8.7 Effective mass of electrons in crystals

8.8 Wavefunctions for zero wavevector

8.9 Orthogonalized plane waves

8.10 Experimental methods in fermi surface studies

Chapter 9

Semiconductor crystals

Chapter 10

Superconductivity

Chapter 11

Magnetism

The magnetic moment of a free atom has three principal sources:

1. The spin of the electrons;
2. Orbital angular momentum of the electrons about the nucleus;
3. Change in orbital moment induced by an applied magnetic field.

Nuclear magnetic moments are of the order of 10^{-3} times smaller than the magnetic moment of the electron.

11.1 Magnetism in a free atom

The magnetisation M is the magnetic moment per unit volume.
The magentic susceptibility χ per unit volume is defined as

$$\chi \equiv \frac{\mu_0 M}{B}$$

where B is the macroscopic magnetic field intensity.

- Substances with a *negative* magnetic susceptibility are called diamagnetic.
- Substances with a *positive* magnetic susceptibility are called paramagnetic.

Chapter 12

Applying an electric field

12.1 Dielectric properties

12.2 Ferroelectric crystals

Chapter 13

Optical phenomena in insulators

Chapter 14

Point defects

Chapter 15

Dislocations

Part XVII

Soft condensed matter physics and biophysics

Chapter 1

Soft condensed matter physics

1.1 Amorphous solids

1.2 Liquid crystals

1.3 Colloids

Chapter 2

Biophysics

2.1 Polymers

Part XVIII

Nanophysics

Chapter 1

Introduction

SKIPPED SLIDES: 23, 34 (Lotus effect), 39, 47-52, 59,60, 76-83, 85-91, 115-124, 131-140, 145-146
Nanoscience studies objects at the nanometer scale. Practically this means objects smaller than 100 nm. Examples of such objects include virusses, clusters (small groups of atoms), DNA, fullerenes, nanotubes ...
Objects in this regime are typically too small for classical physics to make useful predictions and too large for quantum techniques to be feasible.
One major consequence of this is that material properties (such as colour, melting temperature etc.) become dependent on the size of the objects. In bulk this is obviously not the case.
There are different types of nanomaterials. Examples include

1. Nanoclusters. There are many types of nanoclusters, with differing levels of internal structure:
   (a) Simple nanoparticles
   (b) Janus particles: half one material and half another
   (c) Core-shell
   (d) Nano-planets
   (e) DLA fractals
   (f) Nano-prism
   (g) Nano-shell
   (h) Nano-hole
2. Nanostructured materials (3D). This is material with nanoscale features such as dislocations, grain boundaries or defects.
3. Nanostructured film.
4. Multilayered materials.
5. Novel materials such as metamaterials.

Synthesis techniques for nanomaterials can broadly be categorized as either bottom-up or top-down. Bottom-up synthesis includes techniques like nano-positioning and self-assembly. Top-down techniques include nanofabrication and nanolithography.
Applications of nanostructures include the following:

1. Applications based on optical properties (in which case metals such as copper, silver and gold are used).
   (a) Plasmonic sensors
   (b) Plasmonic waveguides
   (c) Nano-antennas
   (d) Iperthermia (cancer therapy)
   (e) Extroordinary transmission of light (EOT)
   (f) Metamaterials
2. Applications based on magnetic properties (in which case metals such as iron, cobalt and nickel are used).
   (a) Super-paramagnetism
   (b) Magneto-optical properties
   (c) Medical diagnosis (Fe-oxides, NMR)
3. Applications based on catalytic properties (in which case metals such as iron, cobalt, nickel, ruthenium, rhodium, palladium, osmium, iridium and platinum are used).
   (a) Catalysis and photo-catalysis
   (b) Water splitting and hydrogen storage.

Chapter 2

Some descriptive parameters

- We denote the number of atoms in a cluster N.
- We denote the (ionic) atomic radius R_0. In a face-centered cubic lattice, with lattice constant a, the distance between atoms is $a/\sqrt{2}$. So

$$R_{0,\text{fcc}} = \frac{a}{4}\sqrt{2}$$

- Cluster volume V. This definition disregards the atomic packing factor (TODO? mainly N dependence as with others).

$$V = \frac{4\pi}{3} R_0^3 N$$

- We can get an effective radius R_{eff} by assuming the volume is spherical: $V = \frac{4\pi}{3} R_{\text{eff}}^3$. Thus we get

$$R_{\text{eff}} = \left(\frac{V}{4\pi/3}\right)^{1/3} = R_0 N^{1/3}$$

- Cluster surface S.

$$S = 4\pi R_{\text{eff}}^2 = 4\pi R_0^2 N^{2/3}$$

- Number of surface atoms N_S.

$$N_S = \frac{S}{S_A} = \frac{4\pi R_0^2 N^{2/3}}{\pi R_0^2} = 4N^{2/3}$$

  With S_A the fraction of the surface that a single atom occupies. This is essentially the cross-sectional surface (hence πR_0^2).

- Finally we define the fraction of surface atoms F:

$$F = \frac{N_S}{N} = \frac{4}{N^{1/3}} = \frac{4R_0}{R_{\text{eff}}}$$

  We also give some values for F:

N	F
10^2	0.86
10^3	0.40
10^4	0.04

Table 2.1: Fraction of surface atoms

Chapter 3

Size equations

When we get down to nano scales, chemical and physical properties start to depend on size. We postulate that this dependency takes the following form, where A represents a generic property:

$$A(N) = A(\infty)\left(1 + \frac{C_N}{N^\alpha}\right) \qquad A(R) = A(\infty)\left(1 + \frac{C_R}{R^\alpha}\right)$$

The equation of the left gives the property in function of the number of atoms, N. The equation on the right gives it in function of the radius R. The bulk limit of A is denoted $A(\infty)$ and C and α are constants.

Chapter 4

Thermodynamic properties of nanostructured materials

4.1 Melting temperature size equation

We want to know how melting temperature T_M depends on the the size of the nanoparticles. Intuitively we can guess that the melting temperature would be lower for nanoparticles, because they have non-negligible fractions of there atoms at the surface and surface atoms have more freedom to move and be excited by thermal energy. We would, however, like a more formal argument.

4.1.1 Liquid drop

Phase transition occurs when there is an equilibrium between the solid (S) and liquid (L) phase. (TODO: why exactly)

$$\mu_L(T_L, P_L) = \mu_S(T_S, P_S)$$

The subscripts L an S are added to the temperature and pressure because they may not be the same for each phase. In fact at equilibrium there must also be thermal equilibrium, so

$$T_L = T_P = T$$

which is the melting temperature we are looking for. The same cannot be said for the pressure. We can expand the chemical potential to first order

$$\mu(T, P) = \mu(T_0, P_0) + \frac{\partial \mu}{\partial T}(T - T_0) + \frac{\partial \mu}{\partial P}(P - P_0) + \dots$$

We can get expressions for the partial derivatives from the Gibbs-Duhem relation

$$S\,\mathrm{d}T - V\,\mathrm{d}P + N\,\mathrm{d}\mu = 0$$

which we can rewrite

$$\mathrm{d}\mu = -\frac{S}{N}\,\mathrm{d}T + \frac{V}{N}\,\mathrm{d}P = -s\,\mathrm{d}T + \frac{1}{\rho}\,\mathrm{d}P$$

where we define

$$\begin{cases} s \equiv \frac{S}{N} = -\left(\frac{\partial \mu}{\partial T}\right)_P \\ \frac{1}{\rho} \equiv \frac{V}{N} = \left(\frac{\partial \mu}{\partial P}\right)_T \end{cases}$$

The equilibrium condition, to first order, is given by

$$0 = \mu_L(T_0, P_0) - s_L(T - T_0) + \frac{1}{\rho_L}(P_L - P_0) - \mu_S(T_0, P_0) + s_S(T - T_0) - \frac{1}{\rho_S}(P_S - P_0) + \dots$$

We also choose the for expansion of the chemical potential to be around the the triple point of the bulk phase. At this point

$$\mu_L(T_0, P_0) = \mu_S(T_0, P_0)$$

further simplifying the equilibrium condition to

$$0 = (s_L - s_S)(T - T_0) - \frac{1}{\rho_L}(P_L - P_0) + \frac{1}{\rho_S}(P_S - P_0) + \dots$$

In order to get expressions for the pressures P_L and P_S, we use the Young-Laplace equation for the pressure difference across an interface between two static fluids. (TODO: justification, assumes Curie-Wulff crystal? Refer Ph. Buffat , J P . Borel , Phys . Rev. A 13 (1976) 2287) For spheres Laplace's law gives us the following

$$\begin{cases} P_L = P_{\text{ext}} + 2\frac{\gamma_L}{R_L} \approx 2\frac{\gamma_L}{R_L} \\ P_S = P_{\text{ext}} + 2\frac{\gamma_S}{R_S} \approx 2\frac{\gamma_S}{R_S} \end{cases}$$

where P_{ext} is the pressure not due to the interface. For small particles this in negligible compared to the pressure gradient across the interface. The constants γ_L and γ_S are the surface tension of the liquid ans solid phase, respectively.
Recognising that the nanoparticle has the same number of atoms, regardless of which phase it is in, we get

$$R_S = \left(\frac{\rho_L}{\rho_S}\right)^{1/3} R_L$$

Incorporating these equations into the equilibrium condition, we get

$$0 = (s_L - s_S)T_0 \left(\frac{T}{T_0} - 1\right) + 2\left(\frac{\gamma_S}{R_S\rho_S} - \frac{\gamma_L}{R_L\rho_L}\right) + \dots$$

Finally we can rewrite this using that the latent heat of fusion per atom is given by

$$L = (s_L - s_S)T_0$$

and assuming $\rho_L \approx \rho_S$, so we can drop the last term:

$$\frac{\Delta}{T_0} \equiv \frac{T - T_0}{T_0} = -\frac{2}{LR_S\rho_S}\left(\gamma_S - \gamma_L\left(\frac{\rho_S}{\rho_L}\right)^{2/3}\right) = -\frac{A}{R} < 0$$

where A is just a constant. In fact it is the constant in the size equation

$$T_M(R) = T_M(\infty)\left(1 - \frac{A}{R}\right)$$

In deriving this result, we made the following assumptions and approximations:

1. The chemical potential was only expanded to first order;
2. Spherical clusters;

3. $\rho_L \approx \rho_S$.

Unfortunately this model does not fit reality very well. (TODO image) The obvious way to improve the model is to expand the chemical potential to second order. To do that we need expressions for the second order partial derivatives of μ:

$$\begin{cases} \frac{\partial^2 \mu}{\partial P^2} = -\frac{1}{\rho^2}\frac{\partial \rho}{\partial P} = -\frac{\chi}{\rho} \\ \frac{\partial^2 \mu}{\partial T^2} = -\frac{\partial s}{\partial T} = -\frac{C_P}{T} \\ \frac{\partial \mu}{\partial P}T = -\frac{1}{\rho^2}\frac{\partial \rho}{\partial T} \approx \frac{3\alpha}{\rho} \end{cases}$$

where χ is the isothermal compressibility coefficient, C_P is the specific heat at constant pressure and α is the linear-expansion coefficient.
We can the equilibrium condition is then

$$L\frac{T_0 - T}{T_0} - \frac{2}{\rho_S R_S}\left[\gamma_S - \gamma_L\left(\frac{\rho_S}{\rho_L}\right)^{2/3}\right] + \frac{C_{P,S} - C_{P,L}}{2}T_0\left(\frac{T_0 - T}{T_0}\right)^2 - \frac{2}{\rho_S R_S}\left[\gamma_S(\eta_S - 2\alpha_S) - \gamma_L(\eta_L - 2\alpha_L)\left(\frac{\rho_S}{\rho_L}\right)^{2/3}\right]$$

where $\eta = -\frac{1}{\gamma}\frac{\partial \gamma}{\partial T}$.
Before considering other improvements to the liquid drop model, we will first consider another way to obtain the same results, based on a different equilibrium condition. We now imagine the nanoparticles suspended in a (vapor, TODO: necessary?) medium. We now write the Gibbs free energy of the entire nanoparticle, in function of the radius R

$$\begin{cases} G_S = N\mu_S + 4\pi R_S^2 \gamma_{SV} \\ G_L = N\mu_L + 4\pi R_L^2 \gamma_{LV} \end{cases}$$

where γ_{SV} and γ_{LV} are now the interfacial energies between vapor and the solid or liquid, respectively (TODO explicate difference).
Using $R_S = \left(\frac{\rho_L}{\rho_S}\right)^{1/3} R_L$ and

$$\begin{aligned} (\mu_L - \mu_S) &= \Delta G_{\text{vol}} = \Delta g_{\text{vol}} \\ &= \Delta h - T\Delta s \\ &= L - T\frac{L}{T_0} \end{aligned}$$

(TODO explain) we can write the difference in Gibbs free energy as

$$\begin{aligned} \Delta G &\equiv G_L - G_S \\ &= N(\mu_L - \mu_S) + 4\pi R^2\left[\gamma_{LV}\left(\frac{\rho_S}{\rho_L}\right)^{2/3} - \gamma_{SV}\right] \\ &= \frac{4\pi R^3}{3}\rho_S L\frac{T_0 - T}{T_0} + 4\pi R^2\left[\gamma_{LV}\left(\frac{\rho_S}{\rho_L}\right)^{2/3} - \gamma_{SV}\right]. \end{aligned}$$

We can now apply a different equilibrium condition, namely $\frac{\partial \Delta G}{\partial R} = 0$. (TODO: why?) This yields

$$\frac{\Delta T}{T_0} = -\frac{2}{\rho_S L R}\left(\gamma_{SV} - \gamma_{LV}\left(\frac{\rho_S}{\rho_L}\right)^{2/3}\right)$$

which is the same result we obtained before.

4.1.2 Liquid layer model (LLM)

Another strategy is to consider not only equilibrium, but also the transition. In a first attempt we consider the transition from solid to the situation where the core of the nanosphere is solid, but a surface layer of thickness δ is liquid. For the calculation we use the method of calculating the total Gibbs free energy and differentiating with respect to the radius as before, except now the first situation is that of a solid nanosphere and the second is the situation with the liquid layer.
In order to simplify calculations, we assume $\rho_S \approx \rho_L \approx \rho$. We can then write

$$\begin{cases} G_1 = N\mu_s + 4\pi R^2 \gamma_{SV} \\ G_2 = (N - N')\mu_S + N'\mu_L + 4\pi R^2 \left[\gamma_{LV} + \gamma_{SL}\left(1 - \frac{\delta}{R}\right)^2\right] \end{cases}$$

where γ_{SL} is the interfacial energy between the solid and the liquid, and N' the number of atoms in the liquid phase. So

$$\Delta G_{2,1} = G_2 - G_1 = N'(\mu_L - \mu_S) + 4\pi R^2 \left[\gamma_{LV} + \gamma_{SL}\left(1 - \frac{\delta}{R}\right)^2 - \gamma_{SV}\right]$$

We can calculate the volume contribution in a similar fashion as before

$$N'(\mu_L - \mu_S) = V_L \rho L \frac{T_0 - T}{T_0} = \frac{4\pi}{3}(R^3 - (R-\delta)^3)\rho L \frac{T_0 - T}{T_0}$$

with $V_L = \frac{4\pi}{3}(R^3 - (R-\delta)^3)$ the volume of the liquid part. This gives

$$\Delta G_{2,1} = \frac{4\pi}{3}(R^3 - (R-\delta)^3)\rho L \frac{T_0 - T}{T_0} + 4\pi R^2 \left[\gamma_{LV} + \gamma_{SL}\left(1 - \frac{\delta}{R}\right)^2 - \gamma_{SV}\right]$$

Applying the condition $\frac{\partial \Delta G_{2,1}}{\partial R} = 0$ and assuming $\gamma_{LV} \approx \gamma_{SV}$ (TODO: is that really what we are doing??), we get

$$\frac{\Delta T}{T_0} = -\frac{2\gamma_{SL}}{\rho_S L(R-\delta)} = -\frac{A}{R-\delta}$$

with A some constant. If we no longer assume $\rho_S \approx \rho_L \approx \rho$, we get

$$\frac{\Delta T}{T_0} = -\frac{2\gamma_{SL}}{\rho_S L(R-\delta)} - \frac{2\gamma_{LV}}{\rho_S L R}\left(1 - \frac{\rho_S}{\rho_L}\right) = -\frac{A}{R-\delta} - \frac{B}{R}$$

with again B some constant. (TODO: why not power of 2/3?)

4.1.3 Smooth interfaces interaction (SII).

This model is a refinement of the liquid layer model. We split the total difference in Gibbs free energy obtained above into a volume contribution,

$$\Delta G_V = \frac{4\pi}{3}(R^3 - (R-\delta)^3)\rho L \frac{T_0 - T}{T_0},$$

and a surface contribution. The difference is that, in order to account for the surfaces not being perfectly sharp, as assumed in the liquid layer model, the surface contribution is diminished by a factor that depends on the coherence length ξ:

$$\Delta G_S = 4\pi R^2 \left[\gamma_{LV} + \gamma_{SL}\left(1 - \frac{\delta}{R}\right)^2 - \gamma_{SV}\right] \cdot \left(1 - e^{-\frac{\delta}{\xi}}\right)$$

This gives us the following

$$\frac{\Delta T}{T_0} = -\frac{2\gamma_{SL}}{\rho_S L(R-\delta)}\left(1-e^{-\frac{\delta}{\xi}}\right) - \frac{\Gamma R^2}{\rho_S L\xi(R-\delta)^2}e^{-\delta/\xi} = -\frac{A}{R-\delta}\left(1-e^{-\frac{\delta}{\xi}}\right) - \frac{B\Gamma R^2}{(R-\delta)^2}$$

where the Γ has been defined for ease of notation as

$$\Gamma \equiv \gamma_{SV} - \left[\gamma_{LV} + \gamma_{SL}\left(1-\frac{\delta}{R}\right)^2\right].$$

We notice that the result of the SII theory approaches that of the LLM as $\xi \to 0$, which is how it should be.

4.1.4 Thermal expansion variation (TEV)

Here we simply give the result (TODO: more explanation)

$$\frac{\Delta T}{T_0} = -\frac{2}{\rho_S RL}\left(\gamma_{SM} - \gamma_{LM}\left(\frac{\rho_S}{\rho_L}\right)^{2/3}\right) + \frac{\Delta E}{L}$$

where ΔE is the strain energy difference (thermal expansion matrix-cluster).

Chapter 5

Synthesis of nanostructures

In order to synthesize nanostructured materials, we can use a variety of techniques:

- Physical techniques:
  - Condensation from vapor phase
  - Free expansion molecular beams
  - Sputtering (physical vapor Deposition)
  - Ion implantation
  - Ball-milling
  - Lithography, nanofabrication
  - Laser ablation
- Chemical techniques:
  - Colloidal chemistry
  - Sol-gel
  - Chemical vapor deposition
- Mixed approaches combining physical and chemical techniques

5.1 Nucleation

TODO: refer to other part??
In this section we describe to formation of droplets in a homogeneous supersaturated solution, i.e. without impurities. We assume the nucleating embryos are spherical.
The first step is to write down the Gibbs free energy of a nucleating droplet, in function of the number of atoms it has already gathered

$$\begin{aligned}\Delta G(N) &= -N\Delta\mu + \gamma 4\pi R^2 \\ &= -N\Delta\mu + \gamma 4\pi R_0^2 N^{2/3}\end{aligned}$$

where $\Delta\mu$ is the bulk change in Gibbs free energy per atom, i.e. the chemical potential. If we plot this in function of N, see figure TODO, we see that ΔG has a peak at N^*, which we call the critical nucleus.

We can find an expression for the critical nucleus using the condition $\frac{\partial \Delta G}{\partial N}\big|_{N^*} = 0$:

$$N^* = \frac{32\pi\gamma^3}{3\rho^2\Delta\mu^3}$$

The difference in Gibbs free energy at this peak is

$$\Delta G(N^*) = \frac{16\pi\gamma^3}{3\rho^2\Delta\mu^2}.$$

Particles that are smaller than the critical nucleus are unstable, because the system tries to minimise ΔG. So in order for nucleation to occur, thermal fluctuations must overcome the energy barrier of $\Delta G(N^*)$. We can then use Arrhenius' equation to get the nucleation speed (TODO proper ref)

$$J = K \exp(-\Delta G(N^*)/k_B T)$$

We can also rewrite these quantities in function of the radius R, using the relation

$$N = \frac{4\pi}{3} R^3 \rho.$$

We obtain

$$\begin{cases} \Delta G(R) = -\frac{4\pi}{3} R^3 \rho \Delta\mu + \gamma 4\pi R^2 \\ R^* = \frac{2\gamma}{\rho\Delta\mu} \end{cases}$$

Finally we would like to give a more explicit expression for $\Delta\mu$. In order to do so we make use of the Gibbs-Thomson equation which gives the change in vapor pressure or chemical potential across a curved surface or interface.

5.1.1 Gibbs-Thomson equation.

We consider a situation where we have a vapor in an equilibrium with the liquid phase at a pressure P_e and a temperature T. In the previous section we saw that there is a barrier to the condensation of the vapor to form clusters via nucleation. In order to overcome that barrier, we need to raise the pressure above the equilibrium value to a new pressure $P(R)$.
To arrive at the Gibbs-Thomson equation, we need to combine several results.

1. **Isothermal compression of an ideal gas**. We start at the ambient external pressure P_e and increase the pressure isothermically until we reach $P(R)$.

   The differential of the Gibbs free energy is given by

   $$\mathrm{d}G = V\,\mathrm{d}P - S\,\mathrm{d}T + \mu\,\mathrm{d}N.$$

   To get the total change in Gibbs free energy between the start and the end of the process, we integrate. Using the fact that there is no change in temperature or number of particles as well as making use of the ideal gas law, we can write

   $$\Delta G = \int_{P_e}^{P(R)} V\,\mathrm{d}P = N k_B T \int_{P_e}^{P(R)} \frac{\mathrm{d}P}{P} = N k_B T \ln\frac{P(R)}{P_e}$$

   Using the relation $G = N\mu$, we can write

   $$\mu_V[P(R)] - \mu_V[P_e] = \frac{\Delta G}{N} = k_B T \ln\frac{P(R)}{P_e}.$$

   The subscript V shows we are working with a vapor.

2. **Isothermal compressions of a liquid**. We assume that the liquid is incompressible, so the compression occurs with no volume variation. We write the chemical potential inside a nanoparticle in function of the radius and the vapor pressure

$$\mu_L(R, P).$$

We have for a liquid layer (TODO WHY???)

$$\mu_L(\infty, P(R)) \approx \mu_L(\infty, P_e).$$

Setting R to infinity, we are effectively neglecting the surface contribution to the free energy.

Previously we calculated the critical radius in function of the volume contribution to the Gibbs free energy per atom. We can invert that relation to get

$$\Delta\mu = \frac{2\gamma}{\rho R}.$$

The condition for the formation of a cluster with radius R at the pressure $P(R)$ is

$$\mu_V(P(R)) = \mu_L(R, P(R))$$

now we can split $\mu_L(R, P(R))$ into a surface and volume term:

$$\begin{aligned}\mu_L(R, P(R)) &= \mu_L(\infty, P(R)) + \frac{2\gamma}{\rho R} \\ &= \mu_L(\infty, P_e) + \frac{2\gamma}{\rho R} \\ &= \mu_V(P_e) + \frac{2\gamma}{\rho R}\end{aligned}$$

where the last step was justified because P_e is the liquid-vapor equilibrium, so $\mu_L(\infty, P_e) = \mu_V(P_e)$. So we get

$$\mu_V(P(R)) - \mu_V(P_e) = \frac{2\gamma}{\rho R}.$$

But we have already found another expression for the left side of this equition, considering the isothermal compression of an ideal gas. So we can write

$$k_B T \ln \frac{P(R)}{P_e} = \frac{2\gamma}{\rho R}$$

or

$$P(R) = P_e e^{\frac{2\gamma}{k_B T \rho R}}.$$

5.1.2 Making nuclei form.

There are multiple ways we can get nuclei to form:

1. By decreasing P_e (i.e. by lowering T):
    - If we expand the gas adiabatically it will cool down. If it is expanded enough nucleation will occur.

2. By getting over the critical radius:
    - Chemical or photo chemical decomposition in the gas phase
    - Sputtering
    - Thermal evaporation
    - Laser ablation
    - Ion implantation

5.2 Growth of nanoparticles

In order for condensation via nucleation to start, we need a supersaturated vapor. The first step then, is for nuclei to form. This can happen spontaneously due to thermal fluctuations or there can be heterogeneous nucleation, such as with ion implantation.
So long as the vapor pressure is well above P_e, different clusters of atoms are free to grow independently and thus non competitively. The vapor pressure goes down as it turns to liquid. This phase is called diffusion limited aggregation (DLA).
At some point the pressure will get low enough that the clusters start competing for resources (remember that the clusters are never stable unless they have grown as beg as they can or have shrunk out of existence). Now the bigger clusters keep on growing by shrinking the smaller clusters. This process is called coarsening, or Ostwald ripening (OR).

5.2.1 Diffusion-limited aggregation

During this phase, clusters grow as

$$R^2(t) = R_0^2 + K_1 D t$$

To calculate the growth, we start with the total flux of atoms through the surface Σ of the cluster (with outbound the positive direction) and assume isotropicity for ease of calculation.

$$\Phi_\Sigma = \left.\frac{\mathrm{d}N}{\mathrm{d}t}\right|_{\text{out}} - \left.\frac{\mathrm{d}N}{\mathrm{d}t}\right|_{\text{in}}$$

we can rewrite this using

$$\begin{cases} \left.\frac{\mathrm{d}N}{\mathrm{d}t}\right|_{\text{in}} = C_p \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{4\pi}{3}R^3\right) = 4\pi R^2 C_p \frac{\mathrm{d}R}{\mathrm{d}t} \\ \left.\frac{\mathrm{d}N}{\mathrm{d}t}\right|_{\text{out}} = C_e \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{4\pi}{3}R^3\right) = 4\pi R^2 C_e \frac{\mathrm{d}R}{\mathrm{d}t} \end{cases}$$

where C_p is the atomic concentration inside the nanoparticle and C_e is the atomic concentration just outside. Figure TODO (+ TODO explanation of curve) gives the atomic concentration in function of the distance from the centre of the nanoparticle.
Using Fick's first law

$$\mathbf{J} = -D\,\nabla C$$

we can write the flux as

$$\Phi_\Sigma = \int_\Sigma \mathbf{J}\cdot\hat{n}\,\mathrm{d}\Sigma = -4\pi R^2 D \left.\frac{\partial C(r,t)}{\partial r}\right|_{r=R}$$

Putting these expressions together, we get

$$(C_p - C_e)\frac{\mathrm{d}R}{\mathrm{d}t} = D \left.\frac{\partial C(r,t)}{\partial r}\right|_{r=R}.$$

Now we want an expression for $\left.\frac{\partial C(r,t)}{\partial r}\right|_{r=R}$. To get one, we linearize the gradient

$$\left.\frac{\partial C(r,t)}{\partial r}\right|_{r=R} \approx \frac{\Delta C}{\Delta r}$$

We can calculate ΔC and Δr by recognizing that the areas A_1 and A_2 in figure TODO must be the same due to conservation of the number of atoms. (TODO: Why??). The result is

$$\frac{\Delta C}{\Delta r} = \frac{(C_s - C_e)^2}{2(C_p - C_s)R}$$

Integrating our previous result with respect to t, we finally get

$$R^2(t) = R_0^2 + \frac{(C_s - C_e)^2}{(C_p - C_e)(C_p - C_s)} Dt$$

or

$$R^2(t) = R_0^2 + K_1 Dt \qquad \text{with} \qquad K_1 \equiv \frac{(C_s - C_e)^2}{(C_p - C_e)(C_p - C_s)} \approx \frac{C_s^2}{C_p^2}$$

5.2.2 Ostwald ripening

In this regime, growth goes as

$$R^3(t) = R_0^3 + K_2 Dt$$

Ostwald ripening happens when the pressure P is close to (but still above) P_e. If there are two particles of radii R_1 and R_2, with $R_1 < R_2$, then at some point the pressure will be below $P(R_1)$ but still higher than $P(R_2)$. At this point the first particle is no longer stable and will shrink while the second particle grows, effectively feeding off the first.
TODO

5.3 Nanostructures embedded in solid matrices

5.3.1 Ion implantation

for synthesis and processing of metallic nanostructures

5.3.2 Verification of nucleation and growth models

Chapter 6

Optical properties

We now turn to some of the optical properties of nanoparticles. They give rise to some strange phenomena. As an example, the Lycurgus cup (Roman from the fourth century AD) looks green is light is reflected of its surface and red if light is shone through the cup (TODO image). This is due to gold nanoparticles in the glass, which is otherwise similar to modern soda-lime glass.

6.1 Lambert-Beer and optical quantities

Prompted by the strange colour of the Lycurgus cup we would like to study what happens to light when it is shone through a sample containing nanoparticles.
We assume the light passing through the sample is attenuated according to the Lambert-Beer law. The intensity I of the light coming out of the other side of the sample, is given by

$$I(z) = I_0 e^{-\gamma z},$$

where z is the distance traveled through the sample, I_0 is the initial intensity of the light entering the sample and γ is a material constant called the extinction coefficient.
There are two main mechanisms that cause the attenuation: scattering and absorption of photons. We also factor out the volumetric density ρ in order to get the extinction cross-section σ_{ext}:

$$\gamma = \rho\sigma_{\text{ext}} = \rho(\sigma_{\text{abs}} + \sigma_{\text{scat}})$$

We can then define the transmittance T:

$$T \equiv \frac{I(z)}{I_0} = e^{-\gamma z}$$

and the absorbance (or optical density) A:

$$A \equiv \log_{10} \frac{1}{T} = \gamma z \log_{10} e = z\rho\sigma_{\text{ext}} \log_{10} e$$

None of the equations so far show any frequency dependence. All of the quantities discussed so far can in principle depend on frequency. If we measure absorbance for example in function of the frequency of incoming light, we notice a peak at a particular frequency. The location and size of this peak seems to depend on the material and the size of the nanoparticles. It turns out this peak is due to surface plasmon resonance. We will discuss this phenomenon now.

6.2 Localised surface plasmon resonance (LSPR)

6.3 Scattering by spheres (Mie theory)

We want to know incoming light rays interact with small homogeneous spheres. We imagine a sphere suspended in a medium with dielectric constant ϵ_m. The electromagnetic rays have their wave vector, of magnitude k, pointing along the z axis. So, in complex notation, we can write

$$\bar{\mathbf{E}} = \bar{\mathbf{E}}_0 e^{ikz} e^{-i\omega t}$$

6.3.1 Quasi-static regime.

If the electric field does not vary too fast (i.e. $R \ll \lambda$), we can use electrostatics and not take the variation of the electric field into account. As we will see, this leads us to a dipole approximation. Because we assuming the electric field to be basically static, we can write

$$\mathbf{E} = E_0 \hat{z}$$

We now switch to using the electric potential Φ (we use Φ to denote teh electric field, because V is used for volume in this section).
The setup (see figure TODO) is axially symmetric. This means we can use the coordinates r and θ.
We can write the potential outside the nanoparticle as

$$\Phi_{\text{out}}(r,\theta) = \sum_{l=0}^{\infty} \left[B_l r^l + C_l r^{-(l+1)} \right] P_l(\cos\theta)$$

where B_l and C_l are constants we still need to determine and P_l are the Legendre polynomials. In effect the terms with the constants C_l are a multipole expansion of the field due to the polarisation of the nanoparticle. Of course there is also the potential due to incoming photons. This we write as an expansion in powers of r, which we can do because the Legendre polynomials form a complete set (TODO ??).
We also write a similar expansion in powers of r for the potential inside the nanoparticle:

$$\Phi_{\text{in}}(r,\theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos\theta).$$

In order to determine the constants A_l, B_l and C_l, we impose boundary conditions.

1. Very far away ($r \to \infty$) we set (TODO why??)

   $$\Phi_{\text{out}} = -E_0 z = -E_0 r\cos(\theta)$$

   Knowing that $P_1(x) = x$, we can immediately see that this forces

   $$\begin{cases} B_1 = -E_0 \\ B_l = 0 \qquad (l \neq 1) \end{cases}$$

2. At the interface, the tangential component of the electric field $\mathbf{E}$ should be continuous (TODO better explanation):

   $$-\frac{1}{R}\left.\frac{\partial \Phi_{\text{in}}}{\partial \theta}\right|_{r=R} = -\frac{1}{R}\left.\frac{\partial \Phi_{\text{out}}}{\partial \theta}\right|_{r=R}$$

3. At the interface, the radial component of the displacement field $\mathbf{D}$ should be continuous:

$$-\epsilon_0\epsilon \left.\frac{\partial \Phi_{\text{in}}}{\partial r}\right|_{r=R} = -\epsilon_0\epsilon_m \left.\frac{\partial \Phi_{\text{out}}}{\partial r}\right|_{r=R}$$

For $l = 1$ these last two conditions give

$$\begin{cases} A_1 R = -E_0 R + \frac{C_1}{R^2} \\ \epsilon A_1 = -\epsilon_m \left(E_0 + 2\frac{C_1}{R^3}\right) \end{cases} \quad \Rightarrow \quad \begin{cases} A_1 = -\frac{3\epsilon_m}{\epsilon+2\epsilon_m} E_0 \\ C_1 = \frac{\epsilon-\epsilon_m}{\epsilon+2\epsilon_m} R^3 E_0 \end{cases}$$

For $l \neq 1$ these conditions give

$$\begin{cases} A_l R^l = \frac{C_l}{R^{l+1}} \\ \epsilon l A_l R^{l-1} = -\epsilon_m (l+1) \frac{C_l}{R^{l+2}} \end{cases} \quad \Rightarrow \quad \begin{cases} A_l = 0 \\ C_l = 0 \end{cases}$$

We can write the potential outside the nanoparticle as

$$\Phi_{\text{out}}(r,\theta) = -E_0 r\cos\theta + \frac{1}{4\pi\epsilon_0\epsilon_m}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3}$$

where $\mathbf{p}$ is the dipolar moment

$$\mathbf{p} = 4\pi R^3 \epsilon_0 \epsilon_m \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m}\mathbf{E_0}.$$

We can write $\mathbf{p}$ in terms of the polerisability α

$$\mathbf{p} = \alpha\epsilon_0\epsilon_m\mathbf{E_0} \qquad \text{with} \qquad \alpha = 4\pi R^3 \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m}$$

We can write the electric field inside the nanoparticle in terms of the local field enhancement f_e

$$\mathbf{E}_{\text{in}} = \frac{3\epsilon_m}{\epsilon + 2\epsilon_m}\mathbf{E_0} = f_e\mathbf{E_0}$$

Now allowing the incoming field $\mathbf{E_0}$ to vary in time again, we get

$$\begin{cases} \mathbf{E}_{\text{in}}(t) = \frac{3\epsilon_m}{\epsilon+2\epsilon_m}\mathbf{E_0}e^{-i\omega t} \\ \mathbf{E}_{\text{out}}(t) = \mathbf{E_0}e^{-i\omega t} + \frac{1}{4\pi\epsilon_0\epsilon_m}\frac{3(\mathbf{p}\cdot\hat{r})\hat{r}-\mathbf{p}}{r^3} \\ \mathbf{p}(t) = 4\pi R^3 \frac{\epsilon-\epsilon_m}{\epsilon+2\epsilon_m}\epsilon_0\epsilon_m\mathbf{E_0}e^{-i\omega t} \end{cases}$$

We can see that we have a resonance phenomenon when

$$\epsilon + 2\epsilon_m = 0$$

this is known as the Frölich condition. For noble metals this condition is met within the visible range.

6.3.2 Far-field properties.

6.4 Quantum confinement and photoluminescence in semiconductor quantum dots

Chapter 7

Magnetic properties

7.1 Super-paramagnetism

Chapter 8

Dynamics of electrons and photons

Chapter 9

Confinement of electrons and photons in nanostructured or periodic materials

Chapter 10

Metamaterials and negative-refractive index materials

Chapter 11

Seeing at nanoscale

TODO: transfer to optics section
Optical microscopes use optical photons whose wavelength is between 400 nm and 800 nm. These photons can obviously not provide a resolution good enough to study nanomaterials. Electron microscopes solve this problem, but are quite expensive.

Part XIX

Plasma physics

http://farside.ph.utexas.edu/teaching/plasma/lectures1/index.html Magnetohydrodynamics Alfvén waves

Part XX

Cosmology

16.1 FRW cosmology from Einstein equations. 16.2 (Anti) de Sitter spacetime

- Homogeneity of spacetime → translational invariance → conservation of momentum (conservation of p^μ)
- Isotropicity of spacetime → rotational invariance → conservation of $\mathbf{L}$

Derivation of Friedman-Robertson-Walker metric.
We assume space that is homogeneous and isotropic. Then the metric has the following form:

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + a(t)^2\,\mathrm{d}\sigma^2$$

$\gamma_{ij}\,\mathrm{d}x^i \otimes \mathrm{d}x^j = \mathrm{d}\sigma^2$ 3d maximally symm space

$$\tilde{R}_{ijkl}(g_\sigma) = K(\gamma_{ik}\gamma_{jl} - \gamma_{jk}\gamma_{il}) \qquad K = 0, \pm 1$$

$$\mathrm{d}\sigma^2 = f(r)\,\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega_{S^2}{}^2 \qquad f(r) = \frac{1}{1 - Kr^2}$$

We compute

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{1}{M_p^2}T_{\mu\nu}$$

$$\begin{aligned}\tilde{R}^{ij} &= K\tilde{e}^i \wedge \tilde{e}^j \qquad \mathrm{d}\sigma^2 = \tilde{e}^i \otimes \tilde{e}^j\delta_{ij}\\ &= \frac{1}{2}\tilde{e}^k \wedge \tilde{e}^l R^{ij}_{kl}\end{aligned}$$

$$\mathrm{d}\tilde{e}^i = -\tilde{\omega}^i{}_j \wedge \tilde{e}^j \qquad \tilde{R}^{ij} = \mathrm{d}\tilde{\omega}^{ij} + \tilde{\omega}^i{}_k \wedge \tilde{\omega}^{kj}$$

$$\begin{cases} e^0 = \mathrm{d}t \\ e^i = a(t)\tilde{e}^i \end{cases}$$

$$\mathrm{d}e^a = -\omega^a{}_b \wedge e^b \qquad a = a, i$$

Lecture 18/01
FRW (In flat space)

$$\begin{cases} \mathrm{d}s^2 = -\,\mathrm{d}t^2 + a^2(t)\,\mathrm{d}\sigma^2 \\ \mathrm{d}\sigma^2 = \tilde{e}^i \otimes \tilde{e}^j\delta_{ij} \qquad \tilde{R}^{ij} = K\tilde{e}^i \wedge \tilde{e}^j \qquad i,j = 1,2,3 \end{cases}$$

$$R_{ab} - \frac{1}{2}\eta_{ab}R = \frac{1}{M_p^2}T_{ab}$$

$$e^0 = \mathrm{d}t \qquad e^j = a(t)\tilde{e}^i$$

$$\mathrm{d}\tilde{e}^i + \tilde{\omega}^i{}_j \wedge \tilde{e}^j = 0 \qquad \tilde{R}^{ij} = \mathrm{d}\tilde{\omega}^{ij} + \tilde{\omega}^i{}_k \wedge \tilde{\omega}^{kj}$$

$$\mathrm{d}e^a + \omega^a{}_b \wedge e^b = 0$$

$$\mathrm{d}e^0 = 0 = -\omega^0_i \wedge e^i$$

$$\mathrm{d}e^i = \dot{a}e^0 \wedge \tilde{e^i} - a(t)\tilde{\omega}^i{}_j\tilde{e^j} = \frac{\dot{a}}{a}e^0 \wedge e^i + (?) = -\frac{\dot{a}}{a}e^i \wedge e^0 + (?) = -\omega^i{}_0 \wedge e^0 - \omega^i{}_j \wedge e^j$$

Using ? and ?

$$R^i{}_0 = \mathrm{d}\omega^i_0 + \omega^i{}_j \wedge \omega^j_0$$
$$R^i{}_j = \mathrm{d}\omega^i{}_j + \omega^i{}_0 \wedge \omega^0_j + \omega^i{}_k \wedge \omega^k_j$$
$$R^i{}_j = \mathrm{d}$$

(Missed some stuff)

$$R^i{}_0 = \frac{1}{2} e^a \wedge e^b R_{ab\ 0}^{\ \ i} = \frac{\ddot{a}}{a} e^i \wedge e_0$$
$$R^i_j = \frac{1}{2} e^a \wedge e^b R_{ab\ j}^{\ \ i} = \frac{K + \dot{a}^2}{a^2} e^i \wedge e_j$$
$$R_{j0\ 0}^{\ \ i} = -\delta^i_j \frac{\ddot{a}}{a} + (?)$$

(missed some stuff)
Ricci tensors

$$R_0 0 = -3\frac{\ddot{a}}{a}$$
$$R_{ij} = \delta_{ij} \left(2\frac{K + \dot{a}^2}{a^2} + \frac{\ddot{a}}{a} \right)$$
$$R = 6 \left(\frac{\ddot{a}}{a} + \frac{K + \dot{a}^2}{a^2} \right)$$

Ricci scalar

$$R = \eta^{ab} R_{ab} = ??$$

Something like moving along in same reference frame as fluid

$$T_{\mu\nu} = (p + \rho) u_\mu u_\nu + p g_{\mu\nu}$$

$u^\mu = (1, 0, 0, 0)$

$$\begin{cases} T_{tt} = \rho \\ T_{ij} = p g_{ij} \end{cases} \quad \rightarrow \quad \begin{cases} T_0 0 = \rho \\ T_{ij} = p \delta_{ij} \end{cases}$$

missed more stuff
Result:

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{\rho}{3M_p^2} - \frac{K}{a^2}$$

and

$$\frac{\ddot{a}}{a} = -\frac{1}{6M_p^2} (\rho + 3p)$$

which are equivalent if we use that $D_\mu T^{\mu\nu} = 0$ (... conserved)
Also we call

$$\begin{cases} H = \frac{\dot{a}}{a} \\ \Omega = \frac{\rho}{3M_p^2 H^2} = \frac{\rho}{\rho_{\mathrm{crit}}} \end{cases} \qquad \Omega - 1 = \frac{K}{a^2 H^2}$$

Chapter 1

Cosmography

1.1 The cosmological principle

1.2 Energy budget of the universe

1.2.1 Dark matter

1.2.2 Dark energy

1.3 Hubble's law

1.4 Friedmann models

Chapter 2

Thermal history of the universe

Chapter 3

Stellar evolution

Chapter 4

Formation of cosmic structures

Part XXI

Reference

Appendix A

Symbols

Appendix B

Formula reference

Appendix C

Useful constants

Appendix D

Seminal experiments

Coulomb torsion balance Michelson-Morley Eotvös experiment
Bucket (Newton) Zero point energy
LIGO Colliders + CERN
Star composition from spectroscopy
2-degree-Field Galaxy Redshift Survey 2-Micron All-Sky Survey (2MASS) Bell test BOOMERanG Camera obscura experiments Cavendish experiment Cosmic Background Explorer (COBE) Cowan–Reines neutrino experiment Davisson–Germer Double-slit (Young) Foucault pendulum Franck–Hertz Gravity Probe A Gravity Probe B Geiger–Marsden Homestake experiment Michelson–Morley (+ finite c?) Oil drop experiment Sloan Digital Sky Survey Stern–Gerlach Wilkinson Microwave Anisotropy Probe

Appendix E

Bibliography

https://en.wikipedia.org/wiki/Outline_of_physics

1. Experimental physics and statistics:
2. Newtonian mechanics:
   - Gregory
   - Goldstein
   - https://en.wikipedia.org/wiki/Principle_of_least_action
   - The calculus of variations, Bruce van Brunt
3. Fluid mechanics:
4. Building blocks of the world:
5. Thermodynamics and statistical mechanics:
   - Schroeder
   - Callen
   - The Concepts and Logic of Classical Thermodynamics as a Theory of Heat Engines: Rigorously Constructed upon the Foundation Laid by S. Carnot and F. Reech; Clifford A. Truesdell, Subramanyam Bharatha
   - Cengel, Boles
   - Phenomenological thermodynamics in a nutshell, Arnold Neumaier
   - Concepts in thermal physics / Stephen J. Blundell and Katherine M. Blundell
   - Statistical Mechanics, Kerson Huang
6. Classical electromagnetism:
   - David J. Griffiths, Introduction to electrodynamics
7. Optics:
   - Fundamentals of optics, Jenkins and White
   - The light fantastic, Kenyon

8. Special relativity:
    - David J. Griffiths, Introduction to electrodynamics
    - R. D'Auria and M. Trigiante, From special relativity to Feynman diagrams.
    - https://www.quora.com/Which-19th-century-physicist-famously-said-that-all-tha
9. General relativity:
    - James B. Hartle, Gravity
    - S. Carroll, Spacetime and Geometry: An Introduction to General Relativity. –: Addison-Wesley, 2003
    - M. Blau, Lecture Notes in General Relativity
    - http://casa.colorado.edu/~ajsh/phys5770_08/grtetrad.pdf
10. Quantum mechanics:
    - B.H. Bransden and C.J. Joachain, Quantum mechanics
    - Van Hove, Léon (1958). "Von Neumann's contributions to quantum theory". Bull. Amer. Math. Soc. 64: 95–100. doi:10.1090/s0002-9904-1958-10206-2.
    - SUPPLEMENTARY NOTES ON DIRAC NOTATION, QUANTUM STATES, ETC.; R. L. Jaffe, 1996
    - A short course on Quantum Mechanics and its Geometry, Elisa Ercolessi
    - Zetilli
    - Townsend
    - Strocchi, F.: An Introduction to the Mathematical Structure of Quantum Mechanics:A Short Course for Mathematicians, vol. 28. World Scientific, New Jersey (2008)
    - Quantum Theory - Peres
    - QM - Schwabl
    - http://www.ueltschi.org/teaching/2022-TCC.html
11. Classical field theory:
    - Nastase, Classical field theory
12. Quantum field theory:
    - Perturbative Algebraic Quantum Field Theory, An Introduction for Mathematicians - Kasia Rejzner
    - Klauber, Student friendly QFT
    - Mandl, Shaw
    - Weinberg
    - Quantum Fields and Strings: A Course for Mathematicians
13. Quantum many particle systems:
    - Quantum theory of many-particle systems, Fetter and Walecka
14. Nuclear physics:

- Introductory nuclear physics, Kenneth Krane

15. Solid state physics:

    - Charles Kittel, Introduction to solid state physics
    - John J. Quinn; Kyung-Soo Yi
    - Solid State Physics; Giuseppe Grosso, Giuseppe Pastori Parravicini
    - Ashcroft, Mermin
    - Magnetism and magnetic materials, J.M.D. Coey

16. Soft condensed matter:

    - Thermodynamics of the glassy state, Leuzzi and Nieuwenhuizen
    - Kinetics of phase transitions, Puri and Wadhawan (ed.)

17. Cosmology: