# A Guide to Physics

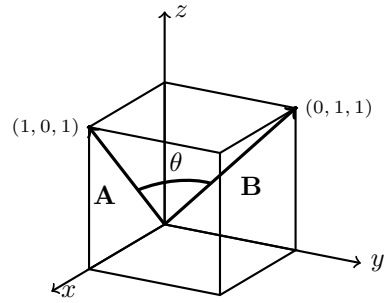
November 17, 2024

# Contents

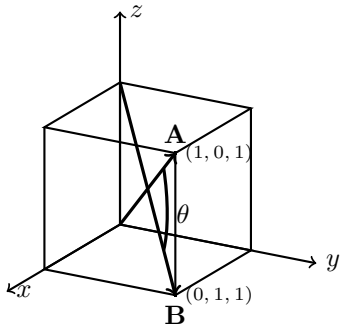
## 0.1 Vector Analysis

### 0.1.1 Vector Algebra

Face diagonals




Body diagonals



## 0.2 Potentials

### 0.2.1 Multipole Expansion

$$\boldsymbol{\mathscr{r}} = |\mathbf{r} - \mathbf{r}'|$$

Where $\boldsymbol{\mathscr{r}}$ denotes the vector from a source point $\mathbf{r}'$ to a field point $\mathbf{r}$.

# Chapter 1

# Introduction to Electrostatics

## 1.1 Charge distributions

### 1.1.1 Surface elements $da$

The following logic comes from [Gri13] chapter 1.4.

Remember that in spherical coordinates we have the **polar angle** $\theta \in [0, \pi]$, which is the angle from the $z$ coordinate to the vector $\mathbf{r}$. And the **azimuthal angle** $\phi$ has the range $\phi \in [0, 2\pi)$, and is the angle from the coordinate $x$ to the line vector $\mathbf{r}$.

In spherical coordinates, an infinitesimal displacement in the $\hat{\mathbf{r}}$ direction is simply $dr$.

$$dl_r = dr$$

An infinitesimal element of length in the $\hat{\boldsymbol{\theta}}$ direction (along the arc from the poles) is

$$dl_\theta = rd\theta$$

And an infinitesimal element of length along the $\hat{\boldsymbol{\phi}}$ direction (along the arc from the equator) is

$$dl_\phi = r\sin\theta d\phi$$

$r\sin\theta$ is the projection of the line vector $\mathbf{r}$ onto the $x-y$ plane. $r\cos\theta$ would be the hight (projection onto the $x$ axis).

The general infinitesimal displacement becomes

$$d\mathbf{l} = dr\hat{\mathbf{r}} + rd\theta\hat{\boldsymbol{\theta}} + r\sin\theta d\phi\hat{\boldsymbol{\phi}}$$

From there one can see how the infinitesimal volume element $d\tau$ is then

$$d\tau = dl_r dl_\theta dl_\phi = r^2\sin\theta dr d\theta d\phi$$

This is a straightforward of justifying the volume element in spherical coordinates. The other way to arrive at it is to use the Jacobian to derive how a volume element morphs when seen changing to spherical coordinates. See 3.0.3.

**However, the inifinitesimal area element depends on the orientation of the surface!**

For example, if we integrate over the surface of a sphere, where r stays contant but $\theta$ and $\phi$ vary,

$$d\mathbf{a} = dl_\theta dl_\phi \hat{\mathbf{r}} = r^2 \sin\theta d\theta d\phi \hat{\mathbf{r}}$$

Where as is the surface lies in the x-y plane, so that $\theta$ is contant, and let's say $\theta = \phi/2$, while $r$ and $\phi$ vary,

$$d\mathbf{a} = dl_r dl_\phi \hat{\boldsymbol{\theta}} = r\sin\theta dr d\phi \hat{\boldsymbol{\theta}} = r dr d\phi \hat{\boldsymbol{\theta}}$$

**The area element on the sphere can be calculated from the cross products of other two elements,**

You should also, definetely, see Integration and differentiation in spherical coordinates.

**Example 1**

One good example to get practice is [Gri13] Example 1.13: find the volume of a sphere of radius $R$.

As we saw above, the optimal way to an answer is to integrate the volume element $d\tau$ in spherical coordinates.

$$\begin{aligned} V = \int d\tau &= \int_{r=0}^{R}\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} dr\, d\theta\, d\phi r^2 \sin\theta \\ &= \left(\int_{r=0}^{R} dr\, r^2\right)\left(\int_{\theta=0}^{\pi} d\theta\, \sin\theta\right)\left(\int_{\phi=0}^{2\pi} d\phi\right) \\ &= \left(\frac{1}{3}R^3\right)\left(-\cos\theta\Big|_0^\pi\right)(2\pi) \\ &= \frac{4}{3}\pi R^3 \end{aligned}$$

### 1.1.2 Fundamental Theorems

This presentation is a quick summary of [Gri13] Sections 1.3.2 to 1.3.5.

The fundamental theorem of calculus states

$$\int_a^b dx \left(\frac{df}{dx}\right) = f(b) - f(a)$$

The fundamental theorem for gradients is

$$\int_{\mathbf{a}}^{\mathbf{b}} (\nabla T) \cdot d\mathbf{l} = T(\mathbf{b}) - T(\mathbf{a})$$

Here $T$ is a scalar-values function, $T : \mathbb{R}^3 \to \mathbb{R}$, and the application of the $\nabla$ operator ($\nabla = \langle \partial_x, \partial_y, \partial_z \rangle$) makes $\nabla T$ a vector-valued function, $\nabla T : \mathbb{R}^3 \to \mathbb{R}^3$. The gradient points in the direction of maximum increase of the function $T$ and it expreses the rate of change of $T$ along a given direction.

The interesting bit of our fundamental theorem for gradients is that $\int_{\mathbf{a}}^{\mathbf{b}} (\nabla T) \cdot d\mathbf{l}$ is path independent. And by consequence, $\oint (\nabla T) \cdot d\mathbf{l} = 0$ (if you go up the stairs, measure how many meters above sea-level you are, then go back down, and come back up, then your change is 0).

The fundamental theorem for divergences (Gauss's of Green's theorem) states that

$$\int_{\mathcal{V}} (\nabla \cdot \mathbf{v}) \; d\tau = \oint_{\mathcal{S}} \mathbf{v} \cdot d\mathbf{a}$$

For the two previous fundamental theorems we saw that the boundary of a line was two points, and now we see that the boundary of a volume $\mathcal{V}$ is the closed region $\mathcal{S}$. Accordingly, if we sum how much quantity $\mathbf{v}$ spreads out over an entire volume, this is equal to how much is $\mathbf{v}$ is crossing through the closed enclosing surface.

The fundamental theorem for curls (Stoke's theorem) states that

$$\int_{\mathcal{S}} (\nabla \times \mathbf{v}) \cdot d\mathbf{a} = \oint_{\mathcal{P}} \mathbf{v} \cdot d\mathbf{l}$$

Now the boundary of a region $\mathcal{S}$ is the perimeter (a line) $\mathcal{P}$. Here we see that the circulation of $\mathbf{v}$, $\oint_{\mathcal{P}} \mathbf{v} \cdot d\mathbf{l}$, over a perimeter, is equal to the sum of the flux of the "twisting" of $\mathbf{v}$ over a surface.

### 1.1.3 Work done by a particle

**Taken from Kline, Chapter 7, Section 7, problem 3**

AM is a straight vertical line. MO is a straight horizontal line. O is to the right of M.

A particle moving along AM is attracted to a fixed point O with a force that varies inversely as the square of the distance from O. Find the work done on the particle as it moves from A to a distance B (between A and M).

Suggestion: let $x$ be the distance from A to any point P on AB and let $r$ be the variable distance from P to O. The force $F$ attracting P to O is $k/r^2$. But only the component of this force along AB serves to move the particle along AB. This component is $(k/r^2) \cos OPM = (k/r^2)(c - x)/r$. Then $dW/dx = k(c - x)/r^3$ and $r = \sqrt{p^2 + (c - x)^2}$.

$$\begin{aligned} W &= \int_A^B \frac{dW}{dx} dx \\ &= \int_A^B \frac{k(c-x)}{r^3} \\ &= \int_A^B \frac{k(c-x)}{\left(p^2+(c-x)^2\right)^{3/2}} dx \end{aligned}$$

The trick here is to note that if we chose $u = p^2 + (c-x)^2$, then $du = -2(c-x)dx$, so that

$$\frac{k(c-x)}{\left(p^2+(c-x)^2\right)^{3/2}} dx \to -\frac{k}{2}\frac{1}{u^{3/2}} du$$

Since $\int u^{-3/2} du = -2u^{-1/2} + C$. So putting everything together,

$$\begin{aligned} W &= \int_A^B \frac{k(c-x)}{\left(p^2+(c-x)^2\right)^{3/2}} dx \\ &= -\frac{k}{2}(-2)\frac{1}{u^{1/2}}\Big|_A^B \\ &= k\left(\frac{1}{\sqrt{p^2+(c-B)^2}} - \frac{1}{\sqrt{p^2+(c-A)^2}}\right) \\ &= \frac{k}{r_B} - \frac{k}{r_A} \end{aligned}$$

**Note:** the two interesting bits to note from this exercise are first, when we were looking for the component of motion that contributed to the work being done, the vertix of the angle used was the point we were "on". Second, note how our the u-substitution we used went, doing it the normal way may confuse you and try to get you for somewhere where to use $u = p^2 + (c-x)^2$, which is not needed.

## 1.2 Gauss's Law

### 1.2.1 Solid Angle

The following is all wikipedia.

Whereas an angle in radians, projected onto a circle, gives a length of a circular arc on the circumference, a solid angle in steradians, projected onto a sphere, gives the area of a spherical cap on the surface.

The point from which the object is viewed is called the apex of the solid angle, and the object is said to subtend its solid angle at that point.

In SI units, a solid angle is expressed in a dimensionless unit called a **steradian**.

One steradian corresponds to one unit of area on the unit sphere surounding the apex, so an object that blocks all rays from the apex would cover a number of steradians equal to the total surface area of the unit sphere, $4\pi$.

Just like a planar angle in radians is the ratio of the length of an arc to its radius, $\theta = s/r$, a solid angle in steradians is the ratio of the area covered on a sphere by an object to the area given by the square of the radius of said sphere. The formula is

$$\Omega = \frac{A}{r^2}$$

where A is the spherical surface area and r is the radius of the considered sphere.

Any area on a sphere which is equal in area to the square of its radius, when observed from its center, subtends precisely one steradian.

The solid angle of a sphere measured from any point in its interior is $4\pi$ sr (steridians), and the solid angle subtended at the center of a cube by one of its faces is one-sixth of that, or $2\pi/3$ sr.

In spherical coordinates there is a formula for the differential,

$$d\Omega = \sin\theta d\theta d\varphi$$

where $\theta$ is the colatitude (angle from the North Pole - z-axis range $[0, \pi]$) and $\varphi$ is the longitude (range $[0, 2\pi]$).

The solid angle for an arbitrary oriented surface S subtended at a point P is equal to the solid angle of the projection of the surface S to the unit sphere with center P, which can be calculated as the surface integral:

$$\Omega = \iint_S \frac{\hat{r} \cdot \hat{n}}{r^2}\, dS = \iint_S \sin\theta\, d\theta\, d\varphi$$

where $\hat{r} = \vec{r}/r$ is the unit vector corresponding to $\vec{r}$, the position vector of an infinitesimal area of surface $dS$ with respect to point P, and where $\hat{n}$ represents the unit normal vector to $dS$. Even if the projection on the unit sphere to the surface S is not isomorphic, the multiple folds are correctly considered according to the surface orientation described by the sign of the scalar product $\hat{r} \cdot \hat{n}$.

Thus one can approximate the solid angle subtended by a small facet having flat surface area $dS$, orientation $\hat{n}$, and distance r from the viewer as:

$$d\Omega = 4\pi \left(\frac{dS}{A}\right)(\hat{r} \cdot \hat{n})$$

where the surface area of a sphere is $A = 4\pi r^2$.

### 1.2.2 Explanation of Jackson

If the electric field makes an angle $\theta$ with the unit normal, then the projection of the infinitesimal area element along the normal of the surface is $\cos\theta da$. Using the solid angle formula, we then get $d\Omega = \cos\theta da/r^2$ or $r^2 d\Omega = \cos\theta da$.

## 1.3 Problems

### 1.3.1 Problem 1.2

The Dirac delta function in three dimensions can be taken as the improper limit as $\alpha \to 0$ of the Gaussian function

$$D(\alpha; x, y, z) = \frac{1}{\left(\alpha\sqrt{2\pi}\right)^3} \exp\left[-\frac{1}{2\alpha^2}\left(x^2 + y^2 + z^2\right)\right]$$

Consider a general orthogonal coordinate system specified by the surfaces $u =$ constant, $v =$ constant, $w =$ constant, with length elements $du/U$, $dv/V$, $dw/W$ in the three perpendicular directions. Show that

$$\delta(\mathbf{x} - \mathbf{x}') = \delta(u - u')\,\delta(v - v')\,\delta(w - w') \cdot UVW$$

by considering the limit of the Gaussian above. Note that as $\alpha \to 0$ only the infinitesimal length element need be used for the distance between the points in the exponent.

Browsing through the web, one will see the following defintions of the delta function,

$$\begin{aligned}
\delta(x) &= \lim_{b\to 0} \frac{1}{|b|\sqrt{\pi}} \exp\left[-\frac{1}{b^2}x^2\right] \\
&= \lim_{b\to 0^+} \frac{1}{2\sqrt{\pi b}} \exp\left[-\frac{1}{4b}x^2\right] \\
&= \lim_{b\to 0} \frac{1}{b\sqrt{2\pi}} \exp\left[-\frac{1}{2b^2}x^2\right]
\end{aligned}$$

The thing to note here is how the arguments of the exponent make it into the constant that's infront of the exponent. And the reason to even look at this now is because we found this amazingly interesting approach to this problem in West Texas CS Baird's site:

We start with

$$D(\alpha; x, y, z) = \frac{1}{\left(\alpha\sqrt{2\pi}\right)^3} \exp\left[-\frac{1}{2\alpha^2}\left(x^2 + y^2 + z^2\right)\right]$$

and then we make a change of variables, a displacement if you will, by changing $x \to x - x'$, etc. This gives us

$$D(\alpha; x - x', y - y', z - z') = \frac{1}{\left(\alpha\sqrt{2\pi}\right)^3} \exp\left[-\frac{1}{2\alpha^2}\left((x - x')^2 + (y - y')^2 + (z - z')^2\right)\right]$$

Now comes the limit taking. First off is whether the limit is some actual number orr just infinity (we are skipping over the question of existence). Since

an exponent grows faster than a ever-increasing number to a power, $\lim_{\alpha\to\infty} \alpha^3 \cdot e^{-\alpha} = 0$, then we would expect our function $D$ to be bounded. But, anyway, if we take the limit as $\alpha \to 0$, and if we look for a case where $D \not\to 0$, then we would want $x - x'$ to "balance out" $\alpha$ just enough so that we have some non-zero bounded value.

Let's assume the above is possible and rewirte $D$ as

$$D\left(\alpha; x-x', y-y', z-z'\right) = \frac{1}{\left(\alpha\sqrt{2\pi}\right)^3} \exp\left[-\frac{1}{2\alpha^2}\left((dx)^2 + (dy)^2 + (dz)^2\right)\right]$$

The infinitesimal displacements in the argument to the exponent look like an arc lenght, so let's rewrite it now as

$$D\left(\alpha; x-x', y-y', z-z'\right) = \frac{1}{\left(\alpha\sqrt{2\pi}\right)^3} \exp\left[-\frac{1}{2\alpha^2}(ds)^2\right]$$

Since an arc length is a concept independent of the coordinate system, we can use our given general orthogonal coordinate system, and rewrite our equation once more,

$$D\left(\alpha; u-u', v-v', w-w'\right) = \frac{1}{\left(\alpha\sqrt{2\pi}\right)^3} \exp\left[-\frac{1}{2\alpha^2}\left(\left(\frac{du}{U}\right)^2 + \left(\frac{dv}{V}\right)^2 + \left(\frac{dw}{W}\right)^2\right)\right]$$

And if we reverse the process, by writing the differentials as differences,

$$D\left(\alpha; u-u', v-v', w-w'\right) = \frac{1}{\left(\alpha\sqrt{2\pi}\right)^3} \exp\left[-\frac{1}{2\alpha^2}\left(\left(\frac{u-u'}{U}\right)^2 + \left(\frac{v-v'}{V}\right)^2 + \left(\frac{w-w'}{W}\right)^2\right)\right]$$

We can now look at the first couple definition of the delta function and make a final change of variahles. For example, if we let $b_1 = \alpha U$, $b_2 = \alpha V$, and $b_3 = \alpha W$,

$$D = \left[\frac{\exp\left[-\frac{\left(u-u'\right)^2}{2b_1^2}\right]}{b_1\sqrt{2\pi}}U\right]\left[\frac{\exp\left[-\frac{\left(v-v'\right)^2}{2b_2^2}\right]}{b_2\sqrt{2\pi}}V\right]\left[\frac{\exp\left[-\frac{\left(w-w'\right)^2}{2b_3^2}\right]}{b_3\sqrt{2\pi}}W\right]$$

If we now take the limit of the $b_i$s, then we get our desired result.

Yet another way to look at this problem is as follows. The general property of the delta function is that

$$\int\int\int \delta(x)\delta(y)\delta(z)dxdydz = 1$$

If we use our limit,

$$\int\int\int \lim_{\alpha\to 0} D(\alpha; x, y, z)dxdydz = 1$$

In our general orthogonal coordinate system the above would look as

$$\int\int\int \lim_{\alpha\to 0} D(\alpha; x, y, z \to u, v, w)\frac{du}{U}\frac{dv}{V}\frac{dw}{W} = 1$$

Note that we do not know how to transform $D$, so we could wrap up our ignorance by defining an intermediate variable as such

$$F(u, v, w) = \lim_{\alpha\to 0} D(\alpha; x, y, z \to u, v, w)\frac{1}{UVW}$$

In which case we have

$$\int\int\int F(u, v, w)dudvdw = 1$$

If we conveniently generalize the above by taking into account the possibility of a translation - because we are integrating over the entire universe anyway - then we have,

$$\int\int\int F(u - u', v - v', w - w')dudvdw = 1$$

Which has the exact same form as the expression that we started with, except we are looking at some other set of integration variables. So we could make the claim that

$$\begin{aligned} F(u - u', v - v', w - w') &= \delta(u - u')\delta(v - v')\delta(w - w') \\ &= \lim_{\alpha\to 0} D(\alpha; x, y, z \to u, v, w)\frac{1}{UVW} \end{aligned}$$

And the most interesting part of this problem is the application!

Le'ts think again about spherical coordinates! The lenght elements along the three orthogonal coordinates are $dr$, $rd\theta$, and $r\sin\theta d\phi$. So $U = 1$, $V = 1/r$, and $W = 1/r\sin\theta$. So a three-dimensional delta function is spherical coordinates becomes

$$\delta\left(\mathbf{x} - \mathbf{x}'\right) = \delta(r - r')\delta(\theta - \theta')\delta(\phi - \phi')\frac{1}{r^2\sin\theta}$$

One interesting way of writing the above is by using $\delta(\cos\theta)$. Since we have the composition of a function as an argument to our delta function, we can use the following formula

$$\delta\left(f(x)\right) = \sum_i \frac{1}{\left|\frac{df}{dx}(x_i)\right|}\delta(x - x_i)$$

where the sum extends over all simple roots.

So,

$$\delta(\cos\theta - \cos\theta') = \sum \frac{1}{|-\sin\theta'|}\delta(\theta - \theta') = \frac{\delta(\theta - \theta')}{\sin\theta'}$$

Thus,

$$\delta\left(\mathbf{x}-\mathbf{x}'\right)=\delta(r-r')\delta(\cos\theta-\cos\theta')\delta(\phi-\phi')\frac{1}{r^2}$$

Now, in cylindrical coordinates, the lenght elements are $dr$, $rd\theta$, and $dz$, so $U=1$, $V=1/r$, and $W=1$. And

$$\delta\left(\mathbf{x}-\mathbf{x}'\right)=\delta(r-r')\delta(\theta-\theta')\delta(z-z')\frac{1}{r}$$

### 1.3.2 Problem 1.3

Using Dirac Delta functions in the appropiate coordinate, express the following charge distributions as three-dimensional charge densities $\rho(\mathbf{x})$.

(a) In spherical coordinates, a charge $Q$ uniformly distributed over a spherical shell of radius $R$.

(b) In cylindrical coordinates, a charge $\lambda$ per unit length uniformly distributed over a cylindrical surface of radius $b$.

(c) In cylindrical coordinates, a charge $Q$ spread uniformly over a flat circular disc of negligible thickness and radius $R$.

(d) The same as part (c), but using spherical coordinates.

# Chapter 2

# Intro to QFT

## 2.1 Inivitation

### 2.1.1 Polarization Vectors

In the invitation, we are presented with the polarization vector $\epsilon^\mu = (0, 1, i, 0)$.

To understand this notation let's go through some examples taken from Lecture 14: Polarization.

We say that a plane wave is **linearly polarized** if there is no phase difference between $E_x$ and $E_y$:

$$\mathbf{E}_0 = (E_x, E_y, 0)$$

Then a linearly polarized plane wave in the x direction looks like $\mathbf{E} = E_0 e^{i(kz-wt)}(1, 0, 0)$. And a linearly polarized plane wave in the y direction looks like $\mathbf{E} = E_0 e^{i(kz-wt)}(0, 1, 0)$.

The thing to remember is that we are implicitly only looking at the real part, so a linearly polarized wave in the x direction is actually $(E_0 \cos(kz - wt), 0, 0)$.

**Circular polarization** is when the electric field components are one-quarter out of phase $(\pi/2)$. Then the field can be written as,

$$\begin{aligned}\mathbf{E}_0 &= \left(E, Ee^{i\pi/2}, 0\right) \\ &= (E, iE, 0) \\ &= E_0 \left(e^{i(kz-wt)}, ie^{i(kz-wt)}, 0\right)\end{aligned}$$

And since we only care about the real parts,

$$\mathbf{E}_0 = (cos(kz - wt), -sin(kz - wt), 0)$$

This is interesting because if you jump to the wikipedia page for "List of trigonometric identities" and look for the "Shift by one quarter period" table,

you'll see that

$$\sin(\theta + \frac{\pi}{2}) = \cos(\theta)$$

and

$$\cos(\theta + \frac{\pi}{2}) = -\sin(\theta)$$

(A shit by a quarter wavelength is essentially differentiation!)

But here is the catch and the connection with P&S: we can also get the same result if we have

$$\mathbf{E} = E_0 \left(\cos(wt - kz)\hat{x} - \sin(wt - kz)\hat{y}\right)$$

We can also write is as,

$$E_0 e^{i(wt-kz)}(0, 1, i, 0)$$

and remembering that you only care about the real part. If you do so, you'll end up with terms such as $\cos\hat{x} - \sin\hat{y}$ which just so happen to again be the a sin and a cos (our original plane wave components) shifted by $\pi/2$.

This polarization vector corresponds to a "**right handed**" field.

To be a "**left handed**" polarization vector, we would write

$$\epsilon^\mu = (0, 1, -i, 0)$$

Corresponding to

$$\mathbf{E} = E_0 \left(\cos(wt - kz)\hat{x} + \sin(wt - kz)\hat{y}\right)$$

J. Binney section 1.4.4 is also a great reference.

### 2.1.2 Cross Sections

The next thing we want to document is how to solve the differential cross section (per unit solid angle). The expression given was

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_{cm}^2}(1 + \cos^2\theta)$$

That when integrated gives the total cross section

$$\sigma_{total} = \frac{4\pi\alpha^2}{3E_{cm}^2}$$

The trick here is to identity $d\Omega = \sin\theta\, d\theta\, d\phi$. So essentially the problem is to integrate

$$\begin{aligned}
\int d\Omega\ (1+\cos\theta) &= \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}(1+\cos\theta)\sin\theta\,d\theta\,d\phi \\
&= \left(\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\sin\theta\,d\theta\,d\phi\right)+\left(\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\cos\theta\sin\theta\,d\theta\,d\phi\right) \\
&= (4\pi)+\left(-\int u^2 du\right) \\
&= (4\pi)+\left(-\frac{1}{3}\cos^3\theta\Big|_0^{\pi}du\right) \\
&= (4\pi)\left(2\pi\frac{8}{3}\right) \\
&= \frac{16\pi}{3}
\end{aligned}$$

Note that we did a $u$-substitution in the second integral: $u = \cos\theta$, so $du = -\sin\theta\,d\theta$.

## 2.2 The Klein-Gordon Field

### 2.2.1 Klein-Gordon Inconsistensies

The Schrodinger equation canr readily be abtain treating the energy and momentum as operators. In quantum mechanics $E = i\partial_t$ and $p = -i\nabla$. Using the realtionship $E = \frac{p^2}{2}$, we get

$$i\partial_t\,|\psi\rangle = -\frac{1}{2m}\nabla^2\,|\psi\rangle$$

Note that we are still in natural units, otherwise there would be an $\hbar$.
The Klein-Gordon equation comes if we instead use $E^2 - p^2 = m^2$,

$$-\partial_t^2\,|\psi\rangle + \nabla^2\,|\psi\rangle = m^2\,|\psi\rangle$$

And remember that $\partial^2 = \partial_t^2 - \nabla^2$, so there comes $(\partial^2 - m)\psi = 0$ equation.

### 2.2.2 A Bit of Formalism: Fourier Transforms

One method to solve the equation is to use Fourier transforms. So we can begin by expressing the wave function $\phi(x,t)$ as a Fourier integral,

$$\psi(x,t) = \int \frac{d^4k}{(2\pi)^4}\tilde{\psi}(k,\omega)e^{-i(k\cdot x-\omega t)}$$

This is actually the **inverse Fourier transform**.

Our convention of the **Fourier transform**, follows [Mic95]:

$$\mathcal{F}\{\psi(x,t)\} = \tilde{\psi}(k,\omega) = \int d^4x\, \psi(x,t) e^{i(k\cdot x - \omega t)}$$

Here, $\tilde{\psi}(k,\omega)$ is the Fourier transform of $\psi(x,t)$. One thing that's often missed is the argument of the exponent: if you look back to examples where we talk about plane waves (go back to 2.1.1), a wave with positive momentum moving in positive direction $x$ has the argument $wt - xk$. The most common convention is to have such a wave in the Fourier transform.

It is also worth stoping here to really take in and mess around with these expressions - we should understand where that factor of $(2\pi)^n$ comes from becuase that will either simplify a ton of expressions or it will get us an answer that's orders or magnitude off.

Let's work in four-dimensional space still but absorb the time time component in our notation (eventhough we have already been doing this in our integration measures).

And consider this, if we take a function, then we compute its Fourier transform, and then we take its inverse Fourier transform, then we should arrive at the original function.

$$\begin{aligned}
\psi(x) &= \int \frac{d^4k}{(2\pi)^4} \tilde{\psi}(k) e^{-ik\cdot x} \\
&= \int \frac{d^4k}{(2\pi)^4} \left( \int d^4x'\, \psi(x') e^{ik\cdot x'} \right) e^{-ik\cdot x} \\
&= \int d^4x'\, \psi(x') \left( \frac{d^4k}{(2\pi)^4} e^{ik\cdot(x'-x)} \right)
\end{aligned}$$

Thus, the only way this all works out is if

$$\int \frac{d^4k}{(2\pi)^4} e^{ik\cdot(x'-x)} = \delta^{(4)}(x'-x)$$

If we try the same exercise but start with a function in momentum space, then do an inverse transform, followed by a Fourier transform, then should again end with the same function. You'll see why it is worth doing this exercise here in a bit. Also note that we are following the convention of ignoring the tilde as we believe that now it will be clear what is going on.

$$\begin{aligned}
\psi(k) &= \int d^4x\, \psi(x) e^{ik\cdot x} \\
&= \int d^4x \left( \int \frac{d^4k'}{(2\pi)^4} \psi(k') e^{-ik'\cdot x} \right) e^{ik\cdot x} \\
&= \int \frac{d^4k'}{(2\pi)^4} \psi(k') \left( \int d^4x\, e^{i(k-k')\cdot x} \right)
\end{aligned}$$

Note that this time we left the $(2\pi)^4$ denominator with the momentum integral in order to be consistent with out convention. That is also why we organized the exponentials to have a leading "+" sign.

Note that in this case, for things to work out, we need

$$\int d^4x\, e^{i(k-k')\cdot x} = (2\pi)^4\, \delta^{(4)}\,(k-k')$$

Which is what [Mic95] call out in the notations and conventions section.

### 2.2.3 A Bit of Formalism: Normalization Factors in Momentum Integrals

We'll dive into a bit for formalism here before proceeding.

The reason we verified the Fourier transofrm above was because one often sees, in the **position representation**, expressions such as

$$\mathcal{I} = \int dx\ |x\rangle\, \langle x|$$

And one could naively plug into their equations the expression $\mathcal{I} = \int dp\ |p\rangle\, \langle p|$. But, the there is an important subtlety in our Fourier transform convention!

If you recall from QM, $|\psi\rangle$ is meant to represent some generic state of a system we are interested about. This **state vector** $|\psi\rangle$ represents the possible quantum amplitudes for some particular complete basis set in some particular coordinate system. See [Jam14], chapters 1.2 and 1.4 on pages 9 and 15.

This way of defining the state vector $|\psi\rangle$ should ring a bell as it implicitly calls out to the **spectral theorem** for Hermitian operators. Remember that the spectral theorem for spectral theorem states that the eigenvectors of a hermitian operator form a complete, orthonormal set of basis sets.

Because any hermitian operator will provide for us a complete basis set, $|\psi\rangle$ as a generic state of a system is a good abstraction. And because of that, we use it to define the **wave function**, which gives us the amplitude to find our system in a specific state.

For example, the wave function $\psi(x)$ is defined as:

$$\psi(x) = \langle x|\psi\rangle$$

This wave function, these amplitudes, can help us find the probability to find our system at position $x \in (-\infty, \infty)$ and thus we can express our generic state in terms of the complete basis set $|x\rangle$:

$$|\psi\rangle = \int d^dx\, \psi(x)\, |x\rangle = \int d^dx\ \langle x|\psi\rangle\, |x\rangle$$

Another interesting wave function - another set of amplitudes - that we will be interested on correspond to the amplitdues that a system with a definite

momentum $p$ is found at $x$. We can define this **position-representation of a momentum eigenstate** wave function as:

$$u_p(x) = \langle x|p\rangle$$

We are using this notation that comes from [Jam14] because it helps us clearly see what a wave function is as a set of amplitudes corresponding to the probability of a measurement but also as a coordinate system. Where in the position representation we denote $\psi(x)$ and $u_p(x)$ as functions of $x$ both defined by having a **dual vector**, $\langle x|$ act on a vector. Which is a great remainder that dual vectors are complex-valued (linear) functions whose domain is some **vector space**.

The reason this brief detour matters is because the wave function $u_p(x)$ appears a lot. It appears so much that we actually begun talking about the Fourier transform which has us work out our integrals in momentum space.

And the thing to know about these wave functions with definite momentum is that they carry a normalization factor of $(2\pi)^{-d/2}$ per wave function, and there are always two (because these are amplitdues). So in QFT we usually absorve this $(2\pi)^{-d}$ factor into the Fourier transofmr instead of carrying around with the wave function. Specially since it isn't obvious, when using natural units, that this factor should be around. Let's remember why it is so.

When you see the integrals in quantum mechanics, the calculation often utilizes the well-known expressions for the overlap of position and momentum eigenstates $\langle \mathbf{x}|\mathbf{p}\rangle$ (amplitude to find at $x$ a particle with well-defined momentum) and $\langle \mathbf{p}|\mathbf{x_0}\rangle$ (amplitude to fina particle at $x_0$ with momentum $p$). These expressions are crucial in converting between position and momentum representations, which allows us to analyze the system's dynamics.

You may be familiar with the solution to the differential equation

$$\langle x|\hat{p}|p\rangle = -i\hbar\partial_x u_p(x) = pu_p(x)$$

see [Jam14] chapter 2.3.2, page 38. The solutions are,

$$\langle \mathbf{x}|\mathbf{p}\rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{x}}$$

and

$$\langle \mathbf{p}|\mathbf{x_0}\rangle = \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{p}\cdot\mathbf{x_0}}$$

As we mention above, Pesking and schroeder use

$$\langle \mathbf{x}|\mathbf{p}\rangle = e^{i\mathbf{p}\cdot\mathbf{x}}$$

and

$$\langle \mathbf{p}|\mathbf{x_0}\rangle = e^{-i\mathbf{p}\cdot\mathbf{x_0}}$$

It'd be rough to remember to find the missing normalization factor when using natural units, people just add it to the momentum integrals, where they should always be.

### 2.2.4 A Bit of Formalism: Bras and the Dual Space

First, this comes from the section "outer products" in Wikipedia: Bra-Ket Notation, we have to keep in mind that $|\psi\rangle\langle\psi|$ defines an **outer product**. In a finite-dimensional vector space the outer product is defined as

$$|\phi\rangle\langle\psi| = \begin{pmatrix}\phi_1\\ \phi_2\\ \vdots\\ \phi_N\end{pmatrix}\begin{pmatrix}\psi_1^* & \psi_2^* & \dots & \psi_N^*\end{pmatrix} = \begin{pmatrix}\phi_1\psi_1^* & \phi_1\psi_2^* & \dots & \phi_1\psi_N^*\\ \phi_2\psi_1^* & \phi_2\psi_2^* & \dots & \phi_2\psi_N^*\\ \vdots & \vdots & \ddots & \vdots\\ \phi_N\psi_1^* & \phi_N\psi_2^* & \dots & \phi_N\psi_N^*\end{pmatrix}$$

One of the uses of the outer product is to construct **projection operators**.

Given a ket $|\psi\rangle$ of norm 1 (being a complete basis set for some space), the orthogonal projection onto the subspace spanned by $|\psi\rangle$ is $|\psi\rangle\langle\psi|$.

Very fancy way of saying that the outer product is a dual vector being used to define an operator. Which may be weird given that thus far we've only mentioned that dual vectors are complex-valued linear functions that feed on kets and spit out complex numbers, see 2.2.3, but in this presentation, the dual vector and the vector waiting for another vector to show up. And when that vector does, the dual vector will appear, convert one of the vectors into a complex number and scale the remaining ket.

The "Unit operator" section on Wikipedia: Bra-Ket Notation, also has this: if we have a complete orthonormal basis $\{e_i | i \in \mathbb{N}\}$, functional analysis tells us that any $|\psi\rangle$ can also be written as

$$|\psi\rangle = \sum_{i\in\mathbb{N}} \langle e_i|\psi\rangle\,|e_i\rangle \tag{2.2.1}$$

This is how we "project" $\psi$ into a new basis.

It mentions that it can also be shown that

$$\mathbb{I} = \sum_{i\in\mathbb{N}} |e_i\rangle\langle\psi|$$

There is also this result called resolution of the identity in Borel functional calculus that allows us to generalized this result to the continuous case,

$$\mathbb{I} = \int dx\ |x\rangle\langle x| = \int dp\ |p\rangle\langle p|$$

The analogous to 2.2.1 in the continuous case is

$$|\psi\rangle = \int dx\,\psi(x)\,|x\rangle$$

where $\psi(x) = \langle x|\psi\rangle$ by analogy.

### 2.2.5 Klein-Gordon Solution

With this Fourier transofrm in place, we can take the Fourier transform of our entire Klein-Gordin equation to reduce it to an algebraic problem.

Technically, what we will do is take the inverse Fourier transform of the first term in the KG equation, $\partial_t\psi$. This comes in handy and you will see why at the end.

The physical reason to go about it this way is because we want to reconstruct a function from its frequency components (inverse transform).

$$\begin{aligned}
\mathcal{F}^{-1}\{\partial_t\psi\} &= \int \frac{d^4k}{(2\pi)^4}\partial_t\psi e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \rightarrow \begin{bmatrix} u = e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} & v = \psi \\ du = i\omega e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,dt & dv = \partial_t\psi\,dt \end{bmatrix} \\
&= \cancelto{0}{\psi(\mathbf{x},t)e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}}\Big|_{t=-\infty}^{t=\infty} - \int \frac{d^4k}{(2\pi)^4}\psi(\mathbf{x},t)\,(i\omega)\,e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \\
&= -i\omega\tilde{\psi}(\mathbf{k},\omega)
\end{aligned}$$

We made use of the integration by parts technique ($\int u\,dv = uv - \int v\,du$) to transfer the time derivative on $\psi$ to the exponent term.

As per the boudnary term, we have a couple words to say: physically, the wave function (or even field) should vanish at infinity. Mathematically, in order for the Fourier transform to converge to a value the function that is being "transformed", $\psi$, must be a **rapidly decreasing function** (a function of Schwartz space).

The forwared Fourier transform would be (decomposing the function into its frequency components):

$$\begin{aligned}
\mathcal{F}\{\partial_t\psi\} &= \int d^4x\,\partial_t\psi e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \rightarrow \begin{bmatrix} u = e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} & v = \psi \\ du = -i\omega e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,dt & dv = \partial_t\psi\,dt \end{bmatrix} \\
&= \cancelto{0}{\psi(\mathbf{x},t)e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}}\Big|_{t=-\infty}^{t=\infty} - \int d^4x\,\psi(\mathbf{x},t)\,(-i\omega)\,e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \\
&= i\omega\tilde{\psi}(\mathbf{k},\omega)
\end{aligned}$$

Similarly then,

$$\begin{aligned}
\mathcal{F}^{-1}\{\partial_t^2\psi\} &= \int \frac{d^4k}{(2\pi)^4}\partial_t^2\psi e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \rightarrow \begin{bmatrix} u = e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} & v = \partial_t\psi \\ du = i\omega e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,dt & dv = \partial_t^2\psi\,dt \end{bmatrix} \\
&= \cancelto{0}{\partial_t\psi(\mathbf{x},t)e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}}\Big|_{t=-\infty}^{t=\infty} - \int \frac{d^4k}{(2\pi)^4}\partial_t\psi(\mathbf{x},t)\,(i\omega)\,e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \\
&= -i\omega\mathcal{F}^{-1}\{\partial_t\psi\} \\
&= -\omega^2\tilde{\psi}(\mathbf{k},\omega)
\end{aligned}$$

The boundary term again goes to zero here. We will skip the physical argument and just mention that a rapidly decreasin function also requires all of its derivatives tend to zero.

For the spatial derivatives, each $\partial_i$ will bring down a factor of $-ik_i$, resulting in a factor of $-k_i^2$ for each $\partial_i^2$. And this is something that always triped me up: one may be inclined to write $-k^2$ in the Fourier transform but that same one ought to remember that $k$ is actually $\vec{k}$ (a vector) and that is why we must write $-|k|^2$, because we want the [L-2] norm.

But anyways, the Fourier transform of the Laplacian is,

$$\begin{aligned}
\mathcal{F}^{-1}\{\nabla\psi\} &= \int \frac{d^4k}{(2\pi)^4}\nabla\psi e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \rightarrow \begin{bmatrix} u = e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} & v = \psi \\ du = -i\mathbf{k}e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,d\mathbf{x} & dv = \nabla\psi\,d\mathbf{x} \end{bmatrix} \\
&= \cancelto{0}{\psi(\mathbf{x},t)e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}}\Big|_{\mathbf{x}=-\infty}^{\mathbf{x}=\infty} - \int \frac{d^4k}{(2\pi)^4}\psi(\mathbf{x},t)\,(-i\mathbf{k})\,e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \\
&= -i\mathbf{k}\tilde{\psi}(\mathbf{k},\omega)
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{F}^{-1}\{\nabla^2\psi\} &= \int \frac{d^4k}{(2\pi)^4}\nabla^2\psi e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \rightarrow \begin{bmatrix} u = e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} & v = \nabla\psi \\ du = -i\mathbf{k}e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,d\mathbf{x} & dv = \nabla^2\psi\,d\mathbf{x} \end{bmatrix} \\
&= \cancelto{0}{\nabla\psi(\mathbf{x},t)e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}}\Big|_{\mathbf{x}=-\infty}^{\mathbf{x}=\infty} - \int \frac{d^4k}{(2\pi)^4}\nabla\psi(\mathbf{x},t)\,(-i\mathbf{k})\,e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \\
&= -i\mathbf{k}\mathcal{F}^{-1}\{\nabla\psi\} \\
&= -|\mathbf{k}|^2\tilde{\psi}(\mathbf{k},\omega)
\end{aligned}$$

And since the above must hold throughout all of space, that's where

$$\left(-\omega^2 + |\mathbf{k}|^2 + m^2\right)\tilde{\psi}(\mathbf{k},\omega) = 0$$

comes from!

Now, in order for us to not have a trivial solution, $\psi(\mathbf{x},t) = 0$, it must be so that $\left(-\omega^2 + |\mathbf{k}|^2 + m^2\right) = 0$, and so our dispersion relation comes about

$$\omega^2 = |\mathbf{k}|^2 + m^2$$

So $\omega = \pm\sqrt{|\mathbf{k}|^2 + m^2}$. So essentially any function will do as long as the dispersion relation (momentum conservation) is respected.

A thing that people do, specially since we are talking about Fourier transforms is to define

$$\psi(\mathbf{k},\omega) = A(\vec{k})\delta\left(\omega - \sqrt{|\mathbf{k}|^2 + m^2}\right) + B(\vec{k})\delta\left(\omega + \sqrt{|\mathbf{k}|^2 + m^2}\right)$$

Note that $A(\vec{k})$ and $B(\vec{k})$ are arbitrary functions that only depend on $\mathbf{k}$ since the Dirac delta functions specifies the value for $\omega$.

With that, we can finally write a solution for the differential equation we started with,

$$\begin{aligned}\psi(\mathbf{x},t) &= \int \frac{d^4k}{(2\pi)^4}\left[A(\vec{k})\delta\left(\omega-\sqrt{|\mathbf{k}|^2+m^2}\right)e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}+B(\vec{k})\delta\left(\omega+\sqrt{|\mathbf{k}|^2+m^2}\right)e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\right]\\ &= \int \frac{d^4k}{(2\pi)^4}\left[A(\vec{k})e^{-i(\mathbf{k}\cdot\mathbf{x}-\sqrt{|\mathbf{k}|^2+m^2}t)}+B(\vec{k})e^{-i(\mathbf{k}\cdot\mathbf{x}+\sqrt{|\mathbf{k}|^2+m^2}t)}\right]\end{aligned}$$

Hopefully thiis result makes sense: we obtained a family of function as solution and only initial conditions or boundary conditions will result in a specific sort of function.

#### 2.2.5.1 Klein-Gordon Negative Density

The expression to compute a probability current is derived by using the analogous continuity equation from fluid dynamics

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{j}=0$$

Using the Schrodinger equation and using the fact that $\rho=|\psi|^2=\psi^*\psi$ one can arrive at the probability density

$$\rho=-\frac{i}{2m}\left(\psi^*\partial_t\psi-\psi\partial_t\psi^*\right)$$

For the Klein Gordon equation, the way to massage it and get an expression for $\partial_t\rho$ and for its probability current is to: take Klein-Gordon multiply it by $\psi^*$, take the complex conjugate multiply it by $\psi$, subtract the two and rearange to get something like the continuity equation.

Following those steps we have,

$$\left(\partial_t^2-\nabla^2+m^2\right)\psi=0$$

and

$$\left(\partial_t^2-\nabla^2+m^2\right)\psi^*=0$$

Multiplying them with $\psi^*$ and $\psi$ respectively we get,

$$\psi^*\partial_t^2\psi-\psi^*\nabla^2\psi+\psi^*m^2\psi=0$$

and

$$\psi\partial_t^2\psi^*-\psi\nabla^2\psi^*+\psi m^2\psi^*=0$$

Subtracting the latter from the former,

$$\begin{aligned}
&\psi^*\partial_t^2\psi - \psi^*\nabla^2\psi + \psi^* m^2\psi - \psi\partial_t^2\psi^* + \psi\nabla^2\psi^* - \psi m^2\psi^* \\
&= \left(\psi^*\partial_t^2\psi - \psi\partial_t^2\psi^*\right) - \left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) + \left(\psi^* m^2\psi - \psi m^2\psi^*\right) \\
&= \left(\psi^*\partial_t^2\psi - \psi\partial_t^2\psi^*\right) - \left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) \\
&= \partial_t\left(\psi^*\partial_t\psi - \psi\partial_t\psi^*\right) - \nabla\cdot\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)
\end{aligned}$$

The time derivative is then equated to $\partial_t\rho$ and the spatial derivatives to the current. though one interesting tangent to take here is concerning the missing $\frac{i}{2m}$ factor that these equations have.

If you remember, the imaginery part of of a complex number can be isolated by taking the different of it with its complex conjugate: $\operatorname{Im} z = \frac{1}{2i}(z-\overline{z})$. (Recall that if $z = a + ib$, then $\operatorname{Im} z = b$, not $ib$.) It just so happens that we are doing the same sort of operation here, so we can throw an $i/2$ factor into our equation to ensure we get a real quantity.

For the $1/m$ factor we have to do some dimensional analysis. To get the probability we have to integrate $\int d^3x\,|\psi|^2$, which is dimensionless. So the $|\psi|^2$ term needs to cancel out the integration over space. Hence $[|\psi|^2] = [L]^{-3} = [M]^3$. From there we can say that $[\psi] = [L]^{-3/2}$ and $[\psi^*\partial_t^2\psi] = [L]^{-3/2}[L]^{-2}[L]^{-3/2} = [L]^{-5} = [M]^5$.

This same term we just evaluated needs to match the dimensions of the time derivative, $[L]^{-1}$,a probability density, $[L]^{-3}$. It should be then that $[\partial_t\rho] = [L]^{-4} = [M^4]$. But hey, what a coincidence that if we were to divide $\psi^*\partial_t^2\psi$ by a mass that we would get just the right dimensions for our probability current! And so it goes that the expression we are used to seeing turns out to be excused as such in the Klein-Gordon case.

Suppose we had a $\psi = e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$ as solution.
Then,

$$\begin{aligned}
\rho &= -\frac{i}{2m}\left(e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\partial_t e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} - e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\partial_t e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\right) \\
&= -\frac{i}{2m}\left(e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,(i\omega)\,e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} - e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\,(-i\omega)\,e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\right) \\
&= -\frac{i}{2m}\,(i\omega)\,(2) \\
&= \frac{\omega}{m}
\end{aligned}$$

And since $\omega = \pm\sqrt{|\mathbf{k}|^+ m^2}$, then the density can be negative!

### 2.2.6 Causality Arguments: The Non-relativistic Case

Consider the amplitude for a free particle to propagate from $\mathbf{x}_0$ to $\mathbf{x}$

$$U(t) = \langle\mathbf{x}|\,e^{-iHt}\,|\mathbf{x}_0\rangle$$

In nonrelativistic quantum mechanics we have $E = \frac{1}{2m}p^2$, so

$$U(t) = \langle \mathbf{x}| \, e^{-i(\mathbf{p}^2/2m)t} \, |\mathbf{x}_0\rangle \tag{2.2.2}$$

$$= \int \frac{d^3p}{(2\pi)^3} \, \langle \mathbf{x}| \, e^{-i(\mathbf{p}^2/2m)t} \, |\mathbf{p}\rangle \, \langle \mathbf{p}|\mathbf{x}_0\rangle \tag{2.2.3}$$

$$= \int \frac{d^3p}{(2\pi)^3} e^{-i(\mathbf{p}^2/2m)t} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)} \tag{2.2.4}$$

$$= \left(\frac{m}{2\pi i t}\right)^{3/2} e^{im(\mathbf{x}-\mathbf{x}_0)^2/2t} \tag{2.2.5}$$

To go from 2.2.2 to 2.2.3, we changed the basis to momentum eigenstates in order to apply the Hamiltonian operator and "extract" the exponential. If we have the operator $\hat{p}$ act on a pure momentum eigenstate $|p\rangle$, then we get $\hat{p}\,|p\rangle = p\,|p\rangle$. So changing basis helps us work with real quantities (an eigenvalue obtained by applying a Hermitian operator on its eigenket).

Note that one may want to use the identity $\mathbb{I} = \int dp \; |p\rangle \, \langle p|$. But as we saw in 2.2.2 and in 2.2.3, in order to stay consistent, we need our integrals over momentum space to have denominators of $(2\pi)^d$. So the actual identity we should use, which is what Peskin and Schroeder use, is identity $\mathbb{I} = \int \frac{d^dp}{(2\pi)^d} \, |p\rangle \, \langle p|$. The additional $1/(2\pi)^d$ factor picks up the normalization for factors such as $\langle p|x_0\rangle$.

When you see the integrals over states in quantum mechanics, especially when involving time evolution and position states, the calculation often utilizes the well-known expressions for the overlap of position and momentum eigenstates $\langle \mathbf{x}|\mathbf{p}\rangle$ and $\langle \mathbf{p}|\mathbf{x_0}\rangle$. These expressions are crucial in converting between position and momentum representations, which allows us to analyze the system's dynamics. In such cases, we have these useful identities:

$$\langle \mathbf{x}|\mathbf{p}\rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{x}}$$

and

$$\langle \mathbf{p}|\mathbf{x_0}\rangle = \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{p}\cdot\mathbf{x_0}}$$

But in QFT we instead use:

$$\langle \mathbf{x}|\mathbf{p}\rangle = e^{i\mathbf{p}\cdot\mathbf{x}}$$

and

$$\langle \mathbf{p}|\mathbf{x_0}\rangle = e^{-i\mathbf{p}\cdot\mathbf{x_0}}$$

This way, it is not the projection into momentum space that carries the appropiate normalization factor but the Fourier integrals. Which is a bit more natural when working with natural units.

We used these identities to go from 2.2.3 to 2.2.4,

$$\int \frac{d^3p}{(2\pi)^3} e^{-i(\mathbf{p}^2/2m)t} \langle \mathbf{x}|\mathbf{p}\rangle \langle \mathbf{p}|\mathbf{x}_0\rangle$$
$$= \int \frac{d^3p}{(2\pi)^3} e^{-i(\mathbf{p}^2/2m)t} \frac{1}{(2\pi)^3} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)}$$

As per the solution, see 2.3.4. The trick is to complete the square for the arguments of the exponentials and extract a Gaussian integral.

### 2.2.7 Causality Arguments: The Relativistic Case

And now let's look at

$$\begin{aligned} U(t) &= \langle \mathbf{x}|\, e^{-it\sqrt{\mathbf{p}^2+m^2}}\, |\mathbf{x}_0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3}\, \langle \mathbf{x}|\, e^{-it\sqrt{\mathbf{p}^2+m^2}}\, |\mathbf{p}\rangle \langle \mathbf{p}|\mathbf{x}_0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3}\, e^{-it\sqrt{\mathbf{p}^2+m^2}} \langle \mathbf{x}|\mathbf{p}\rangle \langle \mathbf{p}|\mathbf{x}_0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3}\, e^{-it\sqrt{\mathbf{p}^2+m^2}} e^{i\mathbf{p}\cdot\mathbf{x}} e^{-i\mathbf{p}\cdot\mathbf{x}_0} \\ &= \int \frac{d^3p}{(2\pi)^3}\, e^{-it\sqrt{\mathbf{p}^2+m^2}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)} \end{aligned}$$

The result then happens to be

$$\frac{1}{2\pi^2|\mathbf{x}-\mathbf{x}_0|} \int_0^\infty dp\, p \sin(p|\mathbf{x}-\mathbf{x}_0|) e^{-it\sqrt{\mathbf{p}^2+m^2}}$$

The trick for this one is to use spherical variables to simplify it a bit.

$$\int \frac{d^3p}{(2\pi)^3}\, e^{-it\sqrt{\mathbf{p}^2+m^2}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)}$$

becomes

$$\int \frac{dp}{(2\pi)^3} d\Omega\, p^2 e^{-it\sqrt{p^2+m^2}} e^{ip\mathbf{r}\cos\theta}$$

where $r = |\mathbf{x}-\mathbf{x}_0|$ and $p = |\mathbf{p}|$.

From here we can do the angular integrals first,

$$\begin{aligned}\int d\Omega\, e^{ipr\cos\theta} &= \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\, \sin\theta e^{ipr\cos\theta} \\ &= 2\pi \int_0^{\pi} d\theta\, \sin\theta e^{ipr\cos\theta} \\ &= -2\pi \int_1^{-1} du\, e^{ipru} \\ &= -2\pi \left(\frac{e^{ipru}}{ipr}\right)\Big|_{u=1}^{u=-1} \\ &= -\frac{2\pi}{ipr}\left(e^{-ipr} - e^{ipr}\right)\end{aligned}$$

Next, we can make use of the formula $\operatorname{Im} z = \frac{z-\overline{z}}{2i}$ and apply it to $z = e^{ix}$. Or just do it manually: $e^{ix} = \cos + i\sin$, $e^{-ix} = \cos - i\sin$, so $e^{-ix} - e^{ix} = -i\sin - i\sin = -2i\sin$.

$$\begin{aligned}\int d\Omega\, e^{ipr\cos\theta} &= -\frac{2\pi}{ipr}\left(e^{-ipr} - e^{ipr}\right) \\ &= -\frac{2\pi}{ipr}\left(-2i\sin(pr)\right) \\ &= \frac{4\pi}{pr}\sin(pr)\end{aligned}$$

Pluging this back in, We got

$$\int \frac{dp}{(2\pi)^3} d\Omega\, |p|^2 e^{-it\sqrt{p^2+m^2}} e^{ipr\cos\theta} = \int \frac{dp}{(2\pi)^3} p\frac{4\pi}{r}\sin(pr) e^{-it\sqrt{p^2+m^2}} \quad (2.2.6)$$

$$= \frac{1}{2\pi^2|\mathbf{x}-\mathbf{x}_0|}\int_0^{\infty} dp\, p\sin(p|\mathbf{x}-\mathbf{x}_0|) e^{-it\sqrt{p^2+m^2}} \quad (2.2.7)$$

The reference to Gradsheteyn and Ryzhik to 3.914 lists the following: on page 491, item 6 has

$$\int_0^{\infty} dx\, x\sin(bx) e^{-\beta\sqrt{x^2+\gamma^2}} = \frac{b\beta\gamma^2}{\beta^2+b^2} K_2\left(\gamma\sqrt{b^2+\beta^2}\right) \quad (2.2.8)$$

ET I 175(35). I refers to volume 1 of the reference, 175 is the page where it should be found, (35) refers to the number of the formula in that work.

We found a PDF of ET and on said place we found ourselves in the Laplace transforms section. This section has two columns, one of them is $f(t)$ and the other one is $g(p) = \int_0^{\infty} dt\, f(t) d^{-pt}$. Forumal 35 had on the first column

$$t^{\alpha} L_n^{\alpha}(\lambda t) L_m^{\alpha}(kt)$$

for $\operatorname{Re}\alpha > -1$.

On the second column was the following:

$$\frac{\Gamma(m+n+\alpha+1)}{m!n!}\frac{(p-\lambda)^n(p-k)^m}{p^{m+n+\alpha+1}}{}_2F_1\left[-m,-n;-m-n-\alpha;\frac{p(p-\lambda-k)}{(p-\lambda)(p-k)}\right]$$

When $\operatorname{Re} p > 0$.

Anyway, backing up a bit, $K_\nu(z)$ is a Bessel function for an imaginary argument. Which is defined in 8.407 and in 8.43, on page 911. Gradsheteyn and Ryzhik have the following definitions. If $-\pi < \arg z \leq \frac{1}{2}\pi$, then

$$K_\nu(z) = \frac{i\pi}{2}e^{i\pi\nu/2}H_\nu^{(1)}\left(ze^{\frac{1}{2}i\pi}\right)$$

If $-\frac{1}{2}\pi < \arg z \leq \pi$, then

$$K_\nu(z) = -\frac{i\pi}{2}e^{-i\pi\nu/2}H_{-\nu}^{(2)}\left(ze^{-\frac{1}{2}i\pi}\right)$$

$H_\nu^{(1)}(z)$ and $H_\nu^{(2)}(z)$ are Bessel function of the third kind, or Hankel functions. Which are defined in terms of the Bessel functions of the first kind, $J_\nu(z)$, and on Bessel functions of the second kind, $Y_\nu(z)$ (also called Neumman functions and written as $N_\nu(z)$).

$$H_\nu^{(1)}(z) = J_\nu(z) + iY_\nu(z)$$
$$H_\nu^{(2)}(z) = J_\nu(z) - iY_\nu(z)$$

Comparing 2.2.7 with 2.2.8,

- $x \to p$
- $b \to |x - x_0|$
- $\beta \to it$
- $\gamma \to m$

So

$$\frac{b\beta\gamma^2}{\beta^2+b^2}K_2\left(\gamma\sqrt{b^2+\beta^2}\right)$$

becomes

$$\frac{itm^2|x-x_0|}{-t^2+|x-x_0|^2}K_2\left(m\sqrt{-t^2+|x-x_0|^2}\right)$$

Now let's look at the modified function $K_2$,

$$\begin{aligned}K_2\left(m\sqrt{-t^2+|x-x_0|^2}\right) &= \frac{i\pi}{2}e^{i\pi}H_2^{(1)}\left(m\sqrt{-t^2+|x-x_0|^2}e^{\frac{1}{2}i\pi}\right)\\ &= -\frac{i\pi}{2}H_2^{(1)}\left(m\sqrt{-t^2+|x-x_0|^2}\right)\end{aligned}$$

Looking at Gradsheteyn and Ryzhik again,

$$H_2^{(1)}\left(m\sqrt{-t^2+|x-x_0|^2}\right) = J_2\left(m\sqrt{-t^2+|x-x_0|^2}\right)+iY_2\left(m\sqrt{-t^2+|x-x_0|^2}\right)$$

And here we stop and appreciate why the method of stationary phase was mentioned and use by the authors.

### 2.2.8 On The Way To The Method of Stationary Phase: Laplace's Method

$$\mathcal{I}(\lambda) = \int_a^b dx\, f(x)e^{i\lambda S(x)}$$

When $\lambda$ is large, the phase factor, the exponential, oscillates very rapidly, except near the stationary points where $S'(x) = 0$.

Let's call a stationary point $x_0$. A Taylor expansion around it would be

$$S(x) = S(x_0) + S'(x_0)(x-x_0) + \frac{1}{2!}S''(x_0)(x-x_0)^2 + \dots$$

The linear term vanishes since we are in a stationary point. Substituting this approximation into our integral, we get

$$\begin{aligned}
\mathcal{I}(\lambda) &\approx \int_a^b dx\, f(x) \exp\left(i\lambda\left(S(x_0) + \frac{1}{2}S''(x_0)(x-x_0)^2\right)\right) \\
&= \int_a^b dx\, f(x)e^{i\lambda S(x_0)} \exp\left(\frac{i\lambda}{2}S''(x_0)(x-x_0)^2\right) \\
&= e^{i\lambda S(x_0)} \int_a^b dx\, f(x) \exp\left(\frac{i\lambda}{2}S''(x_0)(x-x_0)^2\right)
\end{aligned}$$

Another trick that is done is to expand the limits of integration to $\pm\infty$. This can be justified because away from the stationary point, the phase factor oscillations will cancel each other out.

So we now work on

$$\mathcal{I}(\lambda) \approx e^{i\lambda S(x_0)} \int_{-\infty}^{\infty} dx\, f(x_0) \exp\left(\frac{i\lambda}{2}S''(x_0)(x-x_0)^2\right)$$

Note that since we are only interested in the neighborhood of the stationary point we are also approximating $f(x) \approx f(x_0)$.

This integral is a complex Gaussian, see 2.3.2. The general solution is

$$\int e^{-a(t+b)^2} dt = \sqrt{\frac{\pi}{a}}$$

Here $a = -i\frac{\lambda}{2}S''(x_0)$. So

$$\mathcal{I}(\lambda) \approx e^{i\lambda S(x_0)} f(x_0) \sqrt{\frac{2\pi}{-i\lambda S''(x_0)}} \tag{2.2.9}$$

This tells us that that for large $\lambda$, the integral is dominated by contributions near the stationary points, with amplitude decaying as $\lambda^{-1/2}$.

As for why the phase factor oscillates less rapidly near stationary points, we need to look at the phase factor $e^{i\lambda S(x_0)}$ (the oscillating part!).

The rate of oscillation is determined by how quickly the phase $\lambda S(x)$ changes with $x$. This rate of change is

$$\frac{d}{dx}\left(\lambda S(x)\right) = \lambda S'(x)$$

Since $S'(x_0) = 0$, then the phase factor is not changing.

This is why the main contribution to the integral comes from near the stationary points - everywhere else, the rapid oscillations tend to cancel each other out when integrated.

Looking back at 2.2.7,

$$\begin{aligned} U(t) &= \langle \mathbf{x}| \, e^{-iHt} \, |\mathbf{x}_0\rangle && (2.2.10)\\ &= \int \frac{dp}{(2\pi)^3} d\Omega \, |p|^2 e^{-it\sqrt{p^2+m^2}} e^{ipr\cos\theta} && (2.2.11)\\ &= \frac{1}{2\pi^2|\mathbf{x}-\mathbf{x}_0|} \int_0^\infty dp \, p \sin(p|\mathbf{x}-\mathbf{x}_0|) e^{-it\sqrt{p^2+m^2}} && (2.2.12) \end{aligned}$$

Let's defined $x = |\mathbf{x} - \mathbf{x}_0|$. And remeber that $\sin(px) = \frac{1}{2i}\left(e^{ipx} - e^{-ipx}\right)$. So we have,

$$\begin{aligned} U(t) &= \langle \mathbf{x}| \, e^{-iHt} \, |\mathbf{x}_0\rangle && (2.2.13)\\ &= \frac{1}{2\pi^2|\mathbf{x}-\mathbf{x}_0|} \int_0^\infty dp \, p \sin(p|\mathbf{x}-\mathbf{x}_0|) e^{-it\sqrt{p^2+m^2}} && (2.2.14)\\ &= \frac{1}{4i\pi^2 x} \int_0^\infty dp \, p \left(e^{ipx} - e^{-ipx}\right) e^{-it\sqrt{p^2+m^2}} && (2.2.15)\\ &= \frac{1}{4i\pi^2 x} \int_0^\infty dp \, p \left[e^{i\left(px - t\sqrt{p^2+m^2}\right)} - e^{-i\left(px + t\sqrt{p^2+m^2}\right)}\right] && (2.2.16) \end{aligned}$$

There is an identity we can apply here to make the method of stationary phase applicable. Notice that we have an integral like,

$$\int_0^\infty dp \, p \left(e^{ipx} - e^{-ipx}\right) \tag{2.2.17}$$

We make a substitution of $p \to -u$ for the following part of the integral,

$$\int_0^\infty dp\, p\, e^{-ipx} = \int_0^{-\infty} du\, u\, e^{iux}$$

Since $u$ is a dummy integration variable we have,

$$\begin{aligned}\int_0^\infty dp\, p\, e^{-ipx} &= \int_0^{-\infty} dp\, p\, e^{ipx} \\ &= -\int_{-\infty}^0 dp\, p\, e^{ipx}\end{aligned}$$

Putting this back into 2.2.17,

$$\begin{aligned}\int_0^\infty dp\, p\left(e^{ipx} - e^{-ipx}\right) &= \int_0^\infty dp\, p\, e^{ipx} - \int_0^\infty dp\, p\, e^{-ipx} \\ &= \int_0^\infty dp\, p\, e^{ipx} + \int_{-\infty}^0 dp\, p\, e^{ipx} \\ &= \int_{-\infty}^\infty dp\, p\, e^{ipx}\end{aligned}$$

Using this identity in [Mic95], means that we can expand the limits of integration to the entire real line and make use of the method of stationary phase 2.2.9.

$$U(t) = \langle \mathbf{x}|\, e^{-iHt}\, |\mathbf{x}_0\rangle \tag{2.2.18}$$

$$= \frac{1}{4i\pi^2 x}\int dp\, p\, e^{i\left(px - t\sqrt{p^2+m^2}\right)} \tag{2.2.19}$$

This last integral is perfect for the stationary phase method. Let's find the stationary points of the phase factor now,

$$\frac{d}{dp}\left(px - t\sqrt{p^2+m^2}\right) = x - t\frac{p}{\sqrt{p^2+m^2}} = 0$$

To isolate $p_0$, we can try the following

$$x = t\frac{p}{\sqrt{p^2+m^2}}$$

$$x\sqrt{p^2+m^2} = pt$$

$$x^2\left(p^2+m^2\right) = p^2t^2$$

$$x^2p^2 + x^2m^2 = p^2t^2$$

Which can finally be rewritten as

$$p^2\left(x^2-t^2\right)=-m^2x^2$$

or

$$p_0=\pm i\frac{mx}{\sqrt{x^2-t^2}}$$

Substituting this back into the phase factor, we get

$$\exp\left(i\left(p_0x-t\sqrt{p_0^2+m^2}\right)\right)=\exp\left(i\left(\pm i\frac{mx^2}{\sqrt{x^2-t^2}}-t\sqrt{-\frac{m^2x^2}{x^2-t^2}+m^2}\right)\right)$$

Let's drop the exponent from the expression and continue working.

$$\begin{aligned}\mp\frac{mx^2}{\sqrt{x^2-t^2}}-it\sqrt{m^2}\sqrt{-\frac{x^2}{x^2-t^2}+1}&=\mp\frac{mx^2}{\sqrt{x^2-t^2}}-it\sqrt{m^2t^2}\sqrt{-1}\sqrt{\frac{1}{x^2-t^2}}\\&=\mp mx^2\frac{1}{\sqrt{x^2-t^2}}\pm mt^2\frac{1}{\sqrt{x^2-t^2}}\\&=\mp m\sqrt{x^2-t^2}\end{aligned}$$

Since [Mic95] chose to push the contour upward so that it goes through $p_0=+i\frac{mx}{\sqrt{x^2-t^2}}$, then we get the exponent term that looks like

$$e^{-m\sqrt{x^2-t^2}}$$

Recall from 2.2.9, that one of the terms in the answer is $e^{i\lambda S(x_0)}$, the exponential term at the phase factor, what we evaluated above, hence Peskin and schroeder's answer.

#### 2.2.8.1 Stationary Phase Method: Background

To understand this method well it helps to first read Wikipedia: Laplace's Method.

If anything we think these are the highlights from Laplace's method.

If $f(x)$ has a unique global maximum at $x_0$ and $M>0$ and if we deifne the following functions

$$g(x)=Mf(x)$$

and

$$h(x)=e^{g(x)}=e^{Mf(x)}$$

Then look at the ratios

$$\frac{g(x_0)}{g(x)}=\frac{f(x_0)}{f(x)}$$

and

$$\frac{h(x_0)}{h(x)} = e^{M(f(x_0)-f(x))}$$

As $M$ increases, the ratio for $h$ will grow exponentially, while the ratio for $g$ does not change. Since $f(x) \leq f(x_0)$, then the ratio of $h$ will always look like $e^{+xM}$. Said another way, in places where the value of $f(x)$ is a lot smaller than $f(x_0)$, the ratio of $h$ will look like $e^{Mf(x_0)}$, in places where the value of $f(x)$ is close to $f(x_0)$, the ratio will be smaller.

But in general the trick is to expand $f(x)$ around $x_0$ following Taylor's theorem

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{2}{2}f''(x_0)(x - x_0)^2 + R$$

Since $x_0$ is a global maxima, then $f'(x_0) = 0$ and $f''(x_0) < 0$. Hence,

$$f(x) \approx f(x_0) - \frac{1}{2}|f''(x_0)|(x - x_0)^2 + R$$

So we have this approximation,

$$\int_a^b e^{Mf(x)}dx \approx e^{Mf(x_0)} \int_a^b e^{-\frac{1}{2}M|f''(x_0)|(x-x_0)^2}dx$$

So we are approximating the function with a Gaussian. Here the last part of the approximation is to extend the limits of itnegration so that we do indeed have a Gaussian integral, which we assume we can ebcause the tails decay quickly. Since

$$\int dx\, e^{-a(x+b)^2} = \sqrt{\frac{\pi}{a}}$$

$$\int_a^b e^{Mf(x)}dx \approx e^{Mf(x_0)}\sqrt{\frac{2\pi}{M|f''(x_0)|}}$$

The next stepping stone is Wikipedia: the method of steepest descent. This method build upon Laplace's method and aplies to contour integrals in the complex plane. So now we are looking at approximating integrals such as

$$\int_C f(z)e^{\lambda g(z)}dz$$

#### 2.2.8.2 A Bit of Formalism: Understanding Tensors

$$(x \times y)^i = \sum_{j,k=1}^{n=3} \epsilon_{ijk}x_j y_k = \epsilon_{ijk}x_j y_k$$

Expanding the sum it looks like

$$\begin{aligned}(x \times y)^i &= \epsilon_{ijk} x_j y_k \\ &= \epsilon_{ij1} x_j y_1 + \epsilon_{ij2} x_j y_2 + \epsilon_{ij3} x_j y_3 \\ &= \epsilon_{i11} x_1 y_1 + \epsilon_{i12} x_1 y_2 + \epsilon_{i13} x_1 y_3 \\ &\quad \epsilon_{i21} x_2 y_1 + \epsilon_{i22} x_2 y_2 + \epsilon_{i23} x_2 y_3 \\ &\quad \epsilon_{i31} x_3 y_1 + \epsilon_{i32} x_3 y_2 + \epsilon_{i33} x_3 y_3 \\ &= 0 + \epsilon_{i12} x_1 y_2 + \epsilon_{i13} x_1 y_3 \\ &\quad \epsilon_{i21} x_2 y_1 + 0 + \epsilon_{i23} x_2 y_3 \\ &\quad \epsilon_{i31} x_3 y_1 + \epsilon_{i32} x_3 y_2 + 0\end{aligned} \tag{2.2.20}$$

The thing to notice here is that the above operation was just to obtain the $i$-th component of the cross product of $x \times y$.

This also leads us to the notation where a tensor of rank 3 is to be fed 3 vectors,

$$\epsilon(e_i, e_j, e_k) = \epsilon_{ijk} = e_i \cdot (e_j \times e_k)$$

Where, again, we see that the rank 3 tensor takes 3 vectors as input and produces a number. (a linear map or a functional one might say.)

From the explicit computation of the cross product we can also see that we obtain this matrix-like structure from computing each of the $i$-th components

$$\begin{pmatrix} \epsilon_{i11} & \epsilon_{i12} & \epsilon_{i13} \\ \epsilon_{i21} & \epsilon_{i22} & \epsilon_{i23} \\ \epsilon_{i31} & \epsilon_{i32} & \epsilon_{i33} \end{pmatrix}$$

when $i = 1$,

$$\begin{pmatrix} \epsilon_{111} & \epsilon_{112} & \epsilon_{113} \\ \epsilon_{121} & \epsilon_{122} & \epsilon_{123} \\ \epsilon_{131} & \epsilon_{132} & \epsilon_{133} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

and for $n = 3$ there are two other such matrices. And such is the Levi-Civita tensor of rank 3!

The other two "layers" of the tensor are

$$\begin{pmatrix} \epsilon_{211} & \epsilon_{212} & \epsilon_{213} \\ \epsilon_{221} & \epsilon_{222} & \epsilon_{223} \\ \epsilon_{231} & \epsilon_{232} & \epsilon_{233} \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} \epsilon_{311} & \epsilon_{312} & \epsilon_{313} \\ \epsilon_{321} & \epsilon_{322} & \epsilon_{323} \\ \epsilon_{331} & \epsilon_{332} & \epsilon_{333} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

These matrix representation of a tensor is actually the tensor evaluated on the basis $(\hat{x}, \hat{y}, \hat{z})$. So the multilinear function on a vector space that is a tensor can be visualized as a matrix when it is a rank 2 tensor (it eats two vectors as

input; keep reading, it'll make sense) when we are looking at just one component of $x \times y$, that is, by comparing what we just learned with 2.2.8.2 we can see that

$$(x \times y)^i = \begin{pmatrix} x_1, x_2, x_3 \end{pmatrix} \begin{pmatrix} \epsilon_{i11} & \epsilon_{i12} & \epsilon_{i13} \\ \epsilon_{i21} & \epsilon_{i22} & \epsilon_{i23} \\ \epsilon_{i31} & \epsilon_{i32} & \epsilon_{i33} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

And from here you should go and reach 3. We'll wait for you here.

#### 2.2.8.3 On the Way to the Method of Stationary Phase: Method of Steepest Descent

Building on top of the Laplace method, we now want to approximate

$$\mathcal{I}(z) = \int_C f(z) e^{\lambda g(z)} dz$$

Where $f(z)$ and $g(z)$ are analytical and we can deform the contour of integration into any $C'$ without changing the integral. This is the complex analogue to the "fundamental theorem for gradients", where if the function whose gradient we are integrating along a path is continuous, then its line integral along a close path is zero.

In complex analysis, this is the theorem called "Cauchy's integral theorem". The theorem actually talks about holomorphic functions $f(z)$ being integrated on a simply connected domain $\Omega$. But since all holomorphic functions are analytical, this is the topic of Cauchy Riemman's Theorem, and we can assume that we will always be on a simply connected domain (read Abbott's book on analysis if you want to know more about that concept), then

$$\int_C f(z) dz = 0$$

Let's start this adventure by looking at the an analytic function $g(z) = X(z) + iY(z)$, where $X(z)$ and $Y(z)$ are real-valued functions.

The steepest descent direction is the direction in which $\text{Re}(g(z)) = X(z)$ decreases most rapidly, not $g(z)$ itself. This is because we are interested in the behavior of the magnitude of the complex exponential $e^{\lambda g(z)}$, which is governed by the real part of the exponent.

$$e^{\lambda g(z)} = e^{\lambda (X(z) + iY(z))} = e^{\lambda X(z)} e^{i\lambda c}$$

So steepest descent is given by the negative gradient of $X(z)$:

$$-\nabla X = -\left( \frac{\partial X}{\partial x}, \frac{\partial X}{\partial y} \right)$$

The gradient of $Y(z)$ is perpendicular to the level curves of $Y(z) = c$, but we don't need to consider the negative gradient of $Y(z)$ for this proof.

Since $g(z)$ is analytic, the Cauchy-Riemann equations must hold, so that

$$\frac{\partial X}{\partial x} = \frac{\partial Y}{\partial y}$$

and

$$\frac{\partial X}{\partial y} = -\frac{\partial Y}{\partial x}$$

Now, let's use the Cauchy-Riemann equations to show that the constant phase contour is equivalent to the steepest descent contour. The direction of the constant phase contour is perpendicular to the gradient of $Y(z)$, so it is some $\pi/2$ clockwise or counterclockwise.

A counterclockwise rotation of 90 degrees can be expressed as

$$R_{90^\circ} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Applying this rotation matrix to the gradient vector $\nabla Y$:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial Y}{\partial x} \\ \frac{\partial Y}{\partial y} \end{pmatrix} = \begin{pmatrix} -\frac{\partial Y}{\partial y} \\ \frac{\partial Y}{\partial x} \end{pmatrix}$$

By the same logic we, clockwise rotation by 90 degrees would give us

$$\begin{pmatrix} \frac{\partial Y}{\partial y} \\ -\frac{\partial Y}{\partial x} \end{pmatrix}$$

Using the Cauchy-Riemann equations, we can rewrite the counterclockwise-rotated gradient of $Y(z)$ as:

$$\left(-\frac{\partial Y}{\partial y}, \frac{\partial Y}{\partial x}\right) = \left(-\frac{\partial X}{\partial x}, -\frac{\partial X}{\partial y}\right) = -\nabla X$$

Notice that the minus sign points us in the direction of steepest descent, so it is indeed the case that the **direction perpendicular to the gradient of $Y(z)$, which is the direction of the constant phase contour, is the same as the steepest descent direction, $-\nabla X$.**

#### 2.2.8.4 Morse Theory: Intro

Critical points can be degenerate or non-degenerate. A non-degenerate critical point has a Hessian that can be inverted - the Hessian of the function $f$ does not vanish at the critical point.

One associates a number called the **index**, the number of independent directions in which $f$ decreases from a critical point. More precisely, the index of

a non-degenerate critical point $p$ of $f$ is the dimension of the largest subspace of the tangent space to $M$ at $p$, where $M$ is a "landscape surface" function and $f : M \to \mathbb{R}$, on which the Hessian of $f$ is negavite definite.

The **index** also corresponds to the number of negative eigenvalues of the Hessian matrix at the critical cpoint $p$.

In mathematics, a **symmetric matrix** $M$ with real entries is positive-definite if the real number $z^T Mz$ is positive for every nonzero real column vector $z$. More generally, a Hermitian matrix (that is, a complex matrix equal to its conjugate transpose) is **positive-definite** if the real number $z^* Mz$ is positive for every nonzero complex column vector $z$.

#### 2.2.8.5 Morse's Lemma

This lemma is fundamental because it tells us that near a non-degenerate critical point, the function looks like a quadratic form, and this form is determined by the signature of the Hessian matrix (i.e., the number of positive and negative eigenvalues).

For example, suppose we have $f(x, y) = x^2 + y^2 - 2x + 4y$, $f : \mathbb{R}^2 \to \mathbb{R}$.

First, we find the critical point $\nabla f(x, y) = 0$,

$$\begin{aligned} \nabla f &= \langle \partial_x f, \partial_y f \rangle \\ &= \langle 2x - 2, 2y + 4 \rangle \end{aligned}$$

and we set $(2x - 2, 2y + 4) = (0, 0)$. So we have $2x - 2 = 0$ or $x = 1$. And then $2y + 4 = 0$, or $y = -2$. So $p = (1, -2)$.

Now, let's see if the critical point is non-degenerate. For that, let's compute the Hessian. Remember that the Hessian is given by,

$$H_f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

That is, the entry of the $i$-th row and the $j$-th column is

$$(H_f)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

So we want to compute

$$\begin{aligned} H_f &= \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} \\ &= \begin{pmatrix} \partial_x \partial_x f & \partial_x \partial_y f \\ \partial_y \partial_x f & \partial_y \partial_y f \end{pmatrix} \\ &= \begin{pmatrix} \partial_x (2x-2) & \partial_x (2y+4) \\ \partial_y (2x-2) & \partial_y (2y+4) \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned}$$

Given that all the entries on the diagonal are positive and there are no off-diagonal terms, it suggests that $f$ has a parabolic behavior in both the $x$ and $y$ directions independently (recall the second derivative test; both diagonal entries are positive so the function is concave up), which typically indicates a local minimum at the critical point if the function is convex or a saddle point if it is not.

Anyway, the Hessian is invertible and its eigenvalues are positive, so $p$ is a minimum.

Now we can apply Morse' lemma! Since both eigenvalues of our Hessian are positive, then index, $k$, is 0 (no negative eigenvalues), Morse's lemma tells us that near $p$ we can find coordinates $(u, v)$ such that

$$f(u, v) = f(p) + u^2 + v^2 = -5 + u^2 + v^2$$

So near the critical point, $f$ looks like a **paraboloid** downshifted by 5.

$$f(x) = f(p) - x_1^2 - \ldots - x_\gamma^2 + x_{\gamma+1}^2 + \ldots + x_n^2$$

$\gamma$ is the index of $f$ at $p$. The index $\gamma$ essentially determines how many negative squared terms appear in the local form of the function. Our index was 0 so we didn't have any.

In Morse's lemma, when it says there exists a chart $(x_1, x_2, \ldots, x_n)$ it means that near the critical point $p$, you can introduce a new coordinate system where:

- Each $x_i$ for $i = 1, \ldots, \gamma$ corresponds to a direction in which the function decreases (related to the negative eigenvalues of the Hessian).
- Each $x_i$ for $i = \gamma+1, \ldots, n$ corresponds to a direction in which the function increases (related to the positive eigenvalues of the Hessian).

The coordinates $x_i(p) = 0$ imply that the critical point $p$ is at the origin of this new coordinate system. The transformation to this new coordinate system is achieved by a linear change of variables that diagonalizes the Hessian matrix (through a process similar to eigendecomposition).

So here we added $u^2+v^2$ because our Hessian matrix had 2 positive eignevalues.

The complex generalization as stated in wikipedia: method of steepest descent provides another key idea of the Morse lemma.

As stated in the page, if $z^0$ is a saddle point and $S(z)$ is a holomorphic function, then there exist coordinates in terms of which $S(z) - S(z^0)$ is exactly quadratic. We already saw above more or less how it looks for $S(Z) - S(Z^0)$, or $f(x) - f(p)$, to be exactly quadratic. But the kicker is with the rest of the defintion.

If $S$ has a domain $W \in \mathbb{C}^n$, and $z^0$ is such that $\nabla S(z^0) = 0$ and $S''_{zz}(z^0) \neq 0$. Then there exists a neighborhood $U \in \mathbb{C}^n$ of $z^0$ and $V \in \mathbb{C}^n$ of $w = 0$ such that a bijective holomorphic function can be defined

$$\phi : V \to U$$

and $\phi(0) = z^0$. So this function sort of "remaps" our original function so that the saddle point becomes the origin.

The function is defined for all $w \in V$ as

$$S(\phi(w)) = S(Z^0) + \frac{1}{2}\sum_{j=1}^{n} \mu_j w_j^2$$

and $\det(\phi'_w(0)) = 1$. Here $\mu_j$ are the eignevalues of $S''_{zz}(z^0)$.

The use of the complex morse lemma is that we can separate our initial problem into components. Remember that we started with

$$I(\lambda) = \int_C f(z)e^{\lambda g(z)}dz$$

The integral is broken into the domain with the saddle point and the one without it, and then expressed in terms of the bijective function $\phi$, that maps a domain around $z^0$ to some other domain $U$. So we in the limit of $\lambda \to \infty$

$$\begin{aligned} I(\lambda) = I_0(\lambda) &= \int_{C_w} f(\phi(w))e^{\lambda S(\phi(w))}\, d\left(\phi(w)\right) \\ &= \int_{C_w} f(\phi(w)) \left(e^{\lambda S(z^0)} e^{\frac{1}{2}\lambda \sum_j \mu_j w_j^2}\right)\, d\left(\phi(w)\right) \\ &= e^{\lambda S(z^0)} \int_{C_w} f(\phi(w)) \exp\left(\frac{1}{2}\lambda \sum_j \mu_j w_j^2\right) \left|\det \phi'_w\right| dw \end{aligned}$$

To simplify the expression we have to remember how to use the Jacobian to fully do our transformation of variables.

$$f(z) = \int_0^1 \frac{d}{dt} f(tz_1, \ldots, tz_n) dt = \sum_{i=1}^{n} z_i \int_0^1 \frac{\partial f(z)}{\partial z_i}\Big|_{z=(tz_1,\ldots,tz_n)} dt$$

If we apply the chain rule, while keeping in mind that $z_i$ is the actual variable of $f$, not $tz$.

$$\frac{d}{dt} f(tz_1, ..., tz_n) = \sum_{i=1}^{n} \frac{\partial f}{\partial z_i} \cdot \frac{d}{dt}(tz_i) = \sum_{i=1}^{n} z_i \frac{\partial f}{\partial z_i}$$

$$\mathcal{I}_j = \int_{-\infty}^{\infty} e^{\frac{1}{2}\lambda_j y^2} dy = 2\int_0^{\infty} e^{-\frac{1}{2}\lambda\left(\sqrt{-\mu_j}y\right)^2} dy = 2\int_0^{\infty} e^{-\frac{1}{2}\lambda|-\mu_j|^2 y^2 \exp(2i \arg\sqrt{-\mu_j})} dy$$

The last expression is obtained by writing the complex number $\mu$ as

$$-\mu_j = |-\mu_j| e^{i \arg(-\mu_j)}$$

or

$$\sqrt{-\mu_j} = \sqrt{|-\mu_j|} e^{i \arg(\sqrt{-\mu_j})}$$

#### 2.2.8.6 Homotopy

If $M$ is a "landscape surface" function and $f : M \to \mathbb{R}$, and $M^a = f^{-1}(-\infty, a]$, The topology of $M^a$ does not change except when $a$ passes the height of a critical point; at this point, a $\gamma$-cell is attached to $M^a$, where $\gamma$ is the index of the point. This does not address what happens when two critical points are at the same height, which can be resolved by a slight perturbation of $f$. In the case of a landscape or a manifold embedded in Euclidean space, this perturbation might simply be tilting slightly, rotating the coordinate system.

## 2.3 Integrals in QFT

### 2.3.1 Gaussians

The simples one,

$$\int e^{-x^2}\,dx = \sqrt{\pi}$$

The most commn trick seems to have been popularized by Poisson, see wiki: Gaussian integral:

$$\begin{aligned}\left(\int dx\, e^{-x^2}\right)^2 &= \left(\int dx\, e^{-x^2}\right)\left(\int dy\, e^{-y^2}\right)\\ &= \iint e^{-\left(x^2+y^2\right)}\,dxdy\end{aligned}$$

Going to polar coordinates, where $r^2 = x^2 + y^2$ and the $dx\,dy$ measure becomes $rdrd\theta$, we have

$$\begin{aligned}\iint_{\mathbb{R}^2} e^{-\left(x^2+y^2\right)}\,dxdy &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r dr d\theta\\ &= 2\pi\int_0^{\infty} e^{-r^2} r dr\end{aligned}$$

The u-substitution of $u = -r^2$ makes $du = -2rdr$ or $rdr = -\frac{1}{2}du$. With this substitution we also need to evaluate the limits of integration. When $r \to 0$, $u = -r^2 \to 0$. And when $r \to \infty$, $u = -r^2 \to -\infty$. So we have,

$$\begin{aligned}2\pi\int_0^{-\infty} e^{-r^2} r dr &= -\pi\int_0^{\infty} e^{u} du\\ &= -\pi\left(\lim_{x\to\infty} e^{-\infty} - e^0\right)\\ &= -\pi\,(0-1)\\ &= \pi\end{aligned}$$

And thus we have our proof.
If we instead had landed on:

$$\int e^{-ar^2} r dr$$

Then our u-substitution would have looked as $u = -ar^2$ and $du = -2ardr$, so $rdr = -\frac{1}{2a}$. The limits of integration would have looked the same but we would have ended up with a factor of $a$ in the denominator giving us another handy Gaussian.

$$\int e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

And if we had had

$$\int e^{-a(x+b)^2} dx$$

The trick is to do the u-substitution earlier. We now have $u = x + b$ so that $du = dx$, the limits of integration don't change and we'd have

$$\int e^{-a(x+b)^2} dx = \int e^{-au^2} dx$$

Which we already know the answer to, and $b$ seems to not matter at all!

And finally, but very importantly, we will show how to complete the square in order to solve

$$\int e^{-ax^2+bx+c} dx$$

To complete the square, we want to get $ax^2+bx+c$ into the form $a(x+b)+k$ so we can re-use the previous result.

1. factor out $a$: $a\left(x^2 + \frac{b}{a}x\right) + c$
2. take the coefficient of $x$, divide it by 2 and square it: $a\left(x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2\right) + c$
3. by this point we have a perfect square, so separate it: $a\left(x + \frac{b}{a}\right)^2 - a\left(\frac{b}{2a}\right)^2 + c$

With this trcik we can arrive at

$$\int e^{-\left(ax^2+bx+c\right)} dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}-c}$$

But remember how to complete the square! We will be using it more later on.

### 2.3.2 Complex Gaussians: the Fresnel Integral

Convergence of $\int e^{-at^2} dt$.

When $a \in \mathbb{R}$ and $a > 0$, $e^{-at^2}$ decays rapidly as $|t| \to \infty$.

When $a = re^{i\theta}$ we can write the phase factor as

$$e^{-at^2} = e^{-rt^2(\cos\theta + i\sin\theta)} = e^{-rt^2\cos\theta} e^{-irt^2\sin\theta}$$

If we remember Euler's formula, we'll see that the imaginary component just spins around. So convergence is solely controlled by the real part. For the real

part to converge, we need $-rt^2\cos\theta$ to stay negative as $|t| \to \infty$. For this to happen, we need to keep $\theta \in (-\pi/2, \pi/2)$.

With this in mind, we can perform an analytic continuation by rotating $\theta$ from 0 to $\pi/2$. This is because both $\int e^{-at^2}dt$ and its real-solution is $\sqrt{\frac{\pi}{a}}$ are both analytic and we keep the integral from exploding along the path we chose. A fundamental theorem of complex analysis, the identity theorem says that if two analytic functions are equal on a line segment and both are analytic in a larger region then they must be equal throughout that region.

So for example, when we have $\int e^{-\frac{1}{2}it^2}dt$, then $a = re^{i\theta} = \left(\frac{1}{2}\right)(-i)$. Thus,

$$\int e^{-\frac{1}{2}it^2}dt = \sqrt{\frac{\pi}{\left(\frac{1}{2}\right)(-i)}} = \sqrt{2\pi}\left(\sqrt{-\frac{1}{i}}\right) = \sqrt{2\pi}\sqrt{i} = \sqrt{2\pi}e^{i\pi/4}$$

### 2.3.3 Kinetic operator and a Green Function

This problem is part of MontePython: Implementing Quantum Monte Carlo using Python.

There is a kinetic operator defined as

$$G_K\left(\mathbf{R}', \mathbf{R}, t\right) = \frac{1}{\left(2\pi\right)^{3N}}\int e^{-i\mathbf{k}\mathbf{R}'}e^{-Dtk^2}e^{-i\mathbf{k}\mathbf{R}}d\mathbf{k}$$

This integral then is integrated and the following Green function is obtained

$$G_K\left(\mathbf{R}', \mathbf{R}, t\right) = \frac{1}{\left(4\pi Dt\right)^{3N/2}}e^{-\left(\mathbf{R}-\mathbf{R}'\right)/4Dt}$$

And as we saw in Jackson problem 1.2, the Green function we just got is equivalent to $\delta\left(\mathbf{R} - \mathbf{R}'\right)$.

But coming back to the integral, the trick is to complete the square!

$$\begin{aligned} G_K\left(\mathbf{R}', \mathbf{R}, t\right) &= \frac{1}{\left(2\pi\right)^{3N}}\int e^{-i\mathbf{k}\mathbf{R}'}e^{-Dtk^2}e^{-i\mathbf{k}\mathbf{R}}d\mathbf{k} \\ &= \frac{1}{\left(2\pi\right)^{3N}}\int e^{-Dtk^2+-i\mathbf{k}\cdot\left(\mathbf{R}-\mathbf{R}'\right)}d\mathbf{k} \end{aligned}$$

This last integral has the form $\int e^{ax^2+bx}dx$, which can be solved as follows,

$$\begin{aligned} \int e^{ax^2+bx}dx &= \int \exp\left\{a\left(x^2 + \frac{b}{a}x\right)\right\}dx \\ &= \int \exp\left\{a\left(x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2\right)\right\}dx \\ &= \int \exp\left\{a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a}\right\}dx \\ &= e^{-b^2/4a}\int e^{a\left(x+\frac{b}{2a}\right)^2}dx \end{aligned}$$

Now, let's look at $\int e^{a(x+b/2a)^2}dx$. First, we need to make a transformation. Let's define $u = x + b/2a$, then $du = dx$, the measure remains invariant (so do the limits of integration), so $\int e^{a(x+b/2a)^2}dx = \int e^{au^2}du$. Now, if we define $a = -c$, then

$$\begin{aligned}\int e^{a(x+b/2a)^2}dx &= \int e^{au^2}du \\ &= \int e^{-cu^2}du \\ &= \sqrt{\frac{\pi}{c}} = \sqrt{\frac{\pi}{-a}}\end{aligned}$$

This whole thing works if $a < 0$ but it turns out that this is also valid if $\mathrm{Re}\,(a) \leq 0$ but $a \neq 0$.

Putting everything back together,

$$\int e^{ax^2+bx}dx = \sqrt{\frac{\pi}{-a}}e^{-b^2/4a}$$

Looking back at our original problem, we can chose $a = -Dt$ and $b = i\mathbf{r}$, where $\mathbf{r} = \mathbf{R} - \mathbf{R}'$. and so

$$\begin{aligned}G_K\left(\mathbf{R}', \mathbf{R}, t\right) &= \frac{1}{(2\pi)^{3N}}\int e^{-Dtk^2+-i\mathbf{k}\cdot\mathbf{r}}d\mathbf{k} \\ &= \frac{1}{(2\pi)^{3N}}\int e^{-Dt\left(k^2+\frac{i\mathbf{k}}{Dt}\cdot\mathbf{r}\right)}d\mathbf{k} \\ &= \frac{1}{(2\pi)^{3N}}\int e^{-Dt\left(k^2+\frac{i\mathbf{r}}{Dt}\cdot\mathbf{k}+\left(\frac{i\mathbf{r}}{2Dt}\right)^2-\left(\frac{i\mathbf{r}}{2Dt}\right)^2\right)}d\mathbf{k} \\ &= \frac{1}{(2\pi)^{3N}}\int e^{-Dt\left(k+\frac{i\mathbf{r}}{2Dt}\right)^2-\frac{\mathbf{r}^2}{4Dt}}d\mathbf{k} \\ &= \frac{1}{(2\pi)^{3N}}e^{-\mathbf{r}^2/4Dt}\int e^{-Dt\left(k+\frac{i\mathbf{r}}{2Dt}\right)^2}d\mathbf{k}\end{aligned}$$

Again, since $\int e^{-ax^2}dx = \sqrt{\frac{\pi}{a}}$, then

$$\int e^{-Dt\left(k+\frac{i\mathbf{r}}{2Dt}\right)^2}dk = \sqrt{\frac{\pi}{Dt}}$$

Thus,

$$\begin{aligned} G_K\left(\mathbf{R}', \mathbf{R}, t\right) &= \frac{1}{(2\pi)^{3N}} e^{-\mathbf{r}^2/4Dt} \int e^{-Dt\left(k+\frac{i\mathbf{r}}{2Dt}\right)^2} d\mathbf{k} \\ &= \frac{1}{(2\pi)^{3N}} e^{-\mathbf{r}^2/4Dt} \left(\sqrt{\frac{\pi}{Dt}}\right)^{3N} \\ &= \frac{1}{(4\pi^2)^{3N/2}} \frac{\pi^{3N/2}}{(Dt)^{3N/2}} e^{-\mathbf{r}^2/4Dt} \\ &= \frac{1}{(4\pi Dt)^{3N/2}} e^{-\left(\mathbf{R}-\mathbf{R}'\right)^2/4Dt} \end{aligned}$$

### 2.3.4 Klein-gordon Non-relativistic Amplitude

The following integral comes from Peskin and schroeder Section 2.1

$$\int \frac{d^3p}{(2\pi)^3} e^{-i(\mathbf{p}^2/2m)t} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)} = \left(\frac{m}{2\pi it}\right)^{3/2} e^{im(\mathbf{x}-\mathbf{x}_0)^2/2t} \tag{2.3.1}$$

This integral follows the same procedure as 2.3.3.

Here we have $a = -\frac{it}{2m}$ and $b = i(\mathbf{x} - \mathbf{x_0})$. Since $\operatorname{Re} a = -t/2m \leq 0$ and $a \neq 0$ when $t \neq 0$ our solution is valid for $t \neq 0$. Then

$$\sqrt{\frac{\pi}{-a}} = \sqrt{\frac{2\pi m}{it}}$$

and

$$\begin{aligned} -\frac{b^2}{4a} &= \frac{2m\left(\mathbf{x}-\mathbf{x}_0\right)^2}{-4it} \\ &= \frac{im\left(\mathbf{x}-\mathbf{x}_0\right)^2}{2t} \end{aligned}$$

So the solution is

$$\begin{aligned} &\frac{1}{(2\pi)^3} \left(\frac{2\pi m}{it}\right)^{3/2} e^{\frac{im(\mathbf{x}-\mathbf{x}_0)^2}{2t}} \\ &= \left(\frac{m}{2\pi it}\right)^{3/2} e^{im(\mathbf{x}-\mathbf{x}_0)^2/2t} \end{aligned}$$

# Chapter 3

# Tensors

Working with a spacetime metric signature $(+---)$,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$x^\mu = (x^0, \mathbf{x})$ and $x_\mu = g_{\mu\nu}x^\nu = (x^0, -\mathbf{x})$. To see this explciitly, let's carry out the summations.

When $\mu = 0$, $x_0 = g_{00}x^0 + g_{01}x^1 + g_{02}x^2 + g_{03}x^3 = x^0 + 0 + 0 + 0 = x^0$. For $\mu = 1$, $x_1 = g_{10}x^0 + g_{11}x^1 + g_{12}x^2 + g_{13}x^3 = 0 + x^1 + 0 + 0 = -x^0$. Similarly, when $\mu = 2$, $x_2 = -x^2$, and when $\mu = 3$, $x_3 = -x^3$.

The contravariant components of the vector $x^\mu$ are just its components. The covariant compoents $x_\mu$, defined to be equal to $x_\mu = x^\nu g_{\nu\mu}$, are actually the components of the associated dual vector. Why? well, in order to obtain $x_\mu$ we had to use a "non-degenerate hermitian form" (something that looks like an inner product; though it isn't since this metric is not positive definite). This makes it a dual vector by definition.

Also note that $g_{\mu\nu}$ forms a **bilinear form** since $g(x_1, x_2) = g(x_2, x_1)$ (linearity on both sides of the non-degenerate hermitian form).

### 3.0.1 Inner Product

Why does $p \cdot x = g_{\mu\nu}p^\mu x^\nu = g^{\mu\nu}p_\mu x_\nu = p_\mu x^\nu = p^\mu x_\nu = p^0x^0 - \mathbf{p} \cdot \mathbf{x}$?

To start, we ought to know that there is the implicit defintion of $x_\mu = x^\nu g_{\nu\mu}$. So the covariant vector is defined through the application of a non-degenerate hermitian form and it constitutes a dual vector. From there the rest is linear algebra.

The metric tensor defines the geometry of spacetime, including the way distances and angles are measured. The implied summation, einstein summation, effectively 'weights' the components according to the geometry of spacetime.

$g_{\mu\nu}$ lowers indices, while $g^{\mu\nu}$ raises them. $\delta^{\mu}_{\nu}$ is a diagonal matrix with ones on the diagonal and zeros elsewhere. It selects the $\mu$-th component when used in a summation, acting like an identity element.

$g^{\mu\alpha}g_{\alpha\nu} = \delta^{\mu}_{\nu}$

$$g^{\mu\alpha}g_{\alpha\nu} = \sum_{\alpha=0} g^{\mu\alpha}g_{\alpha\nu} = g^{\mu 0}g_{0\nu} + g^{\mu 1}g_{1\nu} + g^{\mu 2}g_{2\nu} + g^{\mu 3}g_{3\nu}$$

This is essentially a matrix multiplication.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Going component by component, take the second row and second column, $\mu = 1$ and $\nu = 1$. The element for that entry is given by

$$g^{1\alpha}g_{\alpha 1} = g^{10}g_{01} + g^{11}g_{11} + g^{12}g_{21} + g^{13}g_{31} = 0 + (-1)(-1) + 0 + 0 = 1$$

### 3.0.2 Derivatives

The examples here come from [Kay11]. The aim here is to see a few examples to get used to the quantities we are going to be seeing later on.

Some useful identities before we carry on:

1.
$$\frac{\partial x_p}{\partial x_q} = \delta_{pq}$$

**Example 1**

If $a_{ij}$ are contants, to compute the partial derivative

$$\frac{\partial}{\partial x_k} a_{ij}x_i x_j$$

We'll firs write out the sums explciitly.

$$a_{ij}x_i x_j = \sum_{i,j} a_{ij}x_i x_j$$

Someone once happened to noticed that the sum could be broken into parts depending on when $x_{\{i,j\}} = x_k$, where $x_k$ is the one we want to differentiate

with respect to,

$$\begin{aligned}\sum_{i,j} a_{ij}x_ix_j &= \sum_{i\neq k, j\neq k} a_{ij}x_ix_j + \sum_{i,j\neq k} a_{ij}x_ix_j + \sum_{i\neq k, j} a_{ij}x_ix_j + \sum_{i=k, j=k} a_{ij}x_ix_j \\ &= C + \sum_{j\neq k} (a_{kj}x_j)\, x_k + \sum_{i\neq k} (a_{ik}x_i)\, x_k + a_{kk}x_kx_k \\ &= C + \sum_{j\neq k} (a_{kj}x_j)\, x_k + \sum_{i\neq k} (a_{ik}x_i)\, x_k + a_{kk}(x_k)^2\end{aligned}$$

We wrote the first sum, where no $x$s are equal to $x_k$ as $C$, because as you may guess this is going to go away with any derivative.

Back to our original task,

$$\begin{aligned}\frac{\partial}{\partial x_k} a_{ij}x_ix_j &= 0 + \sum_{j\neq k} a_{kj}x_j + \sum_{i\neq k} a_{ik}x_i + 2a_{kk}x_k \\ &= \sum_j a_{kj}x_j + \sum_i a_{ik}x_i \\ &= a_{kj}x_j + a_{ik}x_i \\ &= a_{ki}x_i + a_{ik}x_i \\ &= (a_{ik} + a_{ki})x_i\end{aligned}$$

**Example 2**

Another way to work out

$$\frac{\partial}{\partial x_k} a_{ij}x_ix_j$$

goes as follows.

$$\begin{aligned}\frac{\partial}{\partial x_k} a_{ij}x_ix_j &= a_{ij}\frac{\partial}{\partial x_k} x_ix_j \\ &= a_{ij}\left(\frac{\partial x_i}{\partial x_k}x_j + x_i\frac{\partial x_j}{\partial x_k}\right) \\ &= a_{ij}\left(\delta_{ik}x_j + x_i\delta_{jk}\right) \\ &= a_{kj}x_j + a_{ik}x_i \\ &= (a_{ik} + a_{ki})x_i\end{aligned}$$

**Example 3**

If $a_{ij} = a_{ji}$ are constants, compute

$$\frac{\partial^2}{\partial x_k \partial x_l} a_{ij}x_ix_j$$

Let's start with our previous result,

$$\begin{aligned}\frac{\partial^2}{\partial x_k \partial x_l} a_{ij} x_i x_j &= \frac{\partial}{\partial x_k}\left[\frac{\partial}{\partial x_l} a_{ij} x_i x_j\right] \\ &= \frac{\partial}{\partial x_k}(a_{il} + a_{li}) x_i \\ &= 2a_{li}\frac{\partial x_i}{\partial x_k} \\ &= 2a_{li}\delta_{ik} \\ &= 2a_{lk}\end{aligned}$$

### 3.0.3 Changing Variables: the Jacobian

Promise we are not doing the conventional thing where we go from scratch but there is one more pattern we have to covered now in order for some other equations to make sense and this is the Jacobian.

There is a result from calculus that is presented a bit like this (the following presentation comes from [Daw]):

If we define the Jacobian of a transformation $x = g(u, v)$, $y = h(u, v)$ as

$$\begin{aligned}\frac{\partial(x, y)}{\partial(u, v)} &= \begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix} \\ &= \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\end{aligned}$$

Then if we want to integrate $f(x)$ over some surface $R$, this region will become some other surface $S$ following our transformation and the integral can be computed as follows:

$$\int\int_R f(x, y) dA = \int\int_S f\left(g(u, v), h(u, v)\right) \left|\frac{\partial(x, y)}{\partial(u, v)}\right| d\overline{A}$$

For example, when we go from Euclidean to polar coordinates, the transformation we use is $x = r\cos\theta$ and $y = r\sin\theta$ and the Jacobian is $r$ so $dA = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| dxdy = rdrd\theta$:

$$\begin{aligned}\frac{\partial(x,y)}{\partial(r,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\ &= r\cos^2\theta + r\sin^2\theta \\ &= r\end{aligned}$$

In 3-dimensions this extends to

$$\int\int\int_V f(x,y,z)dV = \int\int\int_B f\left(g(u,v,w), h(u,v,w), k(u,v,w)\right)\left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| d\overline{V}$$

For the case of the transformation of Euclidean coordinates to spherical coordinates, the transformation of variables is $x = r\sin\phi\cos\theta$, $y = r\sin\phi\sin\theta$, and $z = r\cos\phi$. In that case the Jacobian is $r^2\sin\phi$ and thus $dV = \left|\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)}\right| dxdydz = r^2\sin\phi dr d\phi d\theta$.

Besides being very widely used, the reason we mentioned it is because this is also that is very well used in [Tom14] to introduce the notation we see so much in QFT. Specifically, the Jacobian as we have been writing it, can also be seen in expressions such as

$$\overline{a}^\mu = \left(\frac{\partial \overline{x}^\mu}{\partial x^\nu}\right) a^\nu$$

$$\frac{\partial \phi}{\partial \overline{x}^\mu} = \left(\frac{\partial x^\nu}{\partial \overline{x}^\mu}\right)\frac{\partial \phi}{\partial x^\nu}$$

The determinant of the Jacobian matrix is used specifically when you are dealing with volume transformations (like integrating over a volume or changing the measure of integration) because it scales the volume elements according to how the coordinate transformation stretches or compresses the space.

When transforming vectors, however, you're not scaling a volume but rather reorienting or rescaling each component of the vector according to how the coordinate axes themselves have changed. This is why the transformation involves a sum over the product of vector components and the appropriate Jacobian elements:

- Contravariant vectors (standard vectors): The components of these vectors are transformed by multiplying with the Jacobian matrix $\partial\overline{x}^\mu/\partial x^\nu$. Here, each new vector component is a linear combination of old components weighted by how much each new coordinate axis changes with respect to each old axis.

- These objects, such as displacement vectors, inherently "follow" the coordinate axes. As the axes stretch or rotate, so do the components of the vector.

- Covariant vectors (gradients and one-forms): These vectors transform using the inverse of the Jacobian matrix $\partial x^\nu / \partial \overline{x}^\mu$. This reflects how changes in the old coordinates map to the new coordinates, suitable for objects that naturally pair with vectors, like gradients.

  - They measure rates of change along these coordinates.
  - A larger coordinate stretch means a smaller gradient in that direction to maintain the same rate of change.

### 3.0.4 From Tensors to Matrices

Tensors can be expressed as matrices only once a basis has been chosen. the example being given is with the angular momentum operator $L_z$ on the spherical harmonics $H_l(\mathbb{R}^3)$.

It says that $H_l(\mathbb{R}^3)$ is the set of all linear functions on $\mathbb{R}^3$ that

$$\{rY_m^l\}_{-1\leq m\leq 1} = \left\{ \frac{1}{\sqrt{2}}(x+iy),\, z,\, -\frac{1}{\sqrt{2}}(x-iy) \right\}$$

and $\{x, y, z\}$ are both basis for the space.

Then having the angular momentum operator $L_z = -i(x\partial_y - y\partial_x)$ on this space we got

$$\frac{1}{\sqrt{2}} L_z(x+iy) = \frac{1}{\sqrt{2}}(x+iy)$$

$$L_z(z) = 0$$

$$-\frac{1}{\sqrt{2}} L_z(x-iy) = \frac{1}{\sqrt{2}}(x-iy)$$

which implies that in the spherical harmonics basis

$$L_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

That tensor representation of $L_z$ is implies by saying that $T(e_i) = \sum_{j=1}^n T_i^j e_j$.

$$\frac{1}{\sqrt{2}} L_z(x+iy) \to L_z\left( \frac{1}{\sqrt{2}}(x+iy) \right)$$

$$L_z(z) = 0$$

$$-\frac{1}{\sqrt{2}} L_z(x-iy) \to L_z\left( -\frac{1}{\sqrt{2}}(x+iy) \right)$$

Since

$$\{rY_m^l\}_{-1\leq m\leq 1} = \Big\{\frac{1}{\sqrt{2}}(x+iy),\, z,\, -\frac{1}{\sqrt{2}}(x-iy)\Big\}$$

then we can say that $e_1 = \frac{1}{\sqrt{2}}(x+iy)$, $e_2 = z$, and $e_3 = -\frac{1}{\sqrt{2}}(x-iy)$.

Meaning that $L_z(e_1) = e_1 = 1\cdot e_1 + 0\cdot e_2 + 0\cdot e_3$, $L_z(e_2) = 0 = 0\cdot e_1 + 0\cdot e_2 + 0\cdot e_3$, and $L_z(e_3) = -e_3 = 0\cdot e_1 + 0\cdot e_2 - 1\cdot e_3$. From here, we can then adopt write this as a matrix

$$T_i^j = L_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

The columns of a matrix representation of a linear operator correspond to the action of the operator on the basis vectors. However, due to the specific diagonal form of the matrix in this example, both the row and column perspectives give the same result, which can indeed be confusing.

The first column represents $L_z(e_1)$, the second column represents $L_z(e_2)$, the third column represents $L_z(e_3)$.

In physics speak this would be read as the wave functions $\frac{1}{\sqrt{2}}(x+iy)$, $z$, and $-\frac{1}{\sqrt{2}}(x-iy)$ have eigenvalues 1, 0, and -1, respectively.

## 3.1 Transformation Matrices

The previous section gave us a good place to begin talking about tensors. Now let's actually talk about them.

Tensors have been historically defined based on how they transform. Think of the conventional axis rotation formula in 2D. If we rotate the axis by an angle $\theta$, then

$$x' = x\cos\theta - y\sin\theta$$

and

$$y' = x\sin\theta + y\cos\theta$$

As we saw in the previous section, where we derived the matrix representation of the angular momentum operator, the columns of a matrix representation of a multilinear map (tensor) correspond to the action of the map (tensor) on the basis vectors. Which is most commonly seen as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

If we envision $x$ and $y$, and $x'$ and $y'$, and some basis sets, such that $x \to e_1$ and $y \to e_2$ (similarly for the primed components). We can rewrite the above as

$$e_{1'} = e_1\cos\theta - e_2\sin\theta$$

and

$$e_{2'} = e_1 \sin\theta + e_2 \cos\theta$$

Meaning that our problem becomes,

$$\begin{pmatrix} e_{1'} \\ e_{2'} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

Following the Einstein summation convention, we could write

$$e_{i'} = A^{j}_{i'} e_j$$

which also means that the original basis can be obtained as follows

$$e_i = A^{j'}_{i} e_{j'}$$

It is interesting to now quickly pursue the exercise of expressing our basis in terms of itself as it will show us some properties about these numbers $A^{j}_{i'}$ and $A^{j'}_{i}$.

Combining the previous two expressions we get,

$$e_i = A^{j'}_{i} e_{j'} = A^{j'}_{i} A^{k}_{j'} e_k$$

From here we can see that

$$A^{j'}_{i} A^{k}_{j'} = \delta^{k}_{i}$$

If we consider the same example but with the primed-basis.

$$e_{i'} = A^{j}_{i'} e_j = A^{j}_{i'} A^{k'}_{j} e_{k'}$$

Meaning that

$$A^{j}_{i'} A^{k'}_{j} = \delta^{k'}_{i'}$$

So the different $A$s are inverse of one another.

It is also worth noting that these $A$s are tensor components.

In expressions such as $A^{j}_{i'} A^{k'}_{j}$, the indices $i'$ and $k'$ refer to the primed basis, while the index $j$ refers to the unprimed basis. This mixing of basis sets in a single expression is why the numbers $A^{j}_{i'}$ and $A^{k'}_{j}$ are not considered tensor components.

The transformation matrices $A^{j}_{i'}$ and $A^{k'}_{j}$ are not tensor components themselves, but rather express the relationship between components in different bases. The fact that these matrices are inverses of each other, which is a consequence of the transformation law for tensor components.

Another useful thing to note is that the transformation for dual vectors looks the same but with uppoer indices.

$$e^{i'} = A^{i'}_{j} e^{j}$$

and

$$e^i = A^i_{j'} e^{j'}$$

With this info at hand, we can guesstimate the matrix transformation laws for tensors,

$$A = \begin{pmatrix} A^{1'}_1 & A^{1'}_2 & \dots & A^{1'}_n \\ A^{2'}_1 & A^{2'}_2 & \dots & A^{2'}_n \\ \vdots & \vdots & \ddots & \vdots \\ A^{n'}_1 & A^{n'}_2 & \dots & A^{n'}_n \end{pmatrix}$$

Its inverse is

$$A^{-1} = \begin{pmatrix} A^1_{1'} & A^1_{2'} & \dots & A^1_{n'} \\ A^2_{1'} & A^2_{2'} & \dots & A^2_{n'} \\ \vdots & \vdots & \ddots & \vdots \\ A^n_{1'} & A^n_{2'} & \dots & A^n_{n'} \end{pmatrix}$$

These are unitary matrices

$$AA^{-1} = A^{-1}A = I$$

Equating the above with our 2D rotation example,

$$A^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

## 3.2 Contravariant and Covariant Vectors

One fun fact about vectors is that they can be seen as tensors of type (0,1), that is, they feed on 1 dual vector. A vector can be viewed as a tensor of type (0, 1) because it transforms according to the contravariant transformation law under a change of coordinates. The key idea is that a vector represents a directed quantity that "eats" or operates on dual vectors (covectors or linear functionals) to produce scalars.

This is mathematically said as,

$$v^{i'} = A^{i'}_j v^j$$

As we saw above, this is the transformation for dual vectors - expressions with contravariant indices.

And as we saw above, these transform according to the Jacobian.

A dual vector in turn transforms as

$$f_{i'} = A^j_{i'} v_j$$

Note how the transformation are the contrary to what we saw for the basis vectors! Where as the components of the dual vector transform in the same way as the basis.

In matrix notation we have contravariant vectors transform as

$$v_{\mathcal{B}'} = A v_{\mathcal{B}}$$

and covariant ones as

$$f_{\mathcal{B}'} = \left(A^{-1}\right)^T f_{\mathcal{B}}$$

But remember that basis and dual basis transform in the opposite ways!

Now, in quantum mechanics, if we consider an orthonormal basis $|e_{i'}\rangle$ in the same Hilbert space. And assume that both $|e_i\rangle$ and $|e_{i'}\rangle$ are complete bases, so there must be a unitary transformation $U$ that relates them. We would write, $|e_{i'}\rangle = U|e_i\rangle$ for this. Consecuently, the dual basis transforms as follows, $\langle e_{i'}| = \langle e_i| U^\dagger$.

Here, $U$ plays the role of $A^{-1}$, or $A^j_{i'}$ in the tensor formalism.

Following the previous approach, we get the following results for transformations into a primed basis set,

$$|e_{i'}\rangle = U|e_i\rangle$$

and

$$\langle e_{i'}| = \langle e_i| U^\dagger$$

The matrix elements of $U$ in the $|e_i\rangle$ basis are given by: $U^j_{i'} = \langle e_{i'}| U |e_j\rangle$. These matrix elements play the same role as the $A^j_{i'}$ in your classical expression.

### 3.2.1 Pauli Matrices

The Pauli matrices actually form a basis for $2 \times 2$ Hermitian matrices. Let's see how: a generic $2 \times 2$ Hermitian matrix would look like

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}$$

Meaning that $a, b \in \mathbb{R}$ and that $b = c^*$. So we could write our Hermitian matrix as

$$\begin{pmatrix} a & x - iy \\ x + iy & d \end{pmatrix}$$

Another requirement that is imposed on $H_2(\mathbb{C})$ is that they must be unitary, that is $U^\dagger U = I$, so

$$\begin{pmatrix} a & x - iy \\ x + iy & d \end{pmatrix} \begin{pmatrix} a & x - iy \\ x + iy & d \end{pmatrix} = \begin{pmatrix} a^2 + x^2 - y^2 & a(x - iy) + b(x - iy) \\ a(x + iy) + b(x + iy) & d^2 + x^2 - y^2 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

One way to make this happen is to set $a = -d$, this will also have the happy side effect of meeting another requirement for the Pauli matrices to be traceless. So now we have our candidate bases that ought to look like

$$\begin{pmatrix} a & x - iy \\ x + iy & -a \end{pmatrix}$$

This then means that our requirement for our matrix $H_2(\mathbb{C})$ to be unitary simplifies to

$$\begin{pmatrix} a^2 + x^2 - y^2 & 0 \\ 0 & a^2 + x^2 - y^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Here is where our 3 Pauli matrices comes!

$$\sigma_1 = \sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

comes from assuming that $x = 1$, so $a^2 + x^2 - y^2 = 1$ can be met by assuming that $a = y = 0$.

$$\sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

comes from assuming that $y = 1$, so $a^2 + x^2 - y^2 = 1$ can be met by assuming that $a = x = 0$.

Finally,

$$\sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Where we took $a = 1$ and $x = y = 0$.

## 3.3 Tensor Product

By definition, a **vector** can be writen as

$$v = v^i e_i$$

See section 2.3. $v^i$ are called the **contravariant** components of $v$.

From section 2.5, we see that we can also define the dual vector $e^i$ as

$$e^i(v) = v^i$$

A **dual vector** $f$ can then be defined as (remember that a dual vector is a linear functional that takes a vector $v$ and returns a scalar)

$$f(v) = f(v^i e_i) = v^i f(e_i) = v^i f_i$$

So while we refer to the compoents of a vector as $v^i$ on a basis $e_i$, we refer to the components of a dual vector as $f_i$. And you might have guessed that the dual vector itself can be writen as

$$f_i = f_i e^i$$

$f_i$ are the **covariant** components of $f$.

Given that a tensor is a multilinear map, the **linear operator** can be seen as a case of a (1, 1) type tensor. So for example, a linear operator $T$

$$T(v) = T(v^i e_i) = v^i T(e_i) = v^i T_i^j e_j$$

Remember that we are using the Einstein summation convention so there is actually a double sum in the last expression, $\sum_{i=1}^{n} \sum_{j=1}^{n} = \sum_{i,j}$. Note that we also used the definition of linearity to "take out" the $v^i$ factor from the action of the tensor.

Components of a tensor are the values of the tensor on a set of basis vectors. For example, the components

$$H_i^j = H(e_i, e^j)$$

A vector can be considered a $(0, 1)$ type tensor (it feeds on a dual vector) and its components transform like

$$v^i = A_j^i v^j$$

And a dual vector can be considered a $(1, 0)$ type tensor (it feeds on a vector) and it compoents transform like

$$f_i = A_i^j f_j$$

### 3.3.1 Formal definition

Given two vectors $v \in V$ and $w \in W$, the tensor product $v \otimes w$ is defined as the element of $V \otimes W$ such that

$$(v \otimes w)(h, g) = h(v)g(w)$$

For all $h \in V^*$ and $g \in W^*$.

If we consider some $T \in V \otimes W$ such that

$$T(h, g) = h_i g_j T(e^i, f^j) = h_i g_j T^{ij}$$

We defined $T^{ij}$ as $T^{ij} = T(e^i, f^j)$, $T$ which is a (0,2) tensor, meaning that it's a multilinear map that takes two dual vectors as inputs and returns a scalar. The vectors it takes as inputs are from the dual spaces (or covectors) $V^*$ and $W^*$. $\{e^i\}$ is a basis for $V^*$ and $\{f^j\}$ is a basis for $W^*$. Both sets of dual basis vectors, which is how we know $T^{ij}$ is a type (0,2) tensor.

The next term in our expression is $e_i \otimes f_j$. This tensor product takes two dual vectors as inputs and returns a scalar,

$$(e_i \otimes f_j)(v^*, w^*) = v^*(e_i) w^*(f_j)$$

The action of the tensor product on these dual vectors is given by the action of each dual vector on its corresponding basis vector.

This is exactly the behavior of a (0,2) tensor - it takes two dual vectors as inputs and returns a scalar. $v^*(e_i)$ is a scalar because - a linear functional acting on a vector. Similarly, $w^*(f_j)$.

If we expand the dual vectors $v^*$ and $w^*$ in terms of their respective dual bases:

$$v^* = v_i e^i$$
$$w^* = w_j f^j$$

Then,

$$\begin{aligned}(e_i \otimes f_j)(v^*, w^*) &= v^*(e_i) w^*(f_j) \\ &= v_k e^k(e_i)\, w_l f^l(f_j) \\ &= v_i w_j\end{aligned}$$

This follows the same approach as in the book when deriving the expression after 3.39.

Now let's put this into practice.

Suppose we have two vectors, $\vec{v} = (v_1, v_2)$ and $\vec{w} = (w_1, w_2, w_3)$. The tensor product of these vectors, denoted $\vec{v} \otimes \vec{w}$, is a new object that lives in a higher-dimensional space. It's defined as:

$$\vec{v} \otimes \vec{w} = (v_1 w_1, v_1 w_2, v_1 w_3, v_2 w_1, v_2 w_2, v_2 w_3)$$

This is a 6-dimensional vector, where each component is a product of a component from $\vec{v}$ and a component from $\vec{w}$.

More generally, if $\vec{v}$ is an $n$-dimensional vector and $\vec{w}$ is an $m$-dimensional vector, their tensor product $\vec{v} \otimes \vec{w}$ is an $nm$-dimensional vector.

Now, let's look at the tensor product of two matrices. If $A$ is an $n \times m$ matrix and $B$ is a $p \times q$ matrix, their tensor product $A \otimes B$ is an $np \times mq$ matrix:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{pmatrix}$$

Each block in this block matrix is a scalar multiple of the matrix $B$.

The above is actually a specialized version of the tensor product called the **Kronecker product**, see wikipedia: kronecker products.

In quantum mechanics, the state space of a composite system is the tensor product of the state spaces of the individual systems. If $|\psi\rangle$ is a state vector in one system and $|\phi\rangle$ is a state vector in another system, the state of the

composite system is $|\psi\rangle \otimes |\phi\rangle$ the notation $|\psi\phi\rangle$ is often used as a shorthand, so that the tensor product $|\psi\rangle \otimes |\phi\rangle = |\psi\phi\rangle$, especially when dealing with composite systems.

The tensor product $|\psi\rangle \otimes |\phi\rangle$ is a formal mathematical operation that combines two state vectors into a single state vector in a higher-dimensional space. This is the correct way to describe the state of a composite system in quantum mechanics.

If $|\psi\rangle$ is the state of one particle and $|\phi\rangle$ is the state of another particle, then $|\psi\phi\rangle$ would typically be understood to mean $|\psi\rangle \otimes |\phi\rangle$.

The tensor product itself is not a Hermitian form, but it can be used to construct a Hermitian form.

A Hermitian form on a complex vector space $V$ is a map $h : V \times V \to \mathbb{C}$. A sesquilinear form then it is a map $h : \overline{V} \times V \to \mathbb{C}$.

The universal property of the tensor product is a key concept that characterizes the tensor product in terms of a unique mapping property. The "universal-property definition" of the tensor product of two vector spaces is the following (see wikipedia: tensor product):

The tensor product of two vector spaces $V$ and $W$ is a vector space denoted as $V \otimes W$, together with a bilinear map $\otimes : (v, w) \mapsto v \otimes w$ from $V \times W$ to $V \otimes W$, such that, for every bilinear map $h : V \times W \to Z$, there is a unique linear map $\tilde{h} : V \otimes W \to Z$, such that $h = \tilde{h} \circ \otimes$, that is, $h(v, w) = \tilde{h}(v \otimes w)$ for every $v \in V$ and $w \in W$.

In other words, the tensor product $U \otimes V$ is the "most general" vector space that can be constructed from $U$ and $V$ using a bilinear map. Any bilinear map from $U \times V$ to another vector space $W$ factors uniquely through the tensor product via a linear map from $U \otimes V$ to $W$.

The phrase "up to isomorphism" is a common one in abstract algebra. It means that two mathematical objects (in this case, vector spaces) are not necessarily identical, but they are isomorphic, meaning that there exists an isomorphism between them.

An isomorphism between two vector spaces $V$ and $W$ is a bijective (one-to-one and onto) linear map $f : V \to W$. If such a map exists, we say that $V$ and $W$ are isomorphic, and we write $V \cong W$.

### 3.3.2 Practical definition

Given two vector spaces $V$ and $W$ with basis vectors $e^i$ and $e^j$, respectively, the tensor product of these spaces, denoted as $V \otimes W$, is **a new vector space spanned** by the basis vectors $e^i \otimes e^j$ for all possible combinations of $i$ and $j$.

For example, if $V$ is a 2D space with basis vectors $\{e^a, e^b\}$ and $W$ is a 3D space with basis vectors $\{e^1, e^2, e^3\}$, then the tensor product space $V \otimes W$ will

be a 6D space with basis vectors:

$$\{e^a \otimes e^1, e^a \otimes e^2, e^a \otimes e^3, e^b \otimes e^1, e^b \otimes e^2, e^b \otimes e^3\}$$

To understand this further, let's look at $e^a \otimes e^1$. Consider the original vector spaces $V$ and $W$.

In the 2D space $V$, $e^a$ might represent the unit vector along the $x$-axis, $(1, 0)$. In the 3D space $W$, $e^1$ might represent the unit vector along the $x$-axis, $(1, 0, 0)$. Remember that when we say vectors, we are actually thinking about column vectors. And it turns out the way to obtain an explicit answer is through the kronecker product again. In this case, our tensor product can be computed as an **outer product**. Remember that the outer product of two vectors $x$ and $y$ is defined as $xy^\dagger$, or simply $xy^T$, if we are dealing with solely real numbers. This is in contrast with an inner product which would be $x^\dagger y$ or $x^T y$.

So now we have a great comparison between non-degenerate Hermitian forms and outer product. (Remember that for a Hermitian form to be an iner product we need the Hermitian form to be positive definite, that is $(x|y) = \langle x|y\rangle$ if $(x|y) > 0$, seewikipedia: inner product space.)

$$u \otimes v = uv^\dagger$$

and

$$\langle u, v\rangle = \langle u|v\rangle = u^\dagger v$$

So when we go on talking about components as in

$$u \otimes v = (u^i e_i) \otimes (v^j b_j) = u^i v^j e_i \otimes b_j$$

It means that

$$(u \otimes v)_{ij} = u^i v^j$$

That is, the tensor product can be pieced together through normal products between components.

For example,

$$u \otimes V = uv^\dagger = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix} = \begin{pmatrix} u_1v_1 & u_1v_2 & \dots & u_1v_m \\ u_2v_1 & u_2v_2 & \dots & u_2v_m \\ \vdots & \vdots & \ddots & \vdots \\ u_nv_1 & u_nv_2 & \dots & u_nv_m \end{pmatrix}$$

In practice, $e^a \otimes e^1$ can be thought of as a 6D unit vector that points in the direction corresponding to the combination of the x-axis direction from space $V$ and the x-axis direction from space $W$. It is a basis vector that allows you to describe vectors in the 6D space that have components along this combined direction.

Using what we noted above,

$$e^a \otimes e^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \end{pmatrix} = \begin{pmatrix} (1)(1) & (1)(0) & (1)(0) \\ (0)(1) & (0)(0) & (0)(0) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Here is where we make our comback to quantum mechanics. If $|\psi\rangle \in V$ and $|\phi\rangle \ in W$, then a member of the tensor product $V \otimes W$ can be written as

$$|\psi\rangle |\phi\rangle = |\psi, \phi\rangle = |\psi\phi\rangle = |\psi\rangle \otimes |\phi\rangle$$

See wikipedia: braket notation - composite bras and kets.

There is one more interesting fact to know! The tensor product $e^i \otimes e^i$, when $\{e^i\}$ represents an orthonormal basis set, will always result in a diagonal matrix. Since it is the product of a matrix with its transpose.

## 3.4 Tensor Contraction

Tensor contractions are used to reduce the rank of a tensor by summing over pairs of indices. A more useful way to think of tensor contractions is an operations that can be used to raise or lower indices when the tensor being contracted with is a metric tensor. In general relativity, the metric tensor $g_{ij}$ is used to raise or lower indices:

$$v^i = g^{ij} v_j$$

$$v_i = g_{ij} v^j$$

Here, $g^{ij}$ is the inverse of the metric tensor $g_{ij}$.

We already saw this example when we began talking about tensors and about the inner product, 3.

An even simpler example, of tensor contraction would be the raising or lowering of indices using the metric tensor in a Euclidean space with a Cartesian coordinate system.

In this case, the metric tensor is just the Kronecker delta, denoted as $\delta_{ij}$, where $i$ and $j$ are the indices.

The Kronecker delta is defined as:

$$\begin{cases} \delta^i_j = 1 & \text{if } i = j \\ \delta^i_j = 0 & \text{if } i \neq j \end{cases}$$

Now, let's say we have a vector $V^i$ (contravariant vector) and we want to find its covariant counterpart $V_i$. We can do this by contracting the vector with the metric tensor:

$$V_i = \delta^i_j V^j$$

Here, we sum over the repeated index $j$, which is the essence of tensor contraction. This operation lowers the index of the vector, transforming it from a contravariant vector to a covariant vector. Similarly, if we have a covariant vector $V_i$ and want to find its contravariant counterpart $V^i$, we can use the

inverse metric tensor $\delta^{ij}$ (which in this case is the same as the metric tensor itself) to raise the index:

$$V^i = \delta^i_j V_j$$

Again, we sum over the repeated index $j$ to perform the contraction.

And note how $\delta^i_j$ is the inverse of $\delta^j_i$.

The summation over indices is what's shown in the book with the more formal syntax.

Given $T \in T^r_s(V)$ which has components $T^{j_1,\ldots,j_s}_{i_1,\ldots,1_r}$ it forms a tensor product of the form

$$V^*_r \times \cdots \times V^*_r \times V_s \times \cdots \times V_s$$

so that,

$$T = T^{j_1,\ldots,j_s}_{i_1,\ldots,1_r} e^{i_1} \otimes \cdots \otimes e^{i_r} \otimes e_{j_1} \otimes \cdots \otimes e_{j_s}$$

The reason that the tensor components have $r$ products of $V^*$ and $s$ product of $V$, is because taking the tensor product with their basis vectors will produce a scalar value, and actual component. And because it allows us to visualize the procedure of a contraction as feeding $e_i$ into the $r$-th slot and $e^i$ into the $(r+s)$-th slot.

Sicne the $r$-th slot corresponds to the basis vector $e^{i_r}$, and the $(r+s)$-th slot corresponds to the basis vector $e_{j_s}$.

$$T^{j_1,\ldots,j_{s-1},i}_{i_1,\ldots,i_{r-1},i} = T^{j_1,\ldots,j_s}_{i_1,\ldots,i_r} e^{i_1} \otimes \cdots \otimes e^{i_{r-1}} \otimes e^i \otimes e_{j_1} \otimes \cdots \otimes e_{j_{s-1}} \otimes e_i$$

Then perform the summation over the repeated index $i$.

Which as we saw in 3, the $e^i \otimes e_i$ component results in a $\delta^i_i$ like term, that then makes sure that only a set of matching indices is kept in sync, while the other indices are used in a summation (one set of indices is kept in sync, the others vary from 0 to $N$ and are then summed at the end).

## 3.5 Frobenius Method

### 3.5.1 A Generic Example

The following example comes from wikipedia: frobenius method:

$$z^2 f'' - z f' + (1 - z) f = 0$$

We assume the solution has the form,

$$f = \sum_{k=0} A_k z^{k+r}$$

Plugging the above series into the ODE and organizing like terms we land in

$$\sum_k \left[ (k+r)(k+r-1) - (k+r) + (1-z) \right] A_k z^{k+r} = 0$$

which can be further simplified to

$$\sum_k \left[ (k+r)(k+r-1) - (k+r) + 1 \right] A_k z^{k+r} - \sum_k A_k z^{k+r+1} = 0$$

The indicial polynomial can then be obtained by considering the lowest power of $z$, which is $z^r$,

$$(r(r-1)+1) A_0 z^r$$

We now set $z = 0$ and obtain

$$r(r-1) + 1 = r^2 - r + 1 = (r-1)^2 = 0$$

Meaning that $r = 1$ (double root).

We now need to find the recurrence relation by plugging $f = \sum_{k=0} A_k z^{k+1}$ back into our original proble, equate coefficients of like powers of $z$, and find a recurrence relation for $A_k$.

We now have,

$$f' = \sum_{k=0} A_k (k+1) z^k$$

and

$$f'' = \sum_{k=0} A_k k (k+1) z^{k-1}$$

And our ODE now becomes,

$$\begin{aligned}
& z^2 \sum_k A_k k(k+1) z^{k-1} - z \sum_k A_k (k+1) z^k + (1-z) \sum_k A_k z^{k+1} \\
&= \sum_k A_k k(k+1) z^{k+1} - \sum_k A_k (k+1) z^{k+1} + \sum_k A_k z^{k+1} - \sum_k A_k z^{k+2} \\
&= \sum_k \left[ k(k+1) - (k+1) + 1 \right] A_k z^{k+1} - \sum_k A_k z^{k+2} \\
&= \sum_k k^2 A_k z^{k+1} - \sum_k A_k z^{k+2}
\end{aligned}$$

To find a recurrence relation we need to align the sums. So we can write

$$\sum_{k=0} k^2 A_k z^{k+1} - \sum_{k=0} A_k z^{k+2}$$
$$= \sum_{k=0} k^2 A_k z^{k+1} - \sum_{k=1} A_{k-1} z^{k+1}$$

We could do a similar shift for the first sum but since the first term $k = 0$ has a $k^2$ term then we can ignore it. Thus,

$$\sum_{k=1} \left(k^2 A_k - A_{k-1}\right) z^{k+1} = 0$$

Since this sum must be zero for all $z$, the coefficients must satisfy,

$$k^2 A_k - A_{k-1} = 0$$

or

$$A_k = \frac{A_{k-1}}{k^2}$$

This recurrence realtionship now gives us all coefficients in terms of $A_0$. The first couple coefficients are,

$$A_1 = \frac{A_0}{1^2} = A_0$$

$$A_2 = \frac{A_1}{2^2} = \frac{A_0}{2^2}$$

$$A_3 = \frac{A_2}{3^2} = \frac{A_0}{2^2 3^2}$$

This means that our general solution looks like

$$f(z) = \sum_{k=0} A_k z^{k+1} = A_0 \sum_{k=0} \frac{z^{k+1}}{k!^2}$$

### 3.5.2 Legendre Equation

Using the nomenclature of Arfken,

$$\mathcal{L}y(x) = -(1 - x^2)y'' + 2xy' = \lambda y$$

Which can also be expressed as

$$(1 - x^2)y'' - 2xy' + \lambda y = 0$$

If we assume the solution is of the form

$$y(x) = \sum_{j=0} a_j x^{j+s}$$

then the Legendre equation can be written as

$$
\begin{aligned}
&(1-x^2)\sum_j a_j(j+s)(j+s-1)x^{j+s-2} - 2x\sum_j a_j(j+s)x^{j+s-1} + \lambda\sum_j a_j x^{j+s} \\
&= \sum_j a_j(j+s)(j+s-1)x^{j+s-2} - \sum_j a_j(j+s)(j+s-1)x^{j+s} - 2\sum_j a_j(j+s)x^{j+s} + \lambda\sum_j x^{j+s} \\
&= \sum_j a_j(j+s)(j+s-1)x^{j+s-2} + \sum_j \left[-(j+s)(j+s-1) - 2(j+s) + \lambda\right] a_j x^{s+j}
\end{aligned}
$$

The indicial polynomial then comes from considering the lowest power of $x$, which is the first sum

$$a_j(j+s)(j+s-1)x^{j+s-2}$$

which for an arbitrary $x$ when $j=0$, it becomes

$$s(s-1)=0$$

whose solutions are $s=0$ and $s=1$.

In this case, the solution will be a linear combination,

$$y(x) = y_1(x) + y_2(x) = \sum_j a_j x^j + \sum_j b_j x^{j+1}$$

Let's start with the $s=0$ case then. Plugging $y_1(x)$ back into our ODE we have

$$
\begin{aligned}
&(1-x^2)\sum_j a_j j(j-1)x^{j-2} - 2x\sum_j a_j j x^{j-1} + \lambda\sum_j a_j x^j \\
&= \sum_j a_j j(j-1)x^{j-2} - \sum_j a_j j(j-1)x^j - 2\sum_j a_j j x^j + \lambda\sum_j a_j x^j \\
&= \sum_j a_j j(j-1)x^{j-2} + \sum_j \left[-j(j-1) - 2j + \lambda\right] a_j x^j
\end{aligned}
$$

Now, to align the powers of $x^j$. First, note that the first two terms in $\sum_j a_j j(j-1)x^{j-2}$ are 0. So we might as well start at $j=2$

$$\sum_{j=2} a_j j(j-1)x^{j-2}$$

If we shift it back to 0, we can align the powers of $x$.

$$\sum_{j=2} a_j j(j-1)x^{j-2} = \sum_{j=0} a_{j+2}(j+2)(j+1)x^j$$

This gives us,

$$\sum_{j=0} a_{j+2}(j+2)(j+1)x^j + \sum_j \left[-j(j-1) - 2j + \lambda\right] a_j x^j = 0$$

Now, we can equate coefficients,

$$a_{j+2}(j+2)(j+1) + [-j(j-1) - 2j + \lambda]\, a_j = 0$$

or

$$a_{j+2} = \frac{j(j-1) + 2j - \lambda}{(j+1)(j+2)} a_j = \frac{j(j+1) - \lambda}{(j+1)(j+2)} a_j$$

In the $s = 1$ case, if we plug in $y_2$ into our ODE,

$$\begin{aligned}
&(1-x^2)\sum_j a_j j(j+1)x^{j-1} - 2x\sum_j a_j(j+1)x^j + \lambda\sum_j a_j x^{j+1} \\
&= \sum_j a_j j(j+1)x^{j-1} - \sum_j a_j j(j+1)x^{j+1} - 2\sum_j a_j(j+1)x^{j+1} + \lambda\sum_j a_j x^{j+1} \\
&= \sum_j a_j j(j+1)x^{j-1} + \sum_j [-j(j+1) - 2(j+1) + \lambda]\, a_j x^{j+1}
\end{aligned}$$

Following the procedure we have been building: look for a series with redundant terms, and then try to shift it to match the other powers of $x$. Here we have

$$\sum_{j=0} a_j j(j+1)x^{j-1}$$

that can be written as

$$\sum_{j=1} a_j j(j+1)x^{j-1}$$

Since the term when $j = 0$ equals to 0. And we can conveniently shift this series back one to match the $x^j$ terms of the other sum:

$$\sum_{j=1} a_j j(j+1)x^{j-1} = \sum_{j=0} a_{j+1}(j+1)(j+2)x^j$$

At this point we have,

$$\sum_{j=0} a_{j+1}(j+1)(j+2)x^j + \sum_j [-j(j+1) - 2(j+1) + \lambda]\, a_j x^{j+1}$$

We need to do another shift and doing it on the first sum seems like the simplest approach. If we do so, then we can arrive at

$$\sum_{j=-1} a_{j+2}(j+2)(j+3)x^{j+1} + \sum_j [-j(j+1) - 2(j+1) + \lambda]\, a_j x^{j+1}$$

But let's separate the first term of the first sum. When $j = -1$, we get $a_1(1)(2)x^0 = 2a_1$. Thus,

$$2a_1 + \sum_{j=0} a_{j+2}(j+2)(j+3)x^{j+1} + \sum_j \left[-j(j+1) - 2(j+1) + \lambda\right] a_j x^{j+1} = 0$$

And this equation must hold for all $x$, so the coefficients of each power of $x$ must be zero. This means that $a_1 must be 0$. And we are left with

$$a_{j+2}(j+2)(j+3) + \left(-j(j+1) - 2(j+1) + \lambda\right) a_j = 0$$

Which simplifies to

$$a_{j+2} = \frac{j^2 + 3j + 2 - \lambda}{(j+2)(j+3)} a_j = \frac{(j+1)(j+2) - \lambda}{(j+2)(j+3)} a_j$$

In order to obtain the eignevalues we need to go back to the beginning. If you remember, this is how the problem was initially framed:

$$\mathcal{L}y(x) = -(1-x^2)y'' + 2xy' = \lambda y$$

So whatever we can do to find a $\lambda$ will give us our eigenvalues. And that is why Arfken and other resources look for the conditions that make both of the recursion relations result in a series solution that converges. In the first case we need a value of lambda that will cancel the numerator $j(j+1)$ and in the second we need a value that will cancel the $(j+1)(j+2)$.

# Chapter 4

# Variational Calculus

Normed linear space - we want to define "closeness" via a norm and continuity to study functionals that look like

$$J[y] = \int_a^b F(x, y, y')dx$$

where the appropiate function space is the set of all functions with two continuous derivatives.

#### 4.0.0.1 A Propagator

Describes the amplitude for a particle to propagate from one point to another.

With a source $j(\mathbf{x}, t)$, the Klein-Gordon equation becomes

$$\left(\partial^2 + m^2\right) \psi(\mathbf{x}, t) = j(\mathbf{x}, t)$$

In momentum space the equation is

$$\left(-\omega^2 + |\mathbf{k}|^2 + m^2\right) \tilde{\psi}(\mathbf{k}, \omega) = \tilde{j}(\mathbf{k}, \boldsymbol{\omega})$$

$$\tilde{\psi}(\mathbf{k}, \omega) = \frac{\tilde{j}(\mathbf{k}, \boldsymbol{\omega})}{\omega^2 - |\mathbf{k}|^2 - m^2 + i\epsilon}$$

$i\epsilon$ Feynman's prescription.

The propagator $D_F(x - y)$ is the inverse transform of $\frac{1}{\omega^2 - |\mathbf{k}|^2 - m^2 + i\epsilon}$

$$D_F(x - y) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{-i(\mathbf{k}\cdot\mathbf{x} - \omega t)}}{\omega^2 - |\mathbf{k}|^2 - m^2 + i\epsilon}$$

# Chapter 5

# Stats

## 5.1 Biased and Unbiased Estimators

The following expressions come from [GBC16] section 5.4.2, which starts at page 122.

Computing the bias of the sample mean:

$$\begin{aligned}
\text{bias}(\hat{\mu}_m) &= \mathbb{E}\left[\hat{\mu}_m\right] - \mu \\
&= \mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m} x^{(i)}\right] - \mu \\
&= \left(\frac{1}{m}\sum_{i=1}^{m} \mathbb{E}\left[x^{(i)}\right]\right) - \mu \\
&= \left(\frac{1}{m}\sum_{i=1}^{m} \mu\right) - \mu \\
&= \mu - \mu = 0
\end{aligned}$$

Then we test the sample variance $\hat{\sigma}_m^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x^{(i)} - \hat{\mu}_m\right)^2$ to see whether it is a biased or an unbiased estimator.

To find the answer, we will need to evaluate the following expression:

$$\mathbb{E}\left[\hat{\sigma}_m^2\right] = \mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m}\left(x^{(i)} - \hat{\mu}_m\right)^2\right] \tag{5.1.1}$$

$$= \frac{m-1}{m}\sigma^2 \tag{5.1.2}$$

where $\hat{\mu}_m = \frac{1}{m}\sum_{i=1}^{m} x^{(i)}$ is the **sample mean**.

To carry out the above computation it helps if we go step by step.

Let's start with the following:

$$\mathbb{E}\left[\sum_{i=1}^{m}\left(x^{(i)}-\hat{\mu}_m\right)^2\right]$$

There is actually a "well known" expression that says that

$$\mathbb{E}\left[\sum_{i=1}^{m}\left(x^{(i)}-\hat{\mu}_m\right)^2\right]=\sum_{i=1}^{m}\left(\mathbb{E}\left[\left(x^{(i)}\right)^2\right]-n\mathbb{E}\left[(\hat{\mu}_m)^2\right]\right)$$

In more conventional notation, this is written as

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]&=\sum_{i=1}^{n}\left(\mathbb{E}\left[X_i^2\right]-\mathbb{E}\left[\overline{X}^2\right]\right)\\&=\sum_{i=1}^{n}\left(\mathbb{E}\left[X_i^2\right]\right)-n\mathbb{E}\left[\overline{X}^2\right]\end{aligned}$$

We will continue with this notation and then switch back at the end because the ML notation is to verbose. But anyway, the trick to seeing why the above result is such is to expand the sum:

$$\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]=\mathbb{E}\left[\left(X_1-\overline{X}\right)^2+\left(X_2-\overline{X}\right)^2+\left(X_3-\overline{X}\right)^2+\dots\right]$$

Let's now expands the terms and group them together,

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]&=\mathbb{E}\left[\left(X_1-\overline{X}\right)^2+\left(X_2-\overline{X}\right)^2+\left(X_3-\overline{X}\right)^2+\dots\right]\\&=\mathbb{E}\left[\left(X_1^2-2X_1\overline{X}+\overline{X}^2\right)+\left(X_2^2-2X_2\overline{X}+\overline{X}^2\right)+\left(X_3^2-2X_3\overline{X}+\overline{X}^2\right)+\dots\right]\\&=\sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]+\mathbb{E}\left[-2\left(\sum_{i=1}^{n}X_i\right)\overline{X}\right]+n\mathbb{E}\left[\overline{X}^2\right]\end{aligned}$$

Now, remember that the sample mean is defined as $\hat{\mu}_m\equiv\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$.
So the middle term can be simplified:

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right] &= \sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]+\mathbb{E}\left[-2\left(\sum_{i=1}^{n}X_i\right)\overline{X}\right]+n\mathbb{E}\left[\overline{X}^2\right]\\ &= \sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]+\mathbb{E}\left[-2n\overline{XX}\right]+n\mathbb{E}\left[\overline{X}^2\right]\\ &= \sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]+\mathbb{E}\left[-2n\overline{X}^2\right]+n\mathbb{E}\left[\overline{X}^2\right]\end{aligned}$$

Since expectation do not depend on the number of samples nor on constant coefficients, the above can be rewritten as:

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right] &= \sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]+\mathbb{E}\left[-2n\overline{X}^2\right]+n\mathbb{E}\left[\overline{X}^2\right]\\ &= \sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]+\mathbb{E}\left[-2n\overline{X}^2+n\overline{X}^2\right]\\ &= \sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]+\mathbb{E}\left[-n\overline{X}^2\right]\\ &= \sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]-n\mathbb{E}\left[\overline{X}^2\right]\end{aligned}$$

Which is the result we wanted to prove.

But there is one more trick! If one just happened to remember that the **population variance** is defined as

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}\left[(X-\mu)^2\right]\\ &= \mathbb{E}\left[X^2\right]-\mathbb{E}\left[X\right]^2\end{aligned}$$

Check wikipedia out, wiki:variance.

If we use this result, we can write:

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right] &= \sum_{i=1}^{n}\left(\mathbb{E}\left[X_i^2\right]\right)-n\mathbb{E}\left[\overline{X}^2\right]\\ &= \sum_{i=1}^{n}\left(\mathrm{Var}\left[X_i\right]+\mathbb{E}\left[X_i\right]^2\right)-n\left(\mathrm{Var}\left[\overline{X}\right]+\mathbb{E}\left[\overline{X}\right]^2\right)\end{aligned}$$

Note that the pattern here is that we are trying to figure out the expectation value of a quantity squared! This is a "trick" one might want to remember!

Now,

$$\text{Var}\left[X_i\right] + \mathbb{E}\left[X_i\right]^2 = \sigma^2 + \mu^2$$

As per the other terms, we have to remember the cetral limit theorem and the sampling distribution of the mean. For example, see [Tri18], Chapter 6-4 on page 267,

$$\mathbb{E}\left[\overline{X}\right] = \mu$$

and

$$\text{Var}\left[\overline{X}\right] = \frac{\sigma^2}{n}$$

Plugging these two into our expression, we get:

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right] &= \sum_{i=1}^{n}\left(\text{Var}\left[X_i\right] + \mathbb{E}\left[X_i\right]^2\right) - n\left(\text{Var}\left[\overline{X}\right] + \mathbb{E}\left[\overline{X}\right]^2\right) \\
&= \sum_{i=1}^{n}\left(\sigma^2 + \mu^2\right) - n\left(\frac{\sigma^2}{n} + \mu^2\right) \\
&= n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2 \\
&= \left(n-1\right)\sigma^2
\end{aligned}$$

If we plug this into 5.1.1, we can see how [GBC16] obtains the result it has for the bias of the sample variance.

The above compuations came from [Spi].

# Chapter 6

# Lattic Simulations of a $\phi^4$ scalar Theory

This document is a guide to help anyone wanting to learn quantum field theory.

This document has come about because I spent a way too much time trying to learn QFT. There are plenty of great books out there but they only made sense after I had spent a very considerable time inmersed in mathematics, engineering, and other sciences.

The one thing that always caused me issues is that almost all textbooks or calss notes build the subject from the ground up - the same way any other subject is taught in school. You probably even thought "duh!" while reading that previous sentence. But recent studies, see [Gra23], made me rethink this approach for QFT! So instead of picking up from whatever may have been taught in your last QM or EM course, I will start with an actual non-trivial research problem. This will have you doing research type work right away. Then we will work backwards and provide you with the background you need!

For a better argument as to why, definetely go and read [Gra23], but the TL;DR is: starting with a real problem right away will help you start building the experience for what matters and what tools are applied when. When we go from the ground up, there is the unspoken rule that you agically have to remember some key fact or formula that someone mentioned years ago and magically know how to apply it to your problem now. Here, we won't have that issue.

We will begin by building upon the work of [Sch], which is one of the few works available that is more or less self-contained but also provides some of the code used for its simulations. A lot of the code snippets in [Sch] does not actually work as in, so we fixed it and made it available in [Ala].

## 6.1 Metropolis Algorithm for the Ising Model

Single site is updated at a time and the probability of generating such as state is $1/N$.

The detailed condition for Markov Chains Monte Carlos is

$$\frac{P(\mu \to \nu)}{P(\nu \to \mu)} = \frac{g(\mu \to \nu)A(\mu \to \nu)}{g(\nu \to \mu)A(\mu \to \nu)} = \frac{A(\mu \to \nu)}{A(\mu \to \nu)}$$

Where $g$ is the probability of generating a given state and $A$ is the probability of accepting the transition.

Remember that the more stringent expression of detailed balance to guarantee static equilibirum is written as

$$p_m u P(\mu \to \nu) = p_n u P(\nu \to \mu)$$

and since the probabilities $p_m u$ and $p_n u$ are $e^{-\beta E_m u}/Z$ and $e^{-\beta E_n u}/Z$ respectively,

$$\frac{P(\mu \to \nu)}{P(\nu \to \mu)} = \frac{A(\mu \to \nu)}{A(\mu \to \nu)} = \frac{p_\nu}{p_\mu} = e^{-\beta(E_\nu - E_\mu)}$$

Note that if $E_\nu \leq E_\mu$, then $e^{-\beta(E_\nu - E_\mu)}$, then the argument of the exponent will be zero or positive, so $e^{+(\dots)} \geq 1$. And if $E_\nu > E_\mu$, then $e^{-\beta(E_\nu - E_\mu)}$ will looke like $e^{-(\dots)} < 1$. Which intuitevely makes sense, transitioning to a state $\nu$ from a state $\mu$ is favored when the end state $\nu$ has a lower energy (rememeber that the exponents are proportional to the probability of the given state).

Thus, the Metropolis algorithm

$$A(\mu \to \nu) = \begin{cases} 1 & \text{if } E_\nu \leq E_\mu \quad (\Delta E \leq 0) \\ e^{-\beta(E_\nu - E_\mu)} = e^{-\beta \Delta E} & \text{if } E_\nu > E_\mu \quad (\Delta E > 0) \end{cases}$$

## 6.2 Metropolis Algorithm for a Scalar Field

We will begin by building upon the work of [Sch], which is one of the few works available that is more or less self-contained and provides some of the code used to obtain its results. A lot of the code snippets in [Sch] do not actually work as-is, so we fixed it and made it available in [Ala].

For the metropolis steps, we need to be able to compute the energy of our system in order to determine whether we accept or reject a change.

As we saw in the Isisng model example, in order to meet the more rigurous condition of detailed balance we need to probabilitically accept a change according to

$$A(\mu \to \nu) = \begin{cases} 1 & \text{if } E_\nu \leq E_\mu \quad (\Delta E \leq 0) \\ e^{-\beta(E_\nu - E_\mu)} = e^{-\beta \Delta E} & \text{if } E_\nu > E_\mu \quad (\Delta E > 0) \end{cases}$$

And this is where we will take a more research-based pedagogy and work backwards - as opposed to sending you off to learn QFT, quantum mechanics, electrodynamics, etc, etc.

The $\phi^4$ theory in four-dimensional Euclidean space has the following expression for its action

$$S_E = \int d^4x_E \mathcal{L}_E = \int d^4x \left[\frac{1}{2}\left(\partial_{E_\mu}\phi\right)^2 + \frac{1}{2}\mu^2\phi^2 + \frac{1}{4!}\lambda\phi^4\right]$$

And it turns out that in Euclidean space the action, or Euclidean action, has the same units as the energy, so from here on out we will refer to the above as the energy.

Just to start out simple, we will work in 2D first. And we will:

- Discretize $\phi$: the continuous function $\phi(x_\mu)$ will be represented by $\phi_n$ for $0 \leq n \leq N$, on an Lattice of zise $N$ and spacing $a$ between points.
- Discretize the derivatives: for example, $\partial\phi/\partial x$ becomes $\frac{1}{a}\left(\phi_n(x+a,t) - \phi_n(x,t)\right)$
- Discretize the integral: $\int dx f$ becomes $\sum_i^N f(n_i)a$

All to define the operations we need in a lattice. Following these steps, and again, starting in 2D, our Euclidean action, the energy, first becomes

$$S_E^{(2)} = \int dxdy \left[\frac{1}{2}\left[\left(\frac{\partial\phi}{\partial t}\right)^2 + \left(\frac{\partial\phi}{\partial x}\right)^2\right] + \frac{1}{2}\mu^2\phi^2 + \frac{1}{4!}\lambda\phi^4\right]$$

Note the kinetic term! But before that, let's look into tensors, because I never saw a class where this material was clearly covered. Most of my physics courses just magically skipped over this or gave you the results. And the math courses I found dove right into the "analysis" (which is pretty damn important but it is not where one ought to start.)

## 6.3 Tensors

Working with a spacetime metric signature $(+---)$,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$x^\mu = (x^0, \mathbf{x})$ and $x_\mu = g_{\mu\nu}x^\nu = (x^0, -\mathbf{x})$. To see this explciitly, let's carry out the summations.

When $\mu = 0$, $x_0 = g_{00}x^0 + g_{01}x^1 + g_{02}x^2 + g_{03}x^3 = x^0 + 0 + 0 + 0 = x^0$. For $\mu = 1$, $x_1 = g_{10}x^0 + g_{11}x^1 + g_{12}x^2 + g_{13}x^3 = 0 + x^1 + 0 + 0 = -x^0$. Similarly, when $\mu = 2$, $x_2 = -x^2$, and when $\mu = 3$, $x_3 = -x^3$.

### 6.3.1 Inner Product

Why does $p \cdot x = g_{\mu\nu} p^\mu x^\nu = g^{\mu\nu} p_\mu x_\nu = p_\mu x^\nu = p^\mu x_\nu = p^0 x^0 - \mathbf{p} \cdot \mathbf{x}$?

The metric tensor defines the geometry of spacetime, including the way distances and angles are measured. The implied summation, einstein summation, effectively 'weights' the components according to the geometry of spacetime.

$g_{\mu\nu}$ lowers indices, while $g^{\mu\nu}$ raises them. $\delta^\mu_\nu$ is a diagonal matrix with ones on the diagonal and zeros elsewhere. It selects the $\mu$-th component when used in a summation, acting like an identity element.

$g^{\mu\alpha} g_{\alpha\nu} = \delta^\mu_\nu$

$$g^{\mu\alpha} g_{\alpha\nu} = \sum_{\alpha=0} g^{\mu\alpha} g_{\alpha\nu} = g^{\mu 0} g_{0\nu} + g^{\mu 1} g_{1\nu} + g^{\mu 2} g_{2\nu} + g^{\mu 3} g_{3\nu}$$

This is essentially a matrix multiplication.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Going component by component, take the second row and second column, $\mu = 1$ and $\nu = 1$. The element for that entry is given by

$$g^{1\alpha} g_{\alpha 1} = g^{10} g_{01} + g^{11} g_{11} + g^{12} g_{21} + g^{13} g_{31} = 0 + (-1)(-1) + 0 + 0 = 1$$

### 6.3.2 Derivatives

The examples here come from [Kay11]. The aim here is to see a few examples to get used to the quantities we are going to be seeing later on.

Some useful identities before we carry on:

1.
$$\frac{\partial x_p}{\partial x_q} = \delta_{pq}$$

**Example 1**
If $a_{ij}$ are contants, to compute the partial derivative

$$\frac{\partial}{\partial x_k} a_{ij} x_i x_j$$

We'll firs write out the sums explciitly.

$$a_{ij} x_i x_j = \sum_{i,j} a_{ij} x_i x_j$$

Someone once happened to noticed that the sum could be broken into parts depending on when $x_{\{i,j\}} = x_k$, where $x_k$ is the one we want to differentiate with respect to,

$$\begin{aligned}\sum_{i,j} a_{ij}x_ix_j &= \sum_{i\neq k, j\neq k} a_{ij}x_ix_j + \sum_{i,j\neq k} a_{ij}x_ix_j + \sum_{i\neq k,j} a_{ij}x_ix_j + \sum_{i=k,j=k} a_{ij}x_ix_j \\ &= C + \sum_{j\neq k} \left(a_{kj}x_j\right) x_k + \sum_{i\neq k} \left(a_{ik}x_i\right) x_k + a_{kk}x_kx_k \\ &= C + \sum_{j\neq k} \left(a_{kj}x_j\right) x_k + \sum_{i\neq k} \left(a_{ik}x_i\right) x_k + a_{kk}(x_k)^2\end{aligned}$$

We wrote the first sum, where no $x$s are equal to $x_k$ as $C$, because as you may guess this is going to go away with any derivative.

Back to our original task,

$$\begin{aligned}\frac{\partial}{\partial x_k} a_{ij}x_ix_j &= 0 + \sum_{j\neq k} a_{kj}x_j + \sum_{i\neq k} a_{ik}x_i + 2a_{kk}x_k \\ &= \sum_j a_{kj}x_j + \sum_i a_{ik}x_i \\ &= a_{kj}x_j + a_{ik}x_i \\ &= a_{ki}x_i + a_{ik}x_i \\ &= (a_{ik} + a_{ki})x_i\end{aligned}$$

**Example 2**

Another way to work out

$$\frac{\partial}{\partial x_k} a_{ij}x_ix_j$$

goes as follows.

$$\begin{aligned}\frac{\partial}{\partial x_k} a_{ij}x_ix_j &= a_{ij}\frac{\partial}{\partial x_k} x_ix_j \\ &= a_{ij}\left(\frac{\partial x_i}{\partial x_k}x_j + x_i\frac{\partial x_j}{\partial x_k}\right) \\ &= a_{ij}\left(\delta_{ik}x_j + x_i\delta_{jk}\right) \\ &= a_{kj}x_j + a_{ik}x_i \\ &= (a_{ik} + a_{ki})x_i\end{aligned}$$

**Example 3**

If $a_{ij} = a_{ji}$ are constants, compute

$$\frac{\partial^2}{\partial x_k \partial x_l} a_{ij}x_ix_j$$

Let's start with our previous result,

$$\begin{aligned}\frac{\partial^2}{\partial x_k \partial x_l} a_{ij} x_i x_j &= \frac{\partial}{\partial x_k}\left[\frac{\partial}{\partial x_l} a_{ij} x_i x_j\right] \\ &= \frac{\partial}{\partial x_k}(a_{il} + a_{li}) x_i \\ &= 2a_{li}\frac{\partial x_i}{\partial x_k} \\ &= 2a_{li}\delta_{ik} \\ &= 2a_{lk}\end{aligned}$$

### 6.3.3 Changing Variables: the Jacobian

Promise we are not doing the conventional thing where we go from scratch but there is one more pattern we have to covered now in order for some other equations to make sense and this is the Jacobian.

There is a result from calculus that is presented a bit like this (the following presentation comes from [Daw]):

If we define the Jacobian of a transformation $x = g(u, v)$, $y = h(u, v)$ as

$$\begin{aligned}\frac{\partial(x, y)}{\partial(u, v)} &= \begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix} \\ &= \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\end{aligned}$$

Then if we want to integrate $f(x)$ over some surface $R$, this region will become some other surface $S$ following our transformation and the integral can be computed as follows:

$$\int\int_R f(x, y) dA = \int\int_S f\left(g(u, v), h(u, v)\right) \left|\frac{\partial(x, y)}{\partial(u, v)}\right| d\overline{A}$$

For example, when we go from Euclidean to polar coordinates, the transformation we use is $x = r\cos\theta$ and $y = r\sin\theta$ and the Jacobian is $r$ so $dA = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| dxdy = rdrd\theta$:

$$\begin{aligned}\frac{\partial(x,y)}{\partial(r,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\ &= r\cos^2\theta + r\sin^2\theta \\ &= r\end{aligned}$$

In 3-dimensions this extends to

$$\int\int\int_V f(x,y,z)dV = \int\int\int_B f\left(g(u,v,w), h(u,v,w), k(u,v,w)\right)\left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| d\overline{V}$$

For the case of the transformation of Euclidean coordinates to spherical coordinates, the transformation of variables is $x = r\sin\phi\cos\theta$, $y = r\sin\phi\sin\theta$, and $z = r\cos\phi$. In that case the Jacobian is $r^2\sin\phi$ and thus $dV = \left|\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)}\right| dxdydz = r^2\sin\phi dr d\phi d\theta$.

Besides being very widely used, the reason we mentioned it is because this is also that is very well used in [Tom14] to introduce the notation we see so much in QFT. Specifically, the Jacobian as we have been writing it, can also be seen in expressions such as

$$\overline{a}^\mu = \left(\frac{\partial \overline{x}^\mu}{\partial x^\nu}\right) a^\nu$$

$$\frac{\partial \phi}{\partial \overline{x}^\mu} = \left(\frac{\partial x^\nu}{\partial \overline{x}^\mu}\right)\frac{\partial \phi}{\partial x^\nu}$$

The determinant of the Jacobian matrix is used specifically when you are dealing with volume transformations (like integrating over a volume or changing the measure of integration) because it scales the volume elements according to how the coordinate transformation stretches or compresses the space.

When transforming vectors, however, you're not scaling a volume but rather reorienting or rescaling each component of the vector according to how the coordinate axes themselves have changed. This is why the transformation involves a sum over the product of vector components and the appropriate Jacobian elements:

- Contravariant vectors (standard vectors): The components of these vectors are transformed by multiplying with the Jacobian matrix $\partial\overline{x}^\mu/\partial x^\nu$. Here, each new vector component is a linear combination of old components weighted by how much each new coordinate axis changes with respect to each old axis.

  - These objects, such as displacement vectors, inherently "follow" the coordinate axes. As the axes stretch or rotate, so do the components of the vector.

- Covariant vectors (gradients and one-forms): These vectors transform using the inverse of the Jacobian matrix $\partial x^\nu / \partial \overline{x}^\mu$. This reflects how changes in the old coordinates map to the new coordinates, suitable for objects that naturally pair with vectors, like gradients.
  - They measure rates of change along these coordinates.
  - A larger coordinate stretch means a smaller gradient in that direction to maintain the same rate of change.

## 6.4 The Kinetic Term

Note how we started with the kinetic term

$$\frac{1}{2}\left(\partial_{E_\mu}\phi\right)^2$$

and in 2D it became

$$\frac{1}{2}\left[\left(\frac{\partial\phi}{\partial t}\right)^2+\left(\frac{\partial\phi}{\partial x}\right)^2\right]$$

The first thing to note is that we are in Euclidean space where the four-vector $x_E^2 = x_0^2 + x_1^2 + x_2^2 + x_3^2$. If we were back in Mikowski space with out metric signature $(+---)$, then $x^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2$.

To see why the kinetic term looks the way it does, it helps to see it other ways.

For example, its definition is

$$\frac{1}{2}(\partial_\mu\phi)^2 = \frac{1}{2}\sum_{\mu=0}^{3}(\partial^\mu\phi)^2$$

and so in 4D it looks like

$$\frac{1}{2}(\partial_\mu\phi)^2 = \frac{1}{2}\left(\left(\frac{\partial\phi}{\partial t}\right)^2+\left(\frac{\partial\phi}{\partial x}\right)^2+\left(\frac{\partial\phi}{\partial y}\right)^2+\left(\frac{\partial\phi}{\partial z}\right)^2\right)$$

However, remember that this is all in Euclidean space. In Mikowski space we have,

$$\begin{aligned}\partial^2 &= \partial^\mu\partial_\mu \\ &= \frac{\partial^2}{\partial t^2}-\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}-\frac{\partial^2}{\partial z^2} \\ &= \frac{\partial^2}{\partial t^2}-\boldsymbol{\nabla}^2\end{aligned}$$

And this is why it sometimes helps to see the metric included in the definition,

$$\begin{aligned}\eta^{\mu\nu}\partial_\mu\partial_\nu\phi &= \partial_\mu\partial^\nu\phi \\ &= (\partial_t\phi)^2 - \delta^{ij}\partial_i\partial_j\phi\end{aligned}$$

If $\eta^{\mu\nu}$ has the signature $(+---)$.

Anyway, back to our problem of Discretization. If we define a lattice, with a spacing $a$, then

$$\lim_{a\to 0}\partial_x\phi = \frac{1}{a}\left(\phi(x+a,t) - \phi(x,t)\right)$$

Dropping the lim notation - trusting in our approximation for the time being - we can then write

$$(\partial_t\phi)^2 = \frac{1}{a^2}\left(\phi^2(x,t+a) + \phi^2(x,t) - 2\phi(x,t+a)\phi(x,t)\right)$$

and

$$(\partial_x\phi)^2 = \frac{1}{a^2}\left(\phi^2(x+a,t) + \phi^2(x,t) - 2\phi(x+a,t)\phi(x,t)\right)$$

If we put this back inside of an integral we have,

$$\int dx\,(\partial_x\phi)^2 \approx \int dx\frac{1}{a^2}\left(\phi^2(x+a,t) + \phi^2(x,t) - 2\phi(x+a,t)\phi(x,t)\right)$$

And so our 2D integral becomes,

$$\begin{aligned}\int dtdx\left[(\partial_t\phi)^2 + (\partial_x\phi)^2\right] \approx \frac{1}{a^2}\int dtdx\,\big[&\phi^2(x+a,t) + \phi^2(x,t) - 2\phi(x+a,t)\phi(x,t) \\ &+\phi^2(x,t+a) + \phi^2(x,t) - 2\phi(x,t+a)\phi(x,t)\big]\end{aligned}$$

The next step we can take in our road to discretization comes from remembering that your run of the mill integral is defined via Riemann sums, such that

$$\int dx f(x) = \lim_{|\Delta x|\to 0}\sum_{i=1}^{n} f(x_i)\Delta x_i$$

In our case, we are doing a double integral so $dt\,dx \to a^2$. Then we have,

$$\begin{aligned}\int dtdx\left[(\partial_t\phi)^2 + (\partial_x\phi)^2\right] \approx \sum_{x,t}\big[&\phi^2(x+a,t) + \phi^2(x,t) - 2\phi(x+a,t)\phi(x,t) \\ &+\phi^2(x,t+a) + \phi^2(x,t) - 2\phi(x,t+a)\phi(x,t)\big]\end{aligned}$$

And here lies yet another trick of the trade. If you remember, we actually want to integrate over these partial derivatives. And it just so happens that

when using periodic boundary conditions (which we always use), a sum over all studies of $\phi(x+a,t)$ is equal to a sum over all sites of $\phi(x,t)$, the only differnce is what the first and last terms are.

For example, imagine a lattice with just 3 sites, where the lattice spacing $a=1$. If we drop the $t$ from the list of arguments, these sums are:

$$\sum_{x=1}^{3}\phi(x+a)=\phi(2)+\phi(3)+\phi(1)$$

and

$$\sum_{x=1}^{3}\phi(x)=\phi(1)+\phi(2)+\phi(3)$$

Since we are working with a double (finite) sum, $\sum_{x,t}=\sum_t\sum_x$, then we can apply the above result twice, allowing us to simplify our integral.

$$\begin{aligned}\int dtdx\left[(\partial_t\phi)^2+(\partial_x\phi)^2\right]\approx\sum_{x,t}&\left[\phi^2(x,t)+\phi^2(x,t)-2\phi(x+a,t)\phi(x,t)\right.\\&\left.+\phi^2(x,t)+\phi^2(x,t)-2\phi(x,t+a)\phi(x,t)\right]\\=\sum_{x,t}&\left[4\phi^2(x,t)-2\phi(x+a,t)\phi(x,t)-2\phi(x,t+a)\phi(x,t)\right]\end{aligned}\tag{6.4.1}$$

Here the sum in $x$ is done for each $x_i=an_x$, where $n_x=0,\ldots,N-1$, and the sum over $t$ is similarly defined with $t_i=an_t$, where $n_t=0,\ldots,N-1$.

The next step now lies in thinking about us actually writing the code. If we think how we would implement a lattice as an array and just wrap things around, we could identify every point in our lattice with an index $n$ and "reindex" our coordinates. To map our double sum into a single dimension we use the mapping $n=x+t\times N$.

This way, by introducing some other indices $i$ and $j$ such tha $n+i=t\times N+x+1$ and $n+j=t\times N+x+N=(t+1)\times N+x$. we can simplify our notation by writing $\phi(x+a,t)=\phi(a(n_x+1),an_t)\to\phi(n+1)$ and $\phi(x,t+a)=\phi(an_x,a(n_t+1))\to\phi(n+j)$. Giving us

$$\begin{aligned}\int dtdx\left[(\partial_t\phi)^2+(\partial_x\phi)^2\right]&\approx\sum_n\left[4\phi^2(n)-2\phi(n+i)\phi(n)-2\phi(n+j)\phi(n)\right]\\&=\sum_n\left[4\phi^2(n)-2\phi(n)\left(\phi(n+i)+\phi(n+j)\right)\right]\end{aligned}$$

If you look at the terms $-2\phi(n)\left(\phi(n+i)+\phi(n+j)\right)$, these look like the nearest-neighbour interactions that are seen in the Ising model Hamiltonina, $H=-\sum_{<ij>}s_is_j$. The other two products, $-2\phi(n)\left(\phi(n-i)+\phi(n-j)\right)$, would again come from performing the sum with periodic boundary conditions. If we use this, we can now write,

$$\int dtdx \left[ (\partial_t \phi)^2 + (\partial_x \phi)^2 \right] \approx -2 \sum_{<ij>} \phi_i \phi_j + 4 \sum_n \phi_n^2 \tag{6.4.2}$$

### 6.4.1 The 4D Kinetic Term

To get the 4-dimensional expression of the kinetic term we need to go back to 6.4.1, which now becomes

$$\begin{aligned}
\int dt\, dx\, dy\, dz \left[ (\partial_t \phi)^2 + (\partial_x \phi)^2 + (\partial_y \phi)^2 + (\partial_z \phi)^2 \right] \approx \\
\sum \big[ & \phi^2 + \phi^2 - 2\phi(t+a,\ldots)\phi \\
& + \phi^2 + \phi^2 - 2\phi(x+a,\ldots)\phi \\
& + \phi^2 + \phi^2 - 2\phi(y+a,\ldots)\phi \\
& + \phi^2 + \phi^2 - 2\phi(z+a,\ldots)\phi \big] \\
& = \sum \left[ 8\phi^2 - 2\phi(t+a)\phi - 2\phi(x+a)\phi - 2\phi(y+a)\phi - 2\phi(z+a)\phi \right] \\
& = 8\sum \phi^2 - 2\sum \phi \left( \phi(t+a) + \phi(x+a) + \phi(y+a) + \phi(z+a) \right) \\
& = 8\sum \phi^2 - 2\sum \phi \left( \phi(n+i) + \phi(n+j) + \phi(n+k) + \phi(n+l) \right)
\end{aligned}$$

Which again, since the summation is over the entire lattice and we have periodic boundary conditions, we end up with

$$\int dt\, dx\, dy\, dz \left[ (\partial_t \phi)^2 + (\partial_x \phi)^2 + (\partial_y \phi)^2 + (\partial_z \phi)^2 \right] \approx -2 \sum_{<ij>} \phi_i \phi_j + 8 \sum_n \phi_n^2$$

Keep in mind that the product of nearest neighbours is now for each $ij$ over each dimension, so it is 8 multiplications in total.

## 6.5 The Potential Term

The next piece of wisdom we need to use is to make our phisical quantities dimensionless. This helps during the simulation to sustain numerical stability. Imagine if we were working with electrons and had to multiply everything by $9.1 \times 10^{-31}$ kg. The electron Mass in MeV is just 0.5 but hopefully you get the point that it could happen that we are working with a quantity that is just right for us to run into a numerical underflow.

Let's now look at our potential. Based on what we did for the kinetic term, we can see right away that

$$\int dt\, dx \left( \frac{1}{2}\mu^2\phi^2 + \frac{1}{4!}\lambda\phi^4 \right) \approx \sum a^2 \left( \frac{1}{2}\mu^2\phi^2 + \frac{1}{4!}\lambda\phi^4 \right)$$

### 6.5.1 Dimensional Analysis

Defining a unitless mass $\mu_0$ is straightforward: in natural units, mass has the units of inverse length. If we square a mass term, then we need to multiply it by a square length term to have a unitless quantity, and it just so happens that we get that term from the double integral. So we can define the unitless quantity $\mu^2 = a^2\mu^2$.

The logic we just followed was a lucky coincidence that happens in 2D. For example, one could naively think that the coupling constant $\lambda$ would have the units of energy, inverse length. But since we have an $a^2$ term, that would make our new coupling constant have units of length.

To get the correct answer, in two and four dimensions we have to go back a bit.

To start, the action $S$ is dimensionless - $S = \int dt\, L$, the lagrangian has units of energy, inverse lenght, so it cancels with $dt$.

In four dimensions, the action is $S = \int d^4x\, L$. The integration "volume" has units of $[M]^{-4}$, so all terms in the lagrangian need to have units of $[M]^4$. (In two dimensions all terms would need units of $[M]^2$.)

The first term in our laganragian in a spacetime derivative $\partial_\mu$. If you remember your calculus days, derivatives are rates of change, i.e., $\dot{y}$ is the velocity which has dimensions of meters per second. And thus $[\partial_m u] = [L^{-1}] = [M^1]$.

From here we can figure out the dimensions of the field $\phi$,

$$[(\partial_\mu \phi)^2] = [M]^2[\phi]^2 = [M]^4$$

Meaning that $[\phi]^2 = [M]^2$, or $[\phi] = [M]$.

This result is often generalized by noting that

$$[(\partial_\mu \phi)^2] = [M]^2[\phi]^2 = [M]^d$$

which makes $[\phi] = [M]^{(d-2)/2}$ in $d$ dimensions. ($[\phi] = 1$ (unitless) in two dimensions.)

Let's keep going we are on a roll! The next term in our lagrangian is,

$$\left[\frac{1}{2}\mu^2\phi^2\right] = [M]^4$$

So it happens that in four dimensions, $[\mu]^2 = [M]^2$, since $[\phi]^2 = [M]^2$. (In two dimensions, $[\mu] = [M]$ as well since $\phi$ is dimensionless.)

Finally, our interaction term, the one that made us start this route,

$$\left[\frac{1}{4!}\lambda\phi^4\right] = [\lambda][M]^4 = [M]^4$$

Therefore, $[\lambda] = 1$, so it is dimensionless in four dimensions. But in two dimensions we have,

$$\left[\frac{1}{4!}\lambda\phi^4\right] = [\lambda][M]^0 = [M]^2$$

Here, $[\lambda] = [M]^2 = [\mu^2]$.

### 6.5.2 Full Lagrangian in 2D

The results we need to highlight from the previous section are that $[\lambda] = [M]^2 = [\mu^2]$. So we can define the dimensionless quantities in 2D $\mu_L^2 = a^2\mu^2$ and $\lambda_L = a^2\lambda$, but we will drop the $L$ subscript since it gets old.

That way the potential term for our Lagrangian becomes

$$\int dt\, dx \left(\frac{1}{2}\mu^2\phi^2 + \frac{1}{4!}\lambda\phi^4\right) \approx \sum \left(\frac{1}{2}\mu^2\phi^2 + \frac{1}{4!}\lambda\phi^4\right)$$

Putting it together with the kinetic term 6.4.2 (and the factor of 1/2 we didn't carry around) we have

$$\begin{aligned}
&\int dt\, dx \left(\frac{1}{2}\left[(\partial_t\phi)^2 + (\partial_x\phi)^2\right] + \frac{1}{2}\mu^2\phi^2 + \frac{1}{4!}\lambda\phi^4\right) \\
&\qquad \approx -\sum_{<ij>} \phi_i\phi_j + 2\sum_n \phi_n^2 + \sum_n \left(\frac{1}{2}\mu^2\phi_n^2 + \frac{1}{4!}\lambda\phi_n^4\right) \qquad (6.5.1) \\
&\qquad = -\sum_{<ij>} \phi_i\phi_j + \sum_n \left(2 + \frac{\mu^2}{2}\right)\phi_n^2 + \frac{\lambda}{4!}\phi_n^4
\end{aligned}$$

# Bibliography


[Mic95] Dan V. Schroeder Michael E. Peskin. *An Introduction To Quantum Field Theory*. Frontiers in Physics, 1995. ISBN: 978-0201503975.

[Kay11] David C. Kay. "Schaums Outline of Tensor Calculus". In: McGraw Hill, 2011. Chap. 1. ISBN: 978-0071756037.

[Gri13] David J. Griffiths. *Introduction to Electrodynamics 4th Edition*. Pearson, 2013. ISBN: 978-0-321-85656-2.

[Jam14] David Skinner James Binney. *The Physics of Quantum mechanics*. Oxford University Press, 2014. ISBN: 978-0-19-968856-2.

[Tom14] Stephen J. Blundell Tom Lancaster. "Quantum Field Theory for the Gifted Amateur". In: Oxford University Press, 2014. Chap. 0.4. ISBN: 978-0199699339.

[GBC16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. `http://www.deeplearningbook.org`. MIT Press, 2016.

[Tri18] Mario F. Triola. *Elementary Statistics 13th edition*. Pearson, 2018. ISBN: 978-0-13-446245-5.

[Gra23] Adam Grant. *Hidden Potential: The Science of Achieving Greater Things*. Viking Press, 2023. ISBN: 978-0593653142.

[Ala] Jorge Alarcon. *github.com/seafoodfry/numerical-analysis/qft/phi-theory*. URL: `https://github.com/seafoodfry/numerical-analysis/blob/main/qft/phi-theory/`. (accessed: 04.17.2024).

[Daw] Paul Dawkins. *Section 15.8 : Change of Variables*. URL: `https://tutorial.math.lamar.edu/classes/calciii/changeofvariables.aspx`. (accessed: 04.17.2024).

[Sch] David Schaich. *Lattice Simulations of Nonperturbative Quantum Field Theories*. URL: `https://inspirehep.net/literature/1386200`. (accessed: 04.17.2024).

[Spi] Adam Spiegler. *Statistical Theory, Section 4.3: Properties of Estimators*. URL: `https://aspiegler.github.io/Statistical-Theory/15-Properties-Estimators.html#theorem-15.1`. (accessed: 10.20.2024).