PHYS 215A: Lectures on Quantum Field Theory

Benjamín Grinstein
UCSD

November 10, 2014

Chapter 1

Introduction

1.1 Why QFT?

As a new student of the subject you may ask Why study Quantum Field Theory (QFT)? Or why do we *need* QFT? There are many reasons, and I will explain some.

Pair Creation. For starters, particle quantum mechanics (QM) does not account for pair creation. In collisions of electrons with sufficiently large energies it is occasionally observed that out of the collision an electron and a positron are created (in addition to the originating electrons): $e^-e^- \to e^-e^-e^+e^-$. Sufficiently large energies means kinetic energies larger than the rest mass of the e^-e^+ pair, namely $2m_ec^2$. Not only does single particle QM not account for this but in this case the colliding electrons must be relativistic. So we need a relativistic QM that accounts for particle creation.

But what if we are atomic physicists, and only care about much smaller energies? If we care about high precision, and atomic physicists do, then the fact that pair creation is possible in principle cannot be ignored. Consider the perturbation theory calculation of corrections to the n-th energy level, E_n, of some atomic state in QM:

$$\delta E_n = \langle n|H'|n\rangle + \sum_{k\neq n} \frac{\langle n|H'|k\rangle\langle k|H'|n\rangle}{E_n - E_k} + \cdots$$

The second order term involves *all* states, regardless of their energy. Hence states involving e^+e^- states give corrections of order

$$\frac{\delta E}{E} \sim \frac{\text{atomic energy spacing}}{m_ec^2}.$$

This is, *a priori*, as important as the relativistic corrections from kinematics, $H = \sqrt{(m_ec^2)^2 + (pc)^2} = m_ec^2 + \frac{1}{2}p^2/m_e - \frac{1}{8}p^4/m_e^3c^2 + \cdots$. The first relativis-

tic correction is $H' = -\frac{1}{4}(p/m_e c)^2(p^2/2m_e)$ so the fractional correction $\delta E/E \sim (p/m_e c)^2 \sim (p^2/2m_e)/(m_e c^2)$, just as in pair creation. We are lucky that for the Hydrogen atom the pair creation correction happens to be small. But in general we have no right to neglect it.

Instability of relativistic QM. Let's insist in single particle QM and explore some consequences. Since $H = \sqrt{(m_e c^2)^2 + (pc)^2}$ does not give us a proper differential Schrödinger equation, let's try a wave equation using the square of the energy, $E^2 = (mc^2)^2 + (pc)^2$, and replacing, as usual, $E \to i\hbar\partial/\partial t$ and $\vec{p} \to -i\hbar\nabla$. This leads to the Klein-Gordon wave equation,

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \frac{\mu^2}{c^2}\right]\phi(\vec{x},t) = 0$$

where $\mu = mc^2/\hbar$. Since this is a free particle we look for plane wave solutions,

$$f^+(\vec{x},t) = a_k^+ e^{-i\omega_k t + i\vec{k}\cdot\vec{x}} \qquad \text{and} \qquad f^-(\vec{x},t) = a_k^- e^{i\omega_k t - i\vec{k}\cdot\vec{x}},$$

where $a^\pm$ are constants (independent of $\vec{x}$ and t) and

$$\omega_k = \sqrt{c^2\vec{k}^2 + \mu^2}.$$

The interpretation of f^+ is clear: using $E \to i\hbar\partial/\partial t$ and $\vec{p} \to -i\hbar\nabla$ we see that it has energy $E = \hbar\omega_k$ and momentum $\vec{p} = \hbar k$ and these are related by $E = \sqrt{(m_e c^2)^2 + (pc)^2}$. However, we now also have solutions, f^- with energy $E = -\hbar\omega_k$ (and momentum $\vec{p} = -\hbar k$). These have negative energy. For the free particle this is not a problem since we can start with a particle of some energy, positive or negative, and it will stay at that energy. But as soon as we introduce interactions, say by making the particle charged and giving it minimal coupling to the electromagnetic field, the particle can radiate into a lower energy state, but there is no minimum energy. There is a catastrophic instability.

Dirac proposed an ingenious solution to this catastrophe. Suppose the particle is a fermion, say, an electron. Then start the system from a state in which all the negative energy states are occupied. We call this the "Dirac sea." Since no two fermions can occupy a state with the same quantum numbers, we cannot have an electron with positive energy radiate to become a negative energy state (that state is occupied). However, an energetic photon can interact with an electron with $E < -m_e c^2$ and bump it to a state with $E > m_e c^2$. What results is a state with a positive energy electron and a hole in the "Dirac sea." The hole signifies the absence of negative energy and negative charge, so it behaves as a particle with positive energy and positive charge. That is, as the anti-particle of the electron,

the positron. So we conclude we cannot have a single particle QM, we are forced to consider pair creation.

A note in passing: as we will discuss later, relativistic fermions are described by the Dirac equation. But a solution of the Dirac equation necessarily solves the Klein-Gordon equation, so the discussion above is in fact appropriate to the fermionic case.

Klein's Paradox. There is another way to see the need to include anti-particles in the solution to the Klein-Gordon equation. As we stated, only if we include interactions will we see a problem. But our description of interactions with an electromagnetic field was heuristic. We can make it more precise by introducing instead a potential. The simplest case, considered by Klein, is that of a particle in one spatial dimension with a step potential, $V(x) = V_0\,\theta(x)$. Here $\theta(x)$ is the Heavyside step-function

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

and $V_0 > 0$ is a constant with units of energy. We look for solutions with a plane wave incident from the left ($x < 0$ with $p = \hbar k > 0$):

$$\psi_L(x,t) = e^{-i\omega t + ik_L x} + Re^{-i\omega t - ik_L x}$$
$$\psi_R(x,t) = Te^{-i\omega t + ik_R x}$$

These solve the Klein-Gordon equation with the shifted energy provided

$$ck_L = \sqrt{\omega^2 - \mu^2} \qquad \text{and} \qquad ck_R = \sqrt{(\omega - \omega_0)^2 - \mu^2}$$

where $\omega_0 = V_0/\hbar$. We now determine the transmission (T) and reflection (T) coefficients matching the solutions at $x = 0$; using $\psi_L(t,0) = \psi_R(t,0)$ and $\frac{\partial \psi_L(t,0)}{\partial t} = \frac{\partial \psi_R(t,0)}{\partial t}$ we get

$$T = \frac{2k_L}{k_L + k_R} \qquad \text{and} \qquad R = \frac{k_L - k_R}{k_L + k_R}.$$

Formally the solution looks just like in the non-relativistic (NR) case. Indeed, for $\omega > \mu + \omega_0$ both wave-vectors k_L and k_R are real, and so are both transmission coefficients and we have a transmitted and a reflected wave. Similarly if $\omega_0 - \mu < \omega < \omega_0 + \mu$ then k_R has a nonzero imaginary part and we have total reflection (the would be transmitted wave is exponentially damped).

But for $V_0 > 2mc^2$, which corresponds to an energy large enough that pair creation is energetically allowed, there is an unusual solution. If $\mu < \omega < \omega_0 - \mu$, which means $\omega - \mu > 0$ and $\omega - \mu < \omega_0 - 2\mu$, we obtain both k_L and k_R are real, and so are both T and R. There is a non-zero probability that the particle, which

has less energy than the height of the barrier ($E - mc^2 - V_0 < 0$), is transmitted. This weird situation, called "Klein's paradox," can be understood in terms of pair creation. A complete treatment of the problem requires (as far as I know) a fully quantum field theoretic treatment, so for that we will have to wait until later in this course. But the result can be easily described: the incident particle is fully reflected, but is accompanied by particle-antiparticle pairs.

Bohr's Box. Above we talked about "localizing" a particle. Let's make this a bit more precise (not much, though). Suppose that in order to localize a free particle we put it in a large box. We don;t know where the particle is other than that it is inside the box. If choose the box large enough, of sides $L_{x,y,z} \gg \lambda_C$, then the uncertainly in its momentum can be small, $\Delta p_{x,y,z} \gtrsim \hbar/L_{x,y,z}$. Then the kinetic energy can be small, $E = \frac{\vec{p}^2}{2m} \sim \frac{\hbar^2}{2m}(L_x^{-2} + L_y^{-2} + L_z^{-2})$. Now suppose one side of the box, parallel to the yz plane, is movable (the box is a cylinder, the movable box is a piston). So we can attempt to localize the particle along the x axis by compressing it, that is, by decreasing L_x. Once $L_x \sim \lambda_C$ the uncertainty in the energy of the particle is $\Delta E \sim mc^2$. Now o localize the particle we have introduced interaction, those of the particle with the walls of the box that keeps it contained. But as we have seen when the energy available exceeds $2mc^2$ interactions require a non-vanishing probability for pair creation. We conclude that as we try to localize the particle to within a Compton wavelength, $\lambda_C = \hbar/mc$, we get instead a state which is a combination of the particle plus a particle-antiparticle pair. Or perhaps two pairs, or three pairs, or .... In trying to localize a particle not only we have lost certainty on its energy, as is always the case in QM, but we have also lost certainty on the number of particles we are trying to localize!

This gives us some physical insight into Klein's paradox. The potential step localizes the particle over distances smaller than a Compton wavelength. The uncertainty principle then requires that the energy be uncertain by more than the energy required for pair production. It can be shown that if the step potential is replaced by a smooth potential that varies slowly between 0 and V_0 over distances d larger than λ_C then transmission is exponentially damped, but as d is made smaller than λ_C Klein's paradox re-emerges (Sauter).

Democracy. It is well known that elementary particles are indistinguishable: every electron in the world has the same mass, charge and magnetic moment. In NRQM we account for this in an ad-hoc fashion. We write

$$H = \sum_i \frac{\vec{p}_i^2}{2m_e} + \frac{e}{m_e c}\vec{A}(\vec{x}_i)\cdot\vec{p}_i + \vec{B}(\vec{x}_i)\cdot\vec{\mu}_i \qquad \text{where } \vec{\mu}_i = \frac{ge\hbar}{2m_e c}\vec{\sigma}_i,$$

but we have put in by hand that the mass m_e, the charge e and the gyromagnetic ratio g are the same for all electrons. Where does this democratic choice come from?

Moreover, photons, which are quanta of the electromagnetic field, are introduced by second quantizing the field. But electrons are treated differently. Undemocratic! If instead we insist that all particles are quanta of field excitations not only we will have a more democratic (aesthetically pleasing?) setup but we will have an explanation for indistinguishability, since all corpuscular excitations of a field carry the same quantum numbers automatically.

Indistinguishability is fundamentally importance in Pauli's exclusion principle. If several electrons had slightly different masses or slightly different charges from each other, then they could be distinguished and they could occupy the same atomic orbital (which in fact would not be the same, but slightly different). More generally, indistinguishability is at the heart of "statistics" in QM. But the choice of Bose-Einstein vs Fermi-Dirac statistics is a recipe in particle QM, it has to be put in by hand. Since QFT will give us indistinguishability automatically, one may wonder if it also has something to say about Bose-Einstein vs Fermi -Dirac statistics. In fact it does. We will see that consistent quantization of a field describing spin-0 corpuscles requires them to obey Bose-Einstein statistics, while quantization of fields that give spin-1/2 particles results in Fermi-Dirac statistics.

Causality. In relativistic kinematics faster than light signal propagation leads to paradoxes. The paradoxes come about because faster than light travel violates our normal notion of causality. A spaceship (it's always a spaceship) moving faster than light from event A to event B is observed in other frames as moving from B to A. You never had to worry about this in NRQM because you never had to worry about it in NR classical mechanics. But you do worry about it in relativistic mechanics so you must be concerned that related problems arise in relativistic QM.

In particle QM we can define an operator $\hat{\vec{x}}$ and we can use eigenstates of this operator, $\hat{\vec{x}}|\vec{x}\rangle = \vec{x}|\vec{x}\rangle$ to describe the particle. For now I will use a hat to denote operators on the Hilbert space, but I will soon drop this. The operator is conjugate to the momentum operator, $\hat{\vec{p}}$, in the sense that $i[\hat{x}_j, \hat{p}_k] = \delta_{jk}$, and this gives rise to a relation between their eigenstates,

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

A state $|\psi\rangle$ at $t = 0$ evolves into $e^{-i\hat{H}t/\hbar}|\psi\rangle$ a time t later. We can ask what does the state of a particle localized at the origin at $t = 0$ evolve into a time t later. Now, in NRQM we know the answer: the particle spreads out. Whether it spreads out faster than light or not is not an issue so we don't ask the question. But in relativistic QM

it matters, so we address this. Note however that for consistency we must use a relativistic Hamiltonian. For a free particle we can use $\hat{H} = \sqrt{(\hat{\vec{p}}c)^2 + (mc^2)^2}$ since its action on states is unambiguously given by its action on momentum eigenstates, $\hat{H}|\vec{p}\rangle = \sqrt{(\vec{p}c)^2 + (mc^2)^2}|\vec{p}\rangle$. The probability amplitude of finding the particle at $\vec{x}$ at time t (given that it started as a localized state at the origin at time $t = 0$) is

$$\begin{aligned}
\langle\vec{x}|e^{-i\hat{H}t/\hbar}|\vec{x}=0\rangle &= \int d^3p\, \langle\vec{x}|\vec{p}\rangle\langle\vec{p}|e^{-i\hat{H}t/\hbar}|\vec{x}=0\rangle \qquad \text{(complete set of states)} \\
&= \int d^3p\, \frac{1}{(2\pi\hbar)^3} e^{i\vec{p}\cdot\vec{x}/\hbar} e^{-iE_pt/\hbar} \qquad \text{(where } E_p = \sqrt{(pc)^2 + (mc^2)^2}\text{)} \\
&= \int_0^\infty \frac{p^2dp}{(2\pi\hbar)^3} \int_{-1}^{1} d\cos\theta \int_0^{2\pi} d\phi\, e^{ipr\cos\theta/\hbar} e^{-iE_pt/\hbar} \qquad (p = |\vec{p}|;\ r = |\vec{x}|) \\
&= -\frac{i}{(2\pi\hbar)^2 r} \int_{-\infty}^{\infty} p\, dp\, e^{irp/\hbar} e^{-iE_pt/\hbar} \qquad (1.1)
\end{aligned}$$

This is not an easy integral to compute. But we can easily show that it does not vanish for $r > ct > 0$. This means that there is a non-vanishing probability of finding the particle at places that require it propagated faster than light to get there in the allotted time. This is a violation of causality.

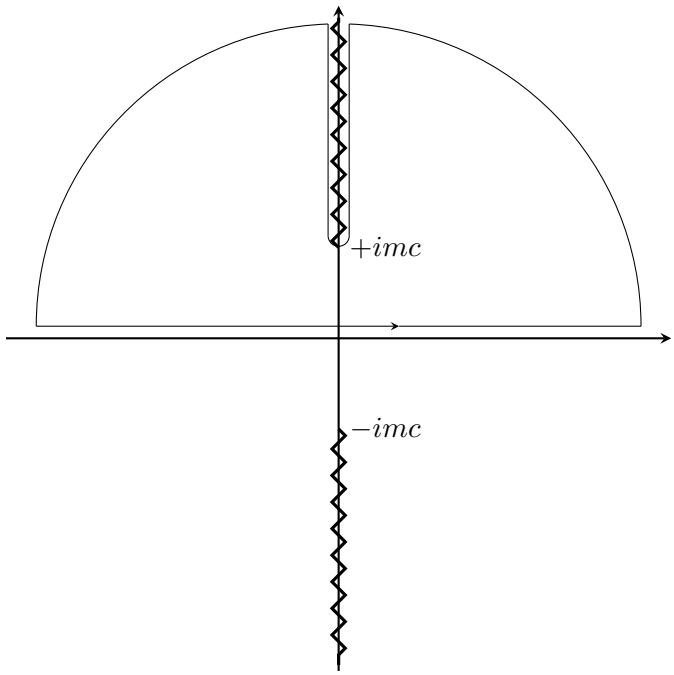

Figure 1.1: Contour integral for evaluating (1.1).

Before discussing this any further let's establish the claim that the integral does not vanish for $r/t > c$. The integral is difficult to evaluate because it is oscillatory. However, we can use complex analysis to relate the integral to one performed over purely imaginary momentum, turning the oscillating factor into an exponential convergence factor. So consider the analytic structure of the integrand. Only the square root defining the energy is non-analytic, with a couple of branch points at $p = \pm imc$. Choose the branch cut to extend from $+imc$ to $+i\infty$ and from $-imc$ to $-i\infty$ along the imaginary axis; see Fig. 1.1. The integral over the contour C vanishes (there are no poles of the integrand), and for $r > ct$ the integral over the semicircle of radius R vanishes exponentially fast as $R \to \infty$. Then the integral we want is related to the one on both sides of the positive imaginary axis branch cut,

$$\langle \vec{x}|e^{-i\hat{H}t/\hbar}|\vec{x}=0\rangle = \frac{i}{(2\pi\hbar)^2 r}\int_{mc}^{\infty} p\,dp\,e^{-rp/\hbar}\left(e^{\sqrt{p^2-(mc)^2}ct/\hbar} - e^{-\sqrt{p^2-(mc)^2}ct/\hbar}\right). \tag{1.2}$$

The integrand is everywhere positive. It decreases exponentially fast as r increases for fixed t, so the violation to causality is small. But any violation to causality is problematic.

1.2 Units and Conventions

You surely noticed the proliferation of c and $\hbar$ in the equations above. The play no role, other than to keep units consistent throughout. So for the remainder of the course we will adopt units in which $c = 1$ and $\hbar = 1$. You are probably familiar with $c = 1$ already: you can measure distance in light-seconds and then x/t has no units. But now, in addition energy momentum and mass and measured in the same units (after all $E^2 = (pc)^2 + (mc^2)^2$). We denote units by square brackets, the units of X are $[X]$; we have $[E] = [p] = [m]$, and $[x] = [t]$.

The choice $\hbar = 1$ may be less familiar. Form the uncertainty condition, $\Delta p \Delta x \geq \hbar$, so $[p] = [x^{-1}]$. This together with the above (from $c = 1$) means that everything can be measured in units of energy, $[E] = [p] = [m] = [x^{-1}] = [t^{-1}]$. In particle physics it is customary to use GeV as the common unit. That's because many elementary particles have masses of the order or a GeV:

particle	symbol	mass(GeV)
proton	m_p	0.938
neutron	m_n	0.940
electron	m_e	5.11×10^{-4}
W-boson	M_W	80.4
Z-boson	M_Z	91.2
higgs boson	M_h	126

To convert units, use $\hbar = 6.582 \times 10^{-25}$ GeV·sec, and often conveniently $\hbar c = 0.1973$ GeV·fm, where fm is a Fermi, or femtometer, 1 fm $= 10^{-15}$ m, a typical distance scale in nuclear physics.

In these units the fine structure constant, $\alpha = e^2/4\pi\hbar c$ is a pure number,

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137}.$$

Since we will study relativistic systems it is useful toset up conventions for our notation. We use the "mostly-minus" metric, $\eta_{\mu\nu} = \text{diag}(+,-,-,-)$. That is, the invariant interval is $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$. The Einstein convention, an implicit sum over repeated indices unless otherwise stated, is adopted. Four-vectors have upper indices, $a = (a^0, a^1, a^2, a^3)$, and the dot product is

$$a \cdot b = \eta_{\mu\nu} a^\mu b^\nu = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 = a^0 b^0 - \vec{a} \cdot \vec{b} = a^0 b^0 - a^i b^i$$

We use latin indices for the spatial component of the 4-vectors, and use again the Einstein convention for repeated latin indices, $\vec{a} \cdot \vec{b} = a^i b^i$. Indices that run from 0 to 3 are denoted by greek letters. We also use $a^2 = a \cdot a$ and $|\vec{a}|^2 = \vec{a}^2 = \vec{a} \cdot \vec{a}$. Sometimes we even use $a^2 = \vec{a} \cdot \vec{a}$, even when there is a 4-vector a^μ; this is confusing, and should be avoided, but when it is used it is always clear from the context whether a^2 is the square of the 4-vector or the 3-vector.

The inverse metric is denoted by $\eta^{\mu\nu}$,

$$\eta^{\mu\nu} \eta_{\nu\lambda} = \delta^\mu_\lambda$$

where δ^μ_ν is a Kronecker-delta, equal to unity when the indices are equal, otherwise zero. Numerically in Cartesian coordinates $\eta^{\mu\nu}$ is the same matrix as the metric $\eta_{\mu\nu}$, $\eta^{\mu\nu} = \text{diag}(+,-,-,-)$, but it is convenient to differentiate among them because they need not be the same in other coordinate systems. We can use the metric and inverse metric to define lower index vectors (I will not use the names "covariant" and "contravariant") and to convert among them:

$$a_\mu = \eta_{\mu\nu} a^\nu, \qquad a^\mu = \eta^{\mu\nu} a_\nu.$$

Then the dot product can be expressed as

$$a \cdot b = \eta_{\mu\nu} a^\mu b^\nu = a^\mu b_\mu = a_\mu b^\mu.$$

Generalized Einstein convention: any type of repeated index is understood as summed, unless explicitly stated. For example, if we have a set of quantities ϕ^a with $a = 1, \ldots, N$, then $\phi^a \phi^a$ stands for $\sum_{a=1}^N \phi^a \phi^a$.

1.3 Lorentz Transformations

Lorentz transformations map vectors into vectors

$$a^\mu \to \Lambda^\mu{}_\nu a^\nu$$

preserving the dot product,

$$a \cdot b \to (\Lambda a) \cdot (\Lambda b) = a \cdot b$$

Since this must hold for any vectors a and b, we must have

$$\Lambda^\lambda{}_\mu \Lambda^\sigma{}_\nu \eta_{\lambda\sigma} = \eta_{\mu\nu} \tag{1.3}$$

Multiplying by the inverse metric, $\eta^{\nu\rho}$

$$\Lambda^\lambda{}_\mu \Lambda^\sigma{}_\nu \eta_{\lambda\sigma} \eta^{\nu\rho} = \delta^\rho_\mu$$

we see that

$$\Lambda_\lambda{}^\rho \equiv \Lambda^\sigma{}_\nu \eta_{\lambda\sigma} \eta^{\nu\rho}$$

is the inverse of $\Lambda^\lambda{}_\rho$, $(\Lambda^{-1})^\rho{}_\lambda = \Lambda_\lambda{}^\rho$. Eq. (1.3) can be written in matrix notation as

$$\Lambda^T \eta \Lambda = \eta \tag{1.4}$$

where the superscript "T" stands for "transpose,"

$$(\Lambda^T)_\mu{}^\lambda = \Lambda^\lambda{}_\mu.$$

Lorentz transformations form a group with multiplication given by composition of transformations (which is just matrix multiplication, $\Lambda_1\Lambda_2$): there is an identity transformation (the unit matrix), an inverse to every transformation (introduced above) and the product of any two transformations is again a transformation. The *Lorentz group* is denoted $O(1,3)$. Taking the determinant of (1.4), and using $\det(AB) = \det(A)\det(B)$ and $\det(A^T) = \det(A)$ we have $\det^2(\Lambda) = 1$, and since Λ is real, $\det(\Lambda) = +1$ or -1. The product of two Lorentz transformations with $\det(\Lambda) = +1$ is again a Lorentz transformation with $\det(\Lambda) = +1$, so the set of transformations with $\det(\Lambda) = +1$ form a subgroup, the group of *Special (or Proper) Lorentz Transformations*, $SO(1,3)$. Among the $\det(\Lambda) = -1$ transformations is the parity transformation, that is, reflection aboutt he origin, $\Lambda = \mathrm{diag}(+1,-1,-1,-1)$.

Taking $\mu = \nu = 0$ in (1.3) we have

$$(\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\Lambda^i{}_0)^2 > 1$$

So any Lorentz transformation has either ${\Lambda^0}_0 \geq 1$ or ${\Lambda^0}_0 \leq 1$. The set of transformations with ${\Lambda^0}_0 \geq 1$ are continuously connected, and so are the ones with ${\Lambda^0}_0 \leq 1$, but no continuous transformation can take one type to the other. Among those with ${\Lambda^0}_0 \geq 1$ is the identity transformation; among those with ${\Lambda^0}_0 \leq 1$ is time reversal, $\Lambda = \text{diag}(-1, +1, +1, +1)$. We will have much more to say about parity and time reversal later in this course. Transformations with ${\Lambda^0}_0 \geq 1$ are called *orthochronous* and they form a subgroup denoted by $O^+(1,3)$, The subgroup of proper, orthochronous transformations, sometimes called the *restricted Lorentz group*, $SO^+(1,3)$.

Examples of Lorentz transformations: boosts along the x-axis

$$\Lambda = \begin{pmatrix} \cosh\theta & \sinh\theta & 0 & 0 \\ \sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.5}$$

and rotations

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$$

where R is a 3×3 matrix satisfying $R^T R = 1$. Rotations form a group, the group of 3×3 orthogonal matrices, $O(3)$. Note that $\det^2(R) = 1$, so the matrices with $\det(R) = +1$ form a subgroup, the group of Special Orthogonal transformations, $SO(3)$.

1.3.1 More conventions

We will use a common shorthand notation for derivatives:

$$\partial_\mu \phi = \frac{\partial \phi}{\partial x^\mu}$$

Note that the lower index on ∂_μ goes with the upper index in the "denominator" in $\frac{\partial \phi}{\partial x^\mu}$. We can see that this works correctly by taking

$$\frac{\partial}{\partial x^\mu}(x^\nu a_\nu) = a_\mu,$$

an lower index vector. Of course, the justification is how the derivative transforms under a Lorentz transformation. If $(x')^\mu = {\Lambda^\mu}_\nu x^\nu$ then

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial}{\partial (x)^\lambda} = \frac{\partial}{\partial x'^\mu}\left({(\Lambda^{-1})^\lambda}_\nu (x')^\nu\right)\frac{\partial}{\partial (x)^\lambda} = {(\Lambda^{-1})^\lambda}_\mu \frac{\partial}{\partial (x)^\lambda}$$

as it should: $\partial'_\mu = (\Lambda^{-1})^\lambda{}_\mu \partial_\lambda$. Sometimes we use also $\partial^\mu = \eta^{\mu\nu}\partial_\nu$. For integrals we use standard notation,

$$\int dt\, dx\, dy\, dz = \int dx^0\, dx^1\, dx^2\, dx^3 = \int d^4x = \int dt \int d^3x.$$

This is a Lorentz invariant (because the Jacobian, $|\det(\Lambda)| = 1$):

$$\int d^4x' = \int d^4x.$$

1.4 Relativistic Invariance

What does it mean to have a relativistic formulation of QM? A QM system is completely defined by its states (the Hilbert space $\mathcal{H}$) and the action of the Hamiltonian $\hat{H}$ on them. Consider again a free particle. Since $H = E$ is in a 4-vector with $\vec{p}$ we define a QM system by

$$\hat{\vec{p}}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle, \qquad \hat{H}|\vec{p}\rangle = \sqrt{\vec{p}^2 + m^2}|\vec{p}\rangle.$$

What do we mean by this being relativistic? That is, how is this invariant under Lorentz transformations?

To answer this it is convenient to first review rotational invariance, which we are more familiar with. In QM for each rotation R there is an operator on $\mathcal{H}$, $\hat{U}(R)$ such that

$$\hat{U}(R)|\vec{p}\rangle = |R\vec{p}\rangle$$

I will stop putting "hats" on operators when it is clear we are speaking of operators. So from here on $U(R)$ stands for $\hat{U}(R)$, etc. Note that

$$\begin{aligned} U(R)\hat{\vec{p}}\,|\vec{p}\,\rangle = \vec{p}\,U(R)|\vec{p}\rangle = \vec{p}\,|R\vec{p}\rangle \\ \Rightarrow U(R)\hat{\vec{p}}\,U(R)^{-1}|R\vec{p}\,\rangle = R^{-1}(R\vec{p}\,|R\vec{p}\,\rangle) = R^{-1}\hat{\vec{p}}\,|R\vec{p}\,\rangle \\ \Rightarrow U(R)\hat{\vec{p}}\,U(R)^{-1} = R^{-1}\hat{\vec{p}} \end{aligned} \quad (1.6)$$

We would like U to be unitary, so that probability of finding a state is the same as that of finding the rotated state (that, and $[U, H] = 0$, is what we mean by a symmetry). We should be able to prove that $U^\dagger U = UU^\dagger = 1$. Let's assume the states $|R\vec{p}\,\rangle$ are normalized by

$$\langle\vec{p}\,'|\vec{p}\rangle = \delta^{(3)}(\vec{p}\,' - \vec{p}).$$

Then to show $UU^\dagger = 1$ we use a neat trick:

$$U(R)U^\dagger(R) = U(R)\int d^3p\,|\vec{p}\rangle\langle\vec{p}|U(R)^\dagger = \int d^3p\,|R\vec{p}\rangle\langle R\vec{p}| = \int d^3p'\,|\vec{p}\,'\rangle\langle\vec{p}\,'| = 1,$$

where we have used $\vec{p}' = R\vec{p}$ and $d^3p = d^3p'$. The latter is non-trivial. It reflects the (clever) choice of normalization of states, which leads to a rotational invariant measure.

You can see things go wrong if one defines the normalization of states differently. Say $|\vec{p}\,\rangle_B = (1+|\vec{a}\cdot\vec{p}\,|)|\vec{p}\,\rangle$ where $\vec{a}$ is a fixed vector (and the subscript "B" stands for "Bad"). Then

$$ {}_B\langle\vec{p}\,'|\vec{p}\,\rangle_B = (1+|\vec{a}\cdot\vec{p}\,|)^2\delta^{(3)}(\vec{p}\,'-\vec{p}) \Rightarrow 1 = \int d^3p \frac{|R\vec{p}\,\rangle_{BB}\langle R\vec{p}\,|}{(1+|\vec{a}\cdot\vec{p}\,|)^2}. $$

The point is that there is a choice of states for which we can show easily $UU^\dagger = 1$. It is true that $UU^\dagger = 1$ even for the bad choice of states, it is just more difficult to prove.

We still have to show that U commutes with H, but this is simple:

$$ U(R)H|\vec{p}\,\rangle = \sqrt{\vec{p}^2+m^2}|\vec{p}\,\rangle \Rightarrow U(R)HU(R)^{-1}|R\vec{p}\,\rangle = \sqrt{(R\vec{p})^2+m^2}|R\vec{p}\,\rangle $$
$$ \Rightarrow U(R)HU(R)^{-1} = H. $$

where we used $(R\vec{p})^2 = \vec{p}^2$ in the second step.

Now consider Lorentz Invariance: $p \to \Lambda p$ with $\Lambda \in O(1,3)$. Actually, for now we will restrict attention to transformations $\Lambda \in SO(1,3)$ with $\Lambda^0{}_0 \geq 1$. That's because time reversal and spatial inversion (parity) present their own subtleties about which we will have much to say later in the course. As before we have $U(\Lambda)|E,\vec{p}\,\rangle = |\Lambda(E,\vec{p}\,)\rangle$. Note, first, that there is no need to specify E in addition to $\vec{p}$; we are being explicit to understand how Λ acts on states. In fact, since $E^2 - \vec{p}^2 = m^2$, E and $\vec{p}$ fall on a hyperboloid. The action of Λ on $(E,\vec{p})$ just moves points around in the hyperboloid. In order to show $UU^\dagger = 1$ we try the same trick:

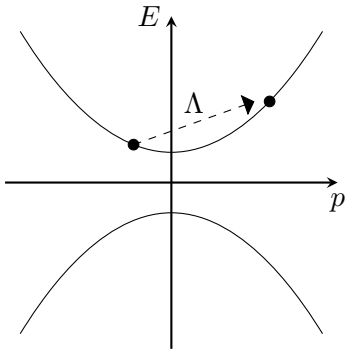

Figure 1.2: The hyperboloid $E^2 - \vec{p}^2 = m^2$; upper(lower) branch has $E > 0 (< 0)$.

$$ U(\Lambda)U^\dagger(\Lambda) \;=\; U(\Lambda)\int d^3p\,|E,\vec{p}\rangle\langle E,\vec{p}\,|\;\; U(\Lambda)^\dagger \;=\; \int d^3p\,|\Lambda(E,\vec{p})\rangle\langle\Lambda(E,\vec{p})|. $$

But now we hit a snag: if $(E',\vec{p}\,') = \Lambda(E,\vec{p})$, then $d^3p' \neq d^3p$. This is easily seen by considering boosts along the x direction, Eq. (1.5). But our experience from the discussion above suggests we find a better basis of states, one chosen

so that the measure of integration is Lorentz invariant. In fact, we can engineer back the normalization of states from requiring a invariant measure. Start from the observation that the 4-dimensional measure is invariant, $d^4p' = d^4p$. Now pick from this the upper hyperboloid in Fig. 1.2 in a manner that explicitly preserves Lorentz invariance. This can be done using a delta function, $\delta(p^2-m^2) = \delta((p^0)^2-\vec{p}^2-m^2)$ and a step function $\theta(p^0)$ to select the upper solution of the δ-function constraint. Note that $\theta(p^0)$ is invariant under transformations with ${\Lambda^0}_0 \geq 1$. So we take our measure to be

$$\begin{aligned} d^4p\ \delta(p^2-m^2)\theta(p^0) &= d^4p\ \delta((p^0)^2 - E_{\vec{p}}^2)\theta(p^0) \quad (\text{where } E_{\vec{p}} \equiv \sqrt{\vec{p}^2+m^2}) \\ &= d^3p\ dp^0\ \frac{1}{2E_{\vec{p}}}\delta(p^0 - E_{\vec{p}}) \\ &= \frac{d^3p}{2E_{\vec{p}}} \end{aligned}$$

For later convenience we introduce a constant factor and compact notation,

$$(dp) = \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}}\,.$$

The corresponding (relativistic) normalization of states is

$$\langle\vec{p}'|\vec{p}\rangle = (2\pi)^3 2E_{\vec{p}}\ \delta^{(3)}(\vec{p}' - \vec{p}) \tag{1.7}$$

Now we have,

$$\begin{aligned} U(\Lambda)U^\dagger(\Lambda) = U(\Lambda)\int (dp)\,|\vec{p}\rangle\langle\vec{p}|\ U(\Lambda)^\dagger &= \int (dp)\,|\Lambda(E,\vec{p}\,)\rangle\langle\Lambda(E,\vec{p})| \\ &= \int (dp)\,|(E',\vec{p}')\rangle\langle(E',\vec{p}')| = \int (dp')\,|\vec{p}'\rangle\langle\vec{p}'| = 1. \end{aligned}$$

From Eq. (1.7) one can easily show $U^\dagger U = 1$. Also,

$$U(\Lambda)(H,\hat{\vec{p}}\,)U(\Lambda)^{-1} = \Lambda^{-1}(H,\hat{\vec{p}}\,)$$

which is what we mean by relativistic co-variance.