QUANTUM FIELD THEORY

by James T. Wheeler

December 22, 2018

Contents

Chapter 1

Review of Hamiltonian Mechanics

1.1 From classical particles to quantum fields

First, let's review the use of the action in classical mechanics. I'll reproduce here a condensation of my notes from classical mechanics, available HERE. If you'd like a copy of the full version, just ask; what's here is more than enough for our purposes.

1.1.1 Hamiltonian Mechanics

Perhaps the most beautiful formulation of classical mechanics, and the one which ties most closely to quantum mechanics, is the canonical formulation. In this approach, the position and velocity velocity variables of Lagrangian mechanics are replaced by the position and conjugate momentum, $p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$. It turns out that by doing this the coordinates and momenta are put on an equal footing, giving the equations of motion a much larger symmetry.

To make the change of variables, we use a Legendre transformation. This may be familiar from thermodynamics, where the internal energy, Gibb's energy, free energy and enthalpy are related to one another by making different choices of the independent variables. Thus, for example, if we begin with

$$dU = TdS - PdV$$

where T and P are regarded as functions of S and V, we can set

$$H = U + VP$$

and compute

$$\begin{aligned} dH &= dU + PdV + VdP \\ &= TdS - PdV + PdV + VdP \\ &= TdS + VdP \end{aligned}$$

to achieve a formulation in which T and V are treated as functions of S and P.

The same technique works here. We have the Lagrangian, $L(q^i, \dot{q}^i)$ and wish to find a function $H(q_i, p_i)$. The differential of L is

$$\begin{aligned} dL &= \sum_{i=1}^{N} \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^{N} \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \\ &= \sum_{i=1}^{N} \dot{p}_i dq_i + \sum_{i=1}^{N} p_i d\dot{q}_i \end{aligned}$$

where the second line follows by using the equations of motion and the definition of the conjugate momentum. Therefore, set

$$H(q_i, p_i) = \sum_{i=1}^{N} p_i \dot{q}_i - L \tag{1.1}$$

so that

$$\begin{aligned}
dH &= \sum_{i=1}^{N} dp_i \dot{q}_i + \sum_{i=1}^{N} p_i d\dot{q}_i - dL \\
&= \sum_{i=1}^{N} dp_i \dot{q}_i + \sum_{i=1}^{N} p_i d\dot{q}_i - \sum_{i=1}^{N} \dot{p}_i dq_i - \sum_{i=1}^{N} p_i d\dot{q}_i \\
&= \sum_{i=1}^{N} dp_i \dot{q}_i - \sum_{i=1}^{N} \dot{p}_i dq_i
\end{aligned}$$

The function H is the *Hamiltonian*. In simple cases, it is of the same form as the energy.

Clearly, H is a function of the momenta. To see that we have really eliminated the dependence on velocity we may compute directly,

$$\begin{aligned}
\frac{\partial H}{\partial \dot{q}_j} &= \frac{\partial}{\partial \dot{q}_j} \left(\sum_{i=1}^{N} p_i \dot{q}_i - L(q_i, \dot{q}_i) \right) \\
&= \sum_{i=1}^{N} p_i \delta_{ij} - \frac{\partial L}{\partial \dot{q}_j} \\
&= p_j - \frac{\partial L}{\partial \dot{q}_j} \\
&= 0
\end{aligned}$$

so we have succeeded in replacing the velocity with the momentum.

The equations of motion are already built into the expression above for dH. Since the differential of H may always be written as

$$dH = \sum_{i=1}^{N} \frac{\partial H}{\partial q_j} dq_i + \sum_{i=1}^{N} \frac{\partial H}{\partial p_j} dp_i$$

we can simply equate the two expressions:

$$dH = \sum_{i=1}^{N} dp_i \dot{q}_i - \sum_{i=1}^{N} \dot{p}_i dq_i = \sum_{i=1}^{N} \frac{\partial H}{\partial q_i} dq_i + \sum_{i=1}^{N} \frac{\partial H}{\partial p_i} dp_i$$

Then, since the differentials dq_i and dp_i are all independent, we can equate their coefficients,

$$\dot{q}_i = \frac{\partial H}{\partial p_j} \tag{1.2}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{1.3}$$

These are *Hamilton's equations.*

1.1.1.1 Poisson brackets

Suppose we are interested in the time evolution of some function of the coordinates, momenta and time, $f(q_i, p_i, t)$. It could be any function – the area of the orbit of a particle, the period of an oscillating system, or one of the coordinates. The total time derivative of f is

$$\frac{df}{dt} = \sum \left(\frac{\partial f}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial f}{\partial p_i} \frac{dp_i}{dt} \right) + \frac{\partial f}{\partial t}$$

Using Hamilton's equations we may write this as

$$\frac{df}{dt} = \sum \left(\frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) + \frac{\partial f}{\partial t}$$

Define the Poisson bracket of H and f to be

$$\{H, f\} \equiv \sum_{i=1}^{N} \left(\frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right)$$

Then the total time derivative is given by

$$\frac{df}{dt} = \{H, f\} + \frac{\partial f}{\partial t} \tag{1.4}$$

If f has no explicit time dependence, so that $\frac{\partial f}{\partial t} = 0$, then the time derivative is given completely by the Poisson bracket:

$$\frac{df}{dt} = \{H, f\}$$

We generalize the Poisson bracket to two arbitrary functions,

$$\{f, g\} \equiv \sum_{i=1}^{N} \left(\frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i} - \frac{\partial g}{\partial p_i} \frac{\partial f}{\partial q_i} \right) \tag{1.5}$$

The importance of the Poisson bracket stems from the underlying invariance of Hamiltonian dynamics. Just as Newton's second law holds in any inertial frame, there is a class of *canonical coordinates* which preserve the form of Hamilton's equations. One central result of Hamiltonian dynamics is that any transformation that preserves certain fundamental Poisson brackets is canonical, and that such transformations preserve all Poisson brackets. Essentially all truly physical properties of a system can be expressed in terms of Poisson brackets.

In particular, we can write the equations of motion as Poisson bracket relations. Using the time evolution relation above we have

$$\begin{aligned} \frac{dq_i}{dt} &= \{H, q_i\} \\ &= \sum_{j=1}^{N} \left(\frac{\partial q_i}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial q_i}{\partial p_j} \frac{\partial H}{\partial q_j} \right) \\ &= \sum_{j=1}^{N} \delta_{ij} \frac{\partial H}{\partial p_j} \\ &= \frac{\partial H}{\partial p_i} \end{aligned}$$

and

$$\begin{aligned} \frac{dp_i}{dt} &= \{H, p_i\} \\ &= \sum_{j=1}^{N} \left(\frac{\partial p_i}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial p_i}{\partial p_j} \frac{\partial H}{\partial q_j} \right) \\ &= -\frac{\partial H}{\partial q_i} \end{aligned}$$

Notice that since q_i, p_i and are all independent, we have $\frac{\partial q_i}{\partial p_j} = \frac{\partial p_i}{\partial q_j} = 0$. Also, as coordinates, they are independent of time, $\frac{\partial q_i}{\partial t} = \frac{\partial p_i}{\partial t} = 0$ (The position of a particle may depend on time; the coordinates do not.)

We list some properties of Poisson brackets. Bracketing with a constant always gives zero

$$\{f, c\} = 0$$

The Poisson bracket is *linear*

$$\{af_1 + bf_2, g\} = a\{f_1, g\} + b\{f_2, g\}$$

and *Leibnitz*

$$\{f_1 f_2, g\} = f_2\{f_1, g\} + f_1\{f_2, g\}$$

These three properties are the defining properties of a *derivation*, which is the formal generalization of differentiation. The action of the Poisson bracket with any given function f on the class of all functions, $\{f, \cdot\}$ is therefore a derivation.

If we take the time derivative of a bracket, we can easily show

$$\frac{\partial}{\partial t}\{f, g\} = \{\frac{\partial f}{\partial t}, g\} + \{f, \frac{\partial g}{\partial t}\}$$

The bracket is antisymmetric

$$\{f, g\} = -\{g, f\}$$

and satisfies the Jacobi identity,

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$$

for all functions f, g and h. These properties are two of the three defining properties of a Lie algebra (the third defining property of a Lie algebra is that the set of objects considered, in this case the space of functions, be a finite dimensional vector space, while the space of functions is infinite dimensional).

Poisson's theorem is of considerable importance not only in classical physics, but also in quantum theory. Suppose f and g are constants of the motion. Then Poisson's theorem states that thier Poisson bracket, $\{f, g\}$, is also a constant of the motion. To prove the theorem, we start with f and g constant:

$$\frac{df}{dt} = \frac{dg}{dt} = 0$$

Then it follows that

$$\begin{aligned} \frac{df}{dt} &= \{H, f\} + \frac{\partial f}{\partial t} = 0 \\ \frac{dg}{dt} &= \{H, g\} + \frac{\partial g}{\partial t} = 0 \end{aligned}$$

Now consider the bracket:

$$\frac{d}{dt}\{f, g\} = \{H, \{f, g\}\} + \frac{\partial}{\partial t}\{f, g\}$$

Using the Jacobi identity on the first term on the right, and the relation for time derivatives on the second term, we have

$$\begin{aligned} \frac{d}{dt}\{f, g\} &= \{H, \{f, g\}\} + \frac{\partial}{\partial t}\{f, g\} \\ &= -\{f, \{g, H\}\} - \{g, \{H, f\}\} + \{\frac{\partial f}{\partial t}, g\} + \{f, \frac{\partial g}{\partial t}\} \\ &= \{f, \{H, g\}\} - \{g, \{H, f\}\} + \{\frac{\partial f}{\partial t}, g\} + \{f, \frac{\partial g}{\partial t}\} \\ &= \{f, \left(-\frac{\partial g}{\partial t}\right)\} - \{g, \left(-\frac{\partial f}{\partial t}\right)\} + \{\frac{\partial f}{\partial t}, g\} + \{f, \frac{\partial g}{\partial t}\} \\ &= 0 \end{aligned}$$

We conclude our discussion of Poisson brackets by using them to characterize canonical transformations.

1.1.1.2 Canonical transformations

Working with the Hamiltonian formulation of classical mechanics, we are allowed more transformations of the variables than with the Newtonian, or even the Lagrangian, formulations. We are now free to redefine our coordinates according to

$$\begin{aligned} q_i &= q_i(x_i, p_i, t) \\ \pi_i &= \pi_i(x_i, p_i, t) \end{aligned}$$

as long as the basic equations still hold.

It is straightforward to show that given any function $f = f(x_i, q_i, t)$ there is a canonical transformation defined by

$$\begin{aligned} p_i &= \frac{\partial f}{\partial x_i} \\ \pi_i &= -\frac{1}{\lambda}\frac{\partial f}{\partial q_i} \\ H' &= \frac{1}{\lambda}\left(H + \frac{\partial f}{\partial t}\right) \end{aligned}$$

The first equation

$$p_i = \frac{\partial f(x_i, q_i, t)}{\partial x_i}$$

gives q_i implicitly in terms of the original variables, while the second determines π_i. Notice that once we pick a function $q_i = q_i(p_i, x_i, t)$, the form of π_i is fixed. The third equation gives the new Hamiltonian in terms of the old one.

Sometimes it is more convenient to specify the new momentum $\pi_i(p_i, x_i, t)$ than the new coordinates $q_i = q_i(p_i, x_i, t)$. A Legendre transformation accomplishes this. Just replace $f = g - \lambda\pi_i q_i$. Then

$$\begin{aligned} df &= dg - d\pi_i q_i - \pi_i dq_i = p_i dx_i - \lambda\pi_i dq_i + (\lambda H' - H)\,dt \\ dg &= p_i dx_i + \lambda q_i d\pi_i + (\lambda H' - H)\,dt \end{aligned}$$

and we see that $g = g(x_i, \pi_i, t)$. In this case, g satisfies

$$\begin{aligned} p_i &= \frac{\partial g}{\partial x_i} \\ q_i &= \frac{1}{\lambda}\frac{\partial g}{\partial \pi_i} \\ H' &= \frac{1}{\lambda}\left(H + \frac{\partial g}{\partial t}\right) \end{aligned}$$

Since canonical transformations can interchange or mix up the roles of x and p, they are called *canonically conjugate*. Within Hamilton's framework, position and momentum lose their independent meaning except that variables always come in conjugate pairs. Notice that this is also a property of quantum mechanics.

Finally, we return to our earlier claim that transformations that preserve certain fundamental Poisson brackets, preserve Hamilton's equations and preserve all Poisson brackets. Specifically, a transformation from one set of phase space coordinates (x_i, π_i) to another (q_i, p_i) as canonical if and only if it preserves the *fundamental Poisson brackets*

$$\begin{aligned} \{q_i, q_j\}_{x\pi} &= \{p_i, p_j\}_{x\pi} = 0 \\ \{p_i, q_j\}_{x\pi} &= -\{q_i, p_j\}_{x\pi} = \delta_{ij} \end{aligned}$$

Here the subscript on the bracket, $\{\}_{x\pi}$ means that the partial derivatives defining the bracket are taken with respect to q_i and p_i. Brackets $\{f,g\}_{qp}$ taken with respect to the new variables (q_i,p_i) are identical to those $\{f,g\}_{x\pi}$ with respect to (x_i,π_i) if and only if the transformation is canonical. In particular, replacing f by H and g by any of the coordinate functions (x_i,π_i), we see that Hamilton's equations are preserved by canonical transformations.

1.1.1.3 Hamilton's equations from the action

It is possible to write the action in terms of x_i and p_i and vary these independently to arrive at Hamilton's equations of motion. We have

$$S = \int L dt \tag{1.6}$$

We can write this in terms of x_i and p_i easily:

$$\begin{aligned} S &= \int L dt \\ &= \int (p_i \dot{x}_i - H)\, dt \\ &= \int (p_i \mathbf{d} x_i - H \mathbf{d} t) \end{aligned}$$

Since S depends on position and momentum (rather than position and velocity), it is these we vary. Thus:

$$\begin{aligned} \delta S &= \delta \int (p_i \dot{x}_i - H)\, dt \\ &= \int \left(\delta p_i \dot{x}_i + p_i \delta \dot{x}_i - \frac{\partial H}{\partial x_i}\delta x_i - \frac{\partial H}{\partial p_i}\delta p_i \right) dt \\ &= p_i \delta x_i |_{t_1}^{t_2} + \int \left(\delta p_i \dot{x}_i - \dot{p}_i \delta x_i - \frac{\partial H}{\partial x_i}\delta x_i - \frac{\partial H}{\partial p_i}\delta p_i \right) dt \\ &= \int \left(\left(\dot{x}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i - \left(\dot{p}_i + \frac{\partial H}{\partial x_i} \right) \delta x_i \right) dt \end{aligned}$$

and since the variations δp_i and δx_i are independent we conclude

$$\dot{x}_i = \frac{\partial H}{\partial p_i} \tag{1.7}$$

$$\dot{p}_i = -\frac{\partial H}{\partial x_i} \tag{1.8}$$

as required.

1.1.1.4 Hamilton's principal function and the Hamilton-Jacobi equation

Properly speaking, the action is a functional, not a function. That is, the action is a function of *curves* rather than a function of *points* in space or phase space. We define Hamilton's principal function $\mathcal{S}$ in the following way. Pick an initial point of space and an initial time, and let $\mathcal{S}(x_i^{(f)},t)$ be the value of the action evaluated along the actual path that a physical system would follow in going from the initial time and place to $x_i^{(f)}$ at time t :

$$\mathcal{S}(x_i^{(f)},t) = S\,|_{physical} = \int_{t_0}^{t} L(x_i(t), \dot{x}_i(t), t) dt$$

where $x_i(t)$ is the solution to the equations of motion and $x_i^{(f)}$ is the final position at time t.

Now consider the variation of the action. Recall that in general,

$$\begin{aligned} \delta S &= \int_{t_0}^{t} \left(\frac{\partial L}{\partial x_i}\delta x_i + \frac{\partial L}{\partial \dot{x}_i}\delta \dot{x}_i \right) dt \\ &= \left[\frac{\partial L}{\partial \dot{x}_i}\delta x_i \right]_{t_0}^{t} + \int_{t_0}^{t} \left(\frac{\partial L}{\partial x_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} \right) \delta x_i dt \end{aligned}$$

Now suppose we hold the action constant at t_0, and require the equations of motion to hold. Then we have simply

$$\delta S \mid_{physical} = \frac{\partial L}{\partial \dot{x}_i}\delta x_i(t) = p_i \delta x_i$$

This means that the change in the $function\ \mathcal{S}$, when we change x_i by dx_i is

$$d\mathcal{S} = \delta S \mid_{physical} = p_i dx_i$$

of

$$\frac{\partial \mathcal{S}}{\partial x_i} = p_i$$

To find the dependence of $\mathcal{S}$ on t, we write $\mathcal{S} = S \mid_{physical} = \int L dt$ as

$$\frac{d\mathcal{S}}{dt} = L$$

But we also have

$$\frac{d\mathcal{S}}{dt} = \frac{\partial \mathcal{S}}{\partial x_i}\dot{x}_i + \frac{\partial \mathcal{S}}{\partial t}$$

Equating these and using $\frac{\partial \mathcal{S}}{\partial x_i} = p_i$ gives

$$\begin{aligned} L &= \frac{\partial \mathcal{S}}{\partial x_i}\dot{x}_i + \frac{\partial \mathcal{S}}{\partial t} \\ &= p_i \dot{x}_i + \frac{\partial \mathcal{S}}{\partial t} \end{aligned}$$

so that the partial of $\mathcal{S}$ with respect to t is

$$\frac{\partial \mathcal{S}}{\partial t} = L - p_i \dot{x}_i = -H$$

Combining the results for the derivatives of $\mathcal{S}$ we may write

$$\begin{aligned} d\mathcal{S} &= \frac{\partial \mathcal{S}}{\partial x_i}dx_i + \frac{\partial \mathcal{S}}{\partial t}dt \\ &= p_i dx_i - H dt \end{aligned}$$

This is a nontrivial condition on the solution of the classical problem. It means that form $p_i dx_i - H dt$ must be a total differential, which cannot be true for arbitrary p_i and H.

We conclude by stating the crowning theorem of Hamiltonian dynamics: for any Hamiltonian dynamical system there exists a canonical transformation to a set of variables on phase space such that the paths of motion reduce to single points. Clearly, this theorem shows the power of canonical transformations! The theorem relies on describing solutions to the Hamilton-Jacobi equation, which we introduce first.

We have the following equations governing Hamilton's principal function.

$$\begin{aligned} \frac{\partial \mathcal{S}}{\partial p_i} &= 0 \\ \frac{\partial \mathcal{S}}{\partial x_i} &= p_i \\ \frac{\partial \mathcal{S}}{\partial t} &= -H \end{aligned}$$

Since the Hamiltonian is a given function of the phase space coordinates and time, $H = H(x_i, p_i, t)$, we combine the last two equations:

$$\frac{\partial \mathcal{S}}{\partial t} = -H(x_i, p_i, t) = -H(x_i, \frac{\partial \mathcal{S}}{\partial x_i}, t)$$

This first order differential equation in $s+1$ variables $(t, x_i; i = 1, \ldots s)$ for the principal function $\mathcal{S}$ is the Hamilton-Jacobi equation. Notice that the Hamilton-Jacobi equation has the same general form as the Schrödinger equation (and is equally difficult to solve!). It is this similarity that underlies Dirac's canonical quantization procedure.

It is not difficult to show that once we have a solution to the Hamiltonian-Jacobi equation, we can immediately solve the entire dynamical problem. Such a solution may be given in the form

$$\mathcal{S} = g(t, x_1, \ldots, x_s, \alpha_1, \ldots, \alpha_s) + A$$

where the α_i are the additional s constants describing the solution. Now consider a canonical transformation from the variables (x_i, p_i) using the solution $g(t, x_i, \alpha_i)$ as the generating function. We treat the α_i as the new momenta, and introduce new coordinates β_i. Since g depends on the old coordinates x_i and the new momenta α_i, we have the relations

$$\begin{aligned}
p_i &= \frac{\partial g}{\partial x_i} \\
\beta_i &= \frac{\partial g}{\partial \alpha_i} \\
H' &= \left(H + \frac{\partial g}{\partial t}\right) \equiv 0
\end{aligned}$$

where the new Hamiltonian vanishes because g satisfies the Hamiltonian-Jacobi equation!. With $H' = 0$, Hamilton's equations in the new canonical coordinates are simply

$$\begin{aligned}
\frac{d\alpha_i}{dt} &= \frac{\partial H'}{\partial \beta_i} = 0 \\
\frac{d\beta_i}{dt} &= -\frac{\partial H'}{\partial \alpha_i} = 0
\end{aligned}$$

with solutions

$$\begin{aligned}
\alpha_i &= const. \\
\beta_i &= const.
\end{aligned}$$

The system remains at the phase space point (α_i, β_i). To find the motion in the original coordinates as functions of time and the $2s$ constants of motion, $x_i = x_i(t; \alpha_i, \beta_i)$, we can algebraically invert the s equations $\beta_i = \frac{\partial g(x_i, t, \alpha_i)}{\partial \alpha_i}$. The momenta may be found by differentiating the principal function, $p_i = \frac{\partial \mathcal{S}(x_i, t, \alpha_i)}{\partial x_i}$. This provides a complete solution to the mechanical problem.

We now apply these results to quantum theory.

1.1.2 Canonical Quantization

One of the most direct ways to quantize a classical system is the method of *canonical quantization* introduced by Dirac. The prescription is remarkably simple. Here we go:

A *dynamical variable* is any function of the phase space coordinates and time, $f(q_i, p_i, t)$. Given any two dynamical variables, we can compute their Poisson bracket,

$$\{f, g\}$$

as described in the previous section. In particular, the time evolution of any dynamical variable is given by

$$\frac{df}{dt} = \{H, f\} + \frac{\partial f}{\partial t}$$

and for any canonically conjugate pair of variables,

$$\{p_i, q_j\} = \delta_{ij}$$

To quantize the classical system, we let the canonically conjugate variables become operators (denoted by a "hat", $\hat{o}$), let all Poisson brackets be replaced by $\frac{i}{\hbar}$ times the *commutator* of those operators, and let all dynamical variables (including the Hamiltonian) become operators through their dependence on the conjugate variables:

$$\begin{aligned} \{\ ,\ \} &\rightarrow \frac{i}{\hbar}[\ ,\ ] && (1.9) \\ (p_i, q_j) &\rightarrow (\hat{p}_i, \hat{q}_j) && (1.10) \\ f(p_i, q_j, t) &\rightarrow \hat{f} = f(\hat{p}_i, \hat{q}_j, t) && (1.11) \end{aligned}$$

The operators are taken to act linearly on a vector space, and the vectors are called "states." This is all often summarized, a bit too succinctly, by saying "replace all Poisson brackets by commutators and put hats on everything."

The space of states

This simple set of rules works admirably, but we must say first what we actually mean by an operator. To do this, we must define a vector space on which they act. In keeping with the usual rules of quantum theory, we require a Hilbert space: a complete, comples, inner product vector space (a complete, normed vector space is called a Banach space). In general, we denote these vectors as kets, $|\psi\rangle$, where the label ψ may be any convenient list of properties. We then know that for any Hermitian operator has eigenstates with real eigenvalues, so we may define position and momentum eigenkets,

$$\begin{aligned} \hat{\mathbf{x}}\,|\mathbf{x}\rangle &= \mathbf{x}\,|\mathbf{x}\rangle \\ \hat{\mathbf{p}}\,|\mathbf{p}\rangle &= \mathbf{p}\,|\mathbf{p}\rangle \end{aligned}$$

and write the *Schrödinger equation*,

$$\hat{H}\,|\psi\rangle = i\hbar\frac{\partial}{\partial t}\,|\psi\rangle \qquad (1.12)$$

with the rules Eqs.(1.9) - (1.11), for canonical quantization leading to the Heisenberg equations of motion.

We can also arrive at the Schrödinger picture by choosing a set of *functions* as our vector space of states by placing the states in a coordinate basis

$$\psi(\mathbf{x}) \equiv \langle \mathbf{x}\,|\psi\rangle$$

Our Hilbert space is now a function space. Let $\psi(x)$ be an element of this vector space. Then we satisfy the fundamental commutators,

$$\begin{aligned} [\hat{p}_i, \hat{x}_j] &= -i\hbar\delta_{ij} \\ [\hat{x}_i, \hat{x}_j] &= 0 \\ [\hat{p}_i, \hat{p}_j] &= 0 \end{aligned}$$

if we represent the operators as

$$\begin{aligned} \hat{x}_i &= x_i \\ \hat{p}_i &= -i\hbar\frac{\partial}{\partial x_i} \\ \hat{H} &= i\hbar\frac{\partial}{\partial t} \end{aligned}$$

These relationships may be derived by acting on infinitesimal expansions of kets with the various operators. See Sakurai for details.

The representation of $\hat{x}_i$ by x_i simply means we replace the operator by the coordinate. Now consider the time evolution of a state ψ. This is given by the action of the Eq.(1.12). If the Hamiltonian is that of a single particle moving in a potential,

$$\hat{H}\psi \quad = \quad \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}})$$

then substitution of the coordinate forms of the operators immediately gives the familiar form of the Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{x})\psi = i\hbar\frac{\partial\hat{\psi}}{\partial t} \tag{1.13}$$

Notice that while ψ is mathematically a field, this is not quantum field theory. The difference between this Schrödinger field theory and quantum field theory is that the dynamical variables of quantum mechanics are positions, momenta and so forth, while the dynamical variables of field theory are the fields. When this is realized, we find little difference between the canonical quantization of mechanics and the canonical quantization of field theory.

This change from particle mechanics to fields is thus not really a change in method at all, the only mathematical difference being in the way we take derivatives. In classical field theory we replace the Lagrangian with a Lagrangian density,

$$L = \int \mathcal{L}(\phi, \partial_\alpha\phi)\, d^3x$$

where the density $\mathcal{L}$ is generally built from the fields and their derivatives, $\mathcal{L} = \mathcal{L}(\phi, \partial_\alpha\phi)$. This means that the Lagrangian is a functional of the fields and partial derivatives become functional derivatives,

$$\frac{\partial L}{\partial \dot{q}} \quad \Longrightarrow \quad \frac{\delta L}{\delta \dot{\phi}}$$

We will define functional derivatives when they are required.

Operator ordering

One of our rules, eq.(1.14), however, still requires modification.

The point requiring caution with Eq.(1.11) is *ordering ambiguity.* The problem arises when the Hamiltonian, or any other dynamical variable of interest, depends in a more complicated way on position and momentum. The simplest example is a Hamiltonian containing a term of the form

$$H_1 = \alpha\mathbf{p}\cdot\mathbf{x}$$

For the classical variables, $\mathbf{p}\cdot\mathbf{x} = \mathbf{x}\cdot\mathbf{p}$, and we could equivalently write

$$H_\beta = \alpha\left(\beta\mathbf{p}\cdot\mathbf{x} + (1-\beta)\,\mathbf{x}\cdot\mathbf{p}\right)$$

for any real number β. But since quantum operators don't commute

$$\begin{aligned}\hat{H}_\beta &= \alpha\left(\beta\hat{\mathbf{p}}\cdot\hat{\mathbf{x}} + (1-\beta)\,\hat{\mathbf{x}}\cdot\hat{\mathbf{p}}\right)\\ &= \alpha\left(\hat{\mathbf{p}}\cdot\hat{\mathbf{x}} - 3i\,(1-\beta)\,\hbar\right)\end{aligned}$$

is a different operator for every β. In many circumstances the symmetric choice

$$\hat{H}_1 = \frac{\alpha}{2}\left(\hat{\mathbf{p}}\cdot\hat{\mathbf{x}} + \hat{\mathbf{x}}\cdot\hat{\mathbf{p}}\right)$$

turns out to be preferable, and certain rules of thumb exist.

This issue also means that, unlike Poisson brackets, commutators are order-specific. Thus, we can write the Leibnitz rule as

$$\left[\hat{A}, \hat{B}\hat{C}\right] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}$$

but must remember that

$$\left[\hat{A}, \hat{B}\hat{C}\right] \neq \left[\hat{A}, \hat{C}\right]\hat{B} + \left[\hat{A}, \hat{B}\right]\hat{C}$$

For now it is enough to be aware of the problem.

A particularly important case is that of the harmonic oscillator. Given the conjugate pair, $(\hat{p}_i, \hat{q}_j)$, we define a new pair of operators,

$$\begin{aligned}
\hat{a}_i &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{q}_i + \frac{i}{m\omega}\hat{p}_i\right) \\
\hat{a}_i^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{q}_i - \frac{i}{m\omega}\hat{p}_i\right)
\end{aligned}$$

which satisfy

$$\begin{aligned}
\left[\hat{a}_j^\dagger, \hat{a}_i\right] &= \frac{m\omega}{2\hbar}\left[\left(\hat{q}_i - \frac{i}{m\omega}\hat{p}_i\right), \left(\hat{q}_j + \frac{i}{m\omega}\hat{p}_j\right)\right] \\
&= \frac{i}{2\hbar}\left[\hat{q}_i, \hat{p}_j\right] - \frac{i}{2\hbar}\left[\hat{p}_i, \hat{q}_j\right] \\
&= \delta_{ij} \\
\left[\hat{a}_i, \hat{a}_j\right] &= \left[\hat{a}_i^\dagger, \hat{a}_j^\dagger\right] = 0
\end{aligned}$$

Rewriting the Hamiltonian in terms of the new variables, we find

$$H = \hbar\omega\left(\sum \hat{a}_i^\dagger \hat{a}_i + \frac{3}{2}\right) \tag{1.14}$$

where the factor of 3 occurs because we have 3 independent oscillators. While are used to the idea of an arbitrary zero point of energy, it is not so arbitrary in quantum mechanics since we know there should be a minimum uncertainty to the momentum. This is not a problem for the simple harmonic oscillator, but when we quantize fields we will find that the each of the infinite number of Fourier modes acts like $\hat{a}_i$ or $\hat{a}_i^\dagger$, meaning that the vacuum energy *diverges*. At this point, we introduce *normal ordering*: a rule for operator ordering chosen to eliminate the infinity.

Exercise: Canonically quantize the 3-dim simple harmonic oscillator. Find the form of Hamilton's equations of motion, and show that the Hamiltonian takes the form given in Eq.(1.14).

1.1.3 One dimensional field theory

Consider a 1-dimensional distribution of equal point masses, m, distributed along the x-axis at positions x_i. As the number of these particles increases, we may define a density.

Consider just one small length Δl. In Δl the density is

$$\lambda(x) \approx \frac{\sum_{i\in\Delta l} m_i}{\Delta l}$$

The kinetic energy of this small length is then giving by its motion relative to its center of mass, x_0. Setting the displacement to $\phi = x - x_0$

$$\begin{aligned}
T &= \frac{1}{2}\left(\lambda(x_0)\Delta l\right)\dot{x}^2 \\
&= \frac{1}{2}\left(\lambda(x_0)\Delta l\right)\dot{\phi}^2
\end{aligned}$$

Now suppose the particles in ΔL move in a slowly changing potential, $V(x)$, that depends only on the displacement of each particle group from equilibrium, defined by

$$\begin{aligned} V(x_0) &= 0 \\ \frac{dV}{dx}(x_0) &= 0 \end{aligned}$$

Then expanding V in a Taylor series, we have

$$\begin{aligned} V(x) &\approx \frac{1}{2}\left.\frac{d^2V}{dx^2}\right|_{x=x_0}(x-x_0)^2 \\ &= \frac{1}{2}(\Delta l)^2\left.\frac{d^2V}{dx^2}\right|_{x=x_0}\frac{(x-x_0)^2}{(\Delta l)^2} \end{aligned}$$

we may treat the ratio $\frac{x-x_0}{\Delta l}$ as a derivative of the displacement, $\phi(x_0) = x - x_0$ of the cell centered on x_0

$$\frac{x-x_0}{\Delta l} \approx \frac{d\phi}{dx}$$

Let the constant be $\sigma \equiv (\Delta l)\left.\frac{d^2V}{dx^2}\right|_{x=x_0}$.

Combining these results, the Lagrangian is

$$\begin{aligned} \Delta L &= \Delta T - \Delta V \\ &= \frac{1}{2}\left[\lambda(x_0)\dot{\phi}^2 - \sigma(x_0)\left(\frac{d\phi}{dx}\right)^2\right]\Delta l(x_0) \end{aligned}$$

Now we add up over all of the small lengths $\Delta l(x_0)$ and take the limit to get an integral,

$$L = \frac{1}{2}\int\left[\lambda(x)\dot{\phi}^2 - \sigma(x)\left(\frac{d\phi}{dx}\right)^2\right]dx$$

For a uniform material with the displacement ϕ small compared to Δl, the functions $\lambda(x)$ and $\sigma(x)$ may be taken constant. Then, defining $\frac{1}{v^2} \equiv \frac{\lambda}{\sigma}$, the Lagrangian becomes

$$L = -\frac{\sigma}{2}\int\left[-\frac{1}{v^2}\dot{\phi}^2 + \left(\frac{d\phi}{dx}\right)^2\right]dx$$

Finally, integrating to find the action,

$$\begin{aligned} S &= \int L dt \\ &= -\frac{\sigma}{2}\int\int dxdt\left[-\frac{1}{v^2}\dot{\phi}^2 + \left(\frac{d\phi}{dx}\right)^2\right] \end{aligned}$$

The action is now expressed as an integral over both space *and* time. Varying ϕ, we find

$$\begin{aligned} 0 &= \delta S \\ &= -\frac{\sigma}{2}\int\int dxdt\left[-2\frac{1}{v^2}\dot{\phi}\delta\dot{\phi} + 2\frac{d\phi}{dx}\frac{d\delta\phi}{dx}\right] \\ &= -\sigma\int\int dxdt\left[-\frac{1}{v^2}\frac{d}{dt}\left(\dot{\phi}\delta\phi\right) + \frac{1}{v^2}\ddot{\phi}\delta\phi + \frac{d}{dx}\left(\frac{d\phi}{dx}\delta\phi\right) - \frac{d^2\phi}{dx^2}\delta\phi\right] \end{aligned}$$

Integrating the total derivative terms to the time or space boundary, the variation vanishes and we are left with

$$0 = -\sigma \int\int dxdt \left[\frac{1}{v^2}\ddot{\phi} - \frac{d^2\phi}{dx^2}\right]\delta\phi$$

Since $\delta\phi$ is arbitrary we must have

$$\frac{1}{v^2}\ddot{\phi} - \frac{d^2\phi}{dx^2} = 0$$

This is the wave equation for a field ϕ that propagates with velocity v.

If we were to do a similar calculation for a 3-dim array of particles we could arrive at

$$S = \frac{\sigma}{2}\int\int dxdt \left[\frac{1}{v^2}\dot{\phi}^2 - (\boldsymbol{\nabla}\phi)^2\right]$$

The variation then leads to the full wave equation

$$\frac{1}{v^2}\ddot{\phi} - \nabla^2\phi = 0$$

If the velocity is the speed of light this is the massless Klein-Gordon equation. This is the simplest relativistic wave equation. If we wish to include a mass for ϕ, we arrive at the Klein-Gordon Lagrangian:

$$S_{KG} = \frac{1}{2}\int\int d^3x dt \left[\frac{1}{c^2}\dot{\phi}^2 - (\boldsymbol{\nabla}\phi)^2 + \frac{m^2c^2}{\hbar^2}\phi^2\right]$$

Exercise: Vary S_{KG} to find the massive Klein-Gordon equation,

$$\Box\phi \equiv \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} - \nabla^2\phi = -\frac{m^2c^2}{\hbar^2}\phi \tag{1.15}$$

The wave operator $\Box \equiv -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2$ is called the *d'Alembertian.*

The Klein-Gordon action is typical of field theories, which have the general form

$$\begin{aligned} S &= \int dt \left(\int \mathcal{L} d^3x\right) \\ &= \int d^4x\, \mathcal{L}\left(\Phi, \dot{\Phi}\right) \end{aligned}$$

where Φ may be any collection of different types of field. $\mathcal{L}$ is called the *Lagrange density.* Note how the integral is over both space and time, making it straightforward to write relativistic theories.

1.1.4 Canonical quantization of a field theory

Without going into careful detail yet, we can see some features of the quantization of a field theory. Let's consider the action for the relativistic scalar field ϕ. Well use Greek indices for spacetime $\alpha, \beta, \ldots = 0, 1, 2, 3$ and Latin for space $i, j, \ldots = 1, 2, 3$. Let's write

$$\partial_\alpha = (\partial_0, \partial_i)$$

where

$$\partial_0 = \frac{1}{c}\frac{\partial}{\partial t}$$

We'll use the metric

$$\eta_{\alpha\beta} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

to raise and lower indices. For example, we can write the d'Alembertian $\Box$ as

$$\begin{aligned} \Box &= \eta^{\alpha\beta}\partial_\alpha\partial_\beta \\ &= \partial^\alpha\partial_\alpha = \partial_\alpha\partial^\alpha \end{aligned}$$

where

$$\partial^\alpha \equiv \eta^{\alpha\beta}\partial_\beta = (\partial_0, -\partial_i)$$

With this notation, the action for the relativistic wave equation is

$$\begin{aligned} S &= \frac{1}{2}\int \left(\dot{\phi}^2 - \nabla\phi\cdot\nabla\phi\right) d^4x \\ &= \frac{1}{2}\int \partial_\alpha\phi\partial^\alpha\phi \; d^4x \end{aligned}$$

The relativistic summation convention *always* involves one raised index and one lowered index. *Euclidean* summations are written with both the repeated indices in the same position. Thus, $\partial_\alpha\partial_\alpha = (\partial_0)^2 + \nabla^2$ is the 4-dimensional *Euclidean* Laplacian.

Now we can illustrate the quantization. We know that the field ϕ is the limit of an uncountable infinity of independent particle coordinates, so all we need to set up the canonical commutator is its conjugate momentum. The usual conjugate momentum is

$$\pi = \frac{\partial L}{\partial\dot{\phi}}$$

but now the Lagrangian is a *functional* of the fields,

$$\begin{aligned} L &= \frac{1}{2}\int \partial_\alpha\phi\partial^\alpha\phi \; d^3x \\ &= \frac{1}{2}\int \left(\dot{\phi}^2 - \nabla\phi\cdot\nabla\phi\right) \; d^3x \end{aligned}$$

so we write a functional derivative, denoted by $\frac{\delta}{\delta\dot{\phi}(x')}$,

$$\begin{aligned} \pi &= \frac{\delta L}{\delta\dot{\phi}(\mathbf{y})} \\ &= \frac{1}{2}\frac{\delta}{\delta\dot{\phi}(\mathbf{y})}\int \left(\dot{\phi}(\mathbf{x})^2 - \nabla\phi(\mathbf{x})\cdot\nabla\phi(\mathbf{x})\right) \; d^3x \\ &= \frac{1}{2}\int \frac{\delta}{\delta\dot{\phi}(y^j)}\dot{\phi}\left(x^i\right)^2 \; d^3x \\ &= \frac{1}{2}\int 2\dot{\phi}\left(x^i\right)\frac{\delta\dot{\phi}\left(x^i\right)}{\delta\dot{\phi}\left(y^j\right)} \; d^3x \end{aligned}$$

The functional derivative behaves here like an ordinary derivative, leaving us with

$$\frac{\delta\dot{\phi}\left(x^i\right)}{\delta\dot{\phi}\left(y^j\right)}$$

This vanishes unless $x^i = y^i$, but to know what it is we make an analogy with the finite case, where

$$\frac{\partial x^i}{\partial y^j} = \delta^i_j$$

Holding j fixed and summing over i gives 1, and we ask the same for the continuous case where the sum becomes an integral:

$$\int d^3x \, \frac{\delta\dot{\phi}\left(x^i\right)}{\delta\dot{\phi}\left(y^j\right)} = 1$$

Since we know that $\frac{\delta\dot{\phi}\left(x^i\right)}{\delta\dot{\phi}\left(y^j\right)}$ vanishes when $\mathbf{x} \neq \mathbf{y}$, this implies

$$\frac{\delta\dot{\phi}\left(x^i\right)}{\delta\dot{\phi}\left(y^j\right)} = \delta^3\left(\mathbf{x}-\mathbf{y}\right)$$

Therefore, the conjugate momentum to ϕ is

$$\begin{aligned}\pi &= \int \dot{\phi}\left(\mathbf{x}\right)\delta^3\left(\mathbf{x}-\mathbf{y}\right)\, d^3x \\ &= \dot{\phi}\left(\mathbf{y}\right)\end{aligned}$$

In quantizing, we change the dynamical variables ϕ and π to operators, $\hat{\phi}$ and $\hat{\pi}$, and their Poisson bracket becomes a commutator

$$\left[\hat{\pi}, \hat{\phi}\right] = -i\hbar \tag{1.16}$$

Before continuing with further details of relativistic quantization, we need two things. First, we prove Noether's theorem, which relates symmetries to conserved quantities. The relationship is central to our understanding of field theory. Second, in the next chapter, we develop group theory both because of the relationship of group symmetries to conservation laws and because it is from group theory that we learn the types of fields that are important in physics, including spinors. Then we will return to quantization.

1.1.5 Special Relativity

Since we have just introduced some relativitistic notation, this seems like a good place to review special relativity, and especially the reason that the notation is meaningful.

1.1.5.1 The invariant interval

The first thing to understand clearly is the difference between physical quantities such as the length of a ruler or the elapsed time on a clock, and the coordinates we use to label locations in the world. In 3-dim Euclidean geometry, for example, the length of a ruler is given in terms of coordinate intervals using the Pythagorean theorem. Thus, if the positions of the two ends of the ruler are (x_1, y_1, z_1) and (x_2, y_2, z_2), the length is

$$L = \sqrt{\left(x_2-x_1\right)^2 + \left(y_2-y_1\right)^2 + \left(z_2-z_1\right)^2}$$

Observe that the actual values of (x_1, y_1, z_1) are irrelevant. Sometimes we choose our coordinates cleverly, say, by aligning the x-axis with the ruler and placing one end at the origin so that the endpoints are at $(0,0,0)$ and $(x_2, 0, 0)$. Then the *calculation* of L is trivial:

$$\begin{aligned}L &= \sqrt{\left(x_2-x_1\right)^2 + \left(y_2-y_1\right)^2 + \left(z_2-z_1\right)^2} \\ &= x_2\end{aligned}$$

but it is still important to recognize the difference between the *coordinates* and the *length*.

With this concept clear, we next need a set of labels for spacetime. Starting with a blank page to represent spacetime, we start to construct a set of labels. First, since all observers agree on the motion of light, let's agree that (with time flowing roughly upward in the diagram and space extending left and right) light beams always move at 45 degrees in a straight line. An inertial observer (whose constant rate of motion has no absolute reality; we only consider the relative motions of two observers) will move in a straight line at a steeper angle than 45 degrees – a lesser angle would correspond to motion faster than the speed of light. For any such inertial observer, we let the time coordinate be the time as measured by a clock they carry. The ticks of this clock provide a time scale along the straight, angled world line of the observer. To set spatial coordinates, we use the constancy of the speed of light. Suppose our inertial observer send out a pulse of light at 3 minutes before noon, and suppose the nearby spacetime is dusty enough that bits of that pulse are reflected back continuously. Then some reflected light will arrive back at the observer at 3 minutes after noon. Since the trip out and the trip back must have taken the same length of time and occurred with the light moving at constant velocity, the reflection of the light by the dust particle must have occurred at noon in our observer's frame of reference. It must have occurred at a distance of 3 light minutes away. If we take the x direction to be the direction the light was initially sent, the location of the dust particle has coordinates $(noon, 3\, light\, minutes, 0, 0)$. In a similar way, we find the locus of all points with time coordinate $t = noon$ and both $y = 0$ and $z = 0$. These points form our x axis. We find the y and z axes in the same way. It is somewhat startling to realize when we draw a careful diagram of this construction, that the x axis seems to make an acute angle with the time axis, as if the time axis has been reflected about the 45 degree path of a light beam. We quickly notice that this must always be the case if all observers are to measure the same speed ($c = 1$ in our construction) for light.

This gives us our labels for spacetime *events*. Any other set of labels would work just as well. In particular, we are interested in those other sets of coordinates we get by choosing a different initial world line of an different inertial observer.

Any observer, in assigning coordinates $x^\alpha = (ct, x, y, z)\,, \alpha = 0, 1, 2, 3$ to an event P, is specifying a *vector*, and just as in classical mechanics we expect to be able to write vector equations for the motions of objects in spacetime. This is an important distinction between special and general relativity. In general relativity, to work with vectors we must use tangent spaces to the spacetime manifold but in special relativity, Minkowski spacetime is a vector space. In order to map one observer's vectors positions to another observer's vector positions, we must use a linear transformation. Thus, if a second observer at the same location but moving with a relative velocity v assigns $x'^\alpha = (ct', x', y', z')$ to the same event P, then there must be a linear relationship of the form

$$x'^\alpha = \sum_{\beta=0}^{3} \Lambda^\alpha{}_\beta x^\beta \tag{1.17}$$

where $\Lambda^\alpha{}_\beta$ is a 4×4 matrix dependent upon the velocity. Suppose a flash of light is emitted from the origin at the moment that the observers pass. Each will describe an sphere of light expanding with velocity c,

$$\begin{aligned} x^2 + y^2 + z^2 &= c^2 t^2 \\ x'^2 + y'^2 + z'^2 &= c^2 t'^2 \end{aligned}$$

Since the quantity $x^2 + y^2 + z^2 - c^2 t^2$ vanishes simultaneously for both observers, we must have

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = \lambda \left(x^2 + y^2 + z^2 - c^2 t^2 \right)$$

The factor λ is a conformal factor, which is allowed but we may set it to 1 here. Then if the relative motion of the two observers lies in the x direction they can directly compare intervals in the $y$and z directions and will find them the same. The transformation between the two reference frames therefore involves only x, t and x', t', and we may write

$$\begin{aligned} x' &= ax + bt \\ y' &= cx + dt \end{aligned}$$

and require

$$x'^2 - c^2t'^2 = x^2 - c^2t^2$$

Let $u = x + ct$ and $v = x - ct$, and similarly for u', v'. Then the transformation may be written as

$$\begin{aligned} u' &= \alpha u + \beta v \\ v' &= \mu u + \nu v \end{aligned}$$

condition becomes

$$\begin{aligned} u'v' &= uv \\ (\alpha u + \beta v)(\mu u + \nu v) &= uv \\ \alpha\mu u^2 + (\alpha\nu + \beta\mu) vu + \beta\nu v^2 &= uv \end{aligned}$$

so we must have either

$$\begin{aligned} \beta &= \mu = 0 \\ \nu &= \frac{1}{\alpha} \end{aligned} \tag{1.18}$$

or

$$\begin{aligned} \alpha &= \nu = 0 \\ \mu &= \frac{1}{\beta} \end{aligned} \tag{1.19}$$

Choosing the first and setting $\alpha = e^\xi$ we have

$$\begin{aligned} u' &= e^\xi u \\ v' &= e^{-\xi} v \end{aligned}$$

and therefore

$$\begin{aligned} x' + ct' &= e^\xi (x + ct) \\ x' - ct' &= e^{-\xi} (x - ct) \end{aligned}$$

Adding and subtracting to solve for x' and y',

$$\begin{aligned} x' &= \frac{1}{2}\left(e^\xi (x + ct) + e^{-\xi} (x - ct)\right) \\ &= x \cosh\xi + ct \sinh\xi \\ ct' &= \frac{1}{2}\left(e^\xi (x + ct) - e^{-\xi} (x - ct)\right) \\ &= x \sinh\xi + ct \cosh\xi \end{aligned}$$

so our matrix is

$$\Lambda^\alpha{}_\beta = \begin{pmatrix} \cosh\xi & \sinh\xi & & \\ \sinh\xi & \cosh\xi & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

Finally, taking the differential of the transformation,

$$\begin{aligned} dx' &= dx \cosh\xi + cdt \sinh\xi \\ cdt' &= dx \sinh\xi + cdt \cosh\xi \end{aligned}$$

and taking the ratio,

$$\begin{aligned}\frac{1}{c}\frac{dx'}{dt'} &= \frac{dx\,\cosh\xi + cdt\,\sinh\xi}{dx\,\sinh\xi + cdt\,\cosh\xi} \\ &= \frac{\frac{1}{c}\frac{dx}{dt}\,\cosh\xi + \sinh\xi}{\frac{1}{c}\frac{dx}{dt}\,\sinh\xi + \cosh\xi} \\ &= \frac{\frac{1}{c}\frac{dx}{dt} + \tanh\xi}{\frac{1}{c}\frac{dx}{dt}\,\tanh\xi + 1}\end{aligned}$$

we identify the velocity of an object in each of the two frames as $u = \frac{dx}{dt}$ and $u' = \frac{dx'}{dt'}$. Multiplying by c, we have

$$u' = \frac{u + \tanh\xi c}{\frac{u}{c}\tanh\xi + 1} \tag{1.20}$$

In the limit of small boost parameter, $\xi \ll 1$, we have $\tanh\xi \approx \xi$. Then taking all velocities small compared to the speed of light and neglecting the very small second order quantity $\frac{u}{c}\xi$ in the denominator, we identify the Newtonian addition of velocities

$$u' = u + v$$

provided we identify

$$\tanh\xi = \frac{v}{c}$$

Then $\cosh\xi = \frac{1}{\sqrt{1-\tanh^2\xi}} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \equiv \gamma$ and $\sinh\xi = \frac{\tanh\xi}{\sqrt{1-\tanh^2\xi}} = \frac{v/c}{\sqrt{1-\frac{v^2}{c^2}}} = \gamma\frac{v}{c}$. The Lorentz transformation becomes,

$$\Lambda^{\alpha}{}_{\beta} = \begin{pmatrix} \gamma & \gamma\frac{v}{c} & & \\ \gamma\frac{v}{c} & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \tag{1.21}$$

and the relationship between the velocities, Eq.(1.20) becomes

$$u' = \frac{u+v}{1+\frac{uv}{c^2}}$$

The next step is the most important: we must find a way to write physically meaningful quantities. These quantities, like length in Euclidean geometry, must be independent of the labels, the coordinates, that we put on different points. If we get on the right track by forming a quadratic expression similar to the Pythagorean theorem, then it doesn't take long to arrive at the correct answer. In spacetime, we have a *pseudo*-Euclidean length interval, given by the *proper time* τ, where

$$c^2\tau^2 = c^2t^2 - x^2 - y^2 - z^2 \tag{1.22}$$

Exercise: Computeτ' in terms of τ by writing it in the primed frame:

$$c^2\tau'^2 = c^2t'^2 - x'^2 - y'^2 - z'^2$$

and substituting using Eqs.(1.17) and (1.21).

Exercise: Find the form of transformation arising from the second solution, Eq.(1.19). What situation does it correspond to?

Tau, τ, is called the proper time, and is invariant under Lorentz transformations. It plays the role of L in spacetime geometry, and becomes the defining property of spacetime symmetry: *we define Lorentz transformations to be those transformations that leave τ invariant.*

1.1.5.2 Lorentz transformations

Notice that with this definition, 3-dim rotations are included as Lorentz transformations because τ only depends on the Euclidean length $x^2 + y^2 + z^2$; any transformation that leaves this length invariant also leaves τ invariant. Lorentz transformations that map the three spatial directions into one another are called rotations, while Lorentz transformations that involve time and velocity are called *boosts*. There are therefore 6 independent Lorentz transformations: three planes $((xy), (yz), (zx))$ of rotation and three planes $((tx), (ty), (tz))$ of boosts.

We will always use the Einstein convention: Lorentz invariant summed indices always occur in pairs with one up and one down. This allows us to omit the summation symbol in Eq.(1.17) and write

$$(x')^\alpha = \sum_{\beta=1}^{3} \Lambda^\alpha{}_\beta x^\beta \equiv \Lambda^\alpha{}_\beta x^\beta \tag{1.23}$$

where we assume a sum on β.

Any object that transforms in this same linear, homogeneous way, where $\Lambda^\alpha{}_\beta$ is any boost or rotation matrix, is called a *Lorentz vector* or a 4-vector. Thus, if

$$w'^\alpha = \Lambda^\alpha{}_\beta w^\beta$$

then w^α is a 4-vector.

We have seen that the proper time is invariant under Lorentz transformations. In fact, a linear transformation is a Lorentz transformation if and only if it leaves all proper times invariant. use this to express Lorentz transformations in terms of a *metric*. Let

$$\eta_{\alpha\beta} \equiv \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{1.24}$$

as given in the previous section. Then the interval spanned by a 4-vector x^α is

$$\begin{aligned} c^2\tau^2 &= \eta_{\alpha\beta} x^\alpha x^\beta \\ &= \begin{pmatrix} ct & x & y & z \end{pmatrix} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \\ &= c^2t^2 - x^2 - y^2 - z^2 \end{aligned}$$

It is convenient to define two different forms of any 4-vector, called covariant (x_α) and contravariant (x^α). These two forms exist anytime we have a metric. We define

$$x_\alpha \equiv \eta_{\alpha\beta} x^\beta \tag{1.25}$$

then we can write invariant intervals as

$$c^2\tau^2 = x_\beta x^\beta = x^\beta x_\beta$$

where the second expression uses the symmetry of the metric, $\eta_{\alpha\beta} = \eta_{\beta\alpha}$.

The defining property of a Lorentz transformation can now be written in a way that doesn't depend on the coordinates. Invariance of the interval requires

$$c^2\tau^2 = \eta_{\alpha\beta} x^\alpha x^\beta = \eta_{\alpha\beta} (x')^\alpha (x')^\beta$$

so that for any Lorentz vector x^β,

$$\begin{aligned}
\eta_{\mu\nu} x^\mu x^\nu &= \eta_{\alpha\beta} (x')^\alpha (x')^\beta \\
&= \eta_{\alpha\beta} \left(\Lambda^\alpha{}_\mu x^\mu\right) \left(\Lambda^\beta{}_\nu x^\nu\right) \\
&= \left(\eta_{\alpha\beta} \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu\right) x^\mu x^\nu
\end{aligned}$$

Since x^μ is arbitrary, and $\eta_{\alpha\beta}$ is symmetric, this implies

$$\eta_{\mu\nu} = \eta_{\alpha\beta} \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \tag{1.26}$$

From now on, we will take this as the *defining property of a Lorentz transformation.*

Suppose w^α is any 4-vector, so that

$$w'^\alpha = \Lambda^\alpha{}_\beta w^\beta \tag{1.27}$$

where $\Lambda^\alpha{}_\beta$ a Lorentz transformatoin. It follows immediately that $w_\alpha w^\alpha$ is invariant under Lorentz transformations.

Exercise Prove that $w^\alpha w_\alpha = \eta_{\alpha\beta} w^\alpha w^\beta$ is invariant under Lorentz transformation (that is, $\eta_{\alpha\beta} w'^\alpha w'^\beta = \eta_{\alpha\beta} w^\alpha w^\beta$, using $w'^\alpha = \Lambda^\alpha{}_\beta w^\beta$ and Eq.(1.26).

Exercise Prove that if w^α transforms according to Eq.(1.27) and the metric according to Eq.(1.26) then w_α transforms as $w'_\alpha = \bar{\Lambda}^\beta{}_\alpha w_\beta$ where $\bar{\Lambda}^\beta{}_\alpha$ is the inverse matrix to $\Lambda^\beta{}_\alpha$. Use this fact to give an alternative proof that $w^\alpha w_\alpha$ is Lorentz invariant.

As long as we are careful to use only quantities that have such simple transformations (i.e., linear and homogeneous) it is easy to construct Lorentz invariant quantities by "contracting" indices. Any time we sum one contravariant vector index with one covariant vector index, we produce an invariant.

It is not hard to derive dynamical variables which are Lorentz vectors. Suppose we have a path in spacetime (perhaps the path of a particle), specified parametrically by $x^\beta(\lambda)$, so as λ increases, $x^\beta(\lambda)$ gives the coordinates of the particle. We can even let λ be the proper time along the world line of the particle, since this increases monotonically as the particle moves along. In fact, this is an excellent choice. To compute the parameter, consider an infinitesimal displacement along the path, dx^β. Then the change in the proper time for that displacement is

$$\begin{aligned}
d\tau &= \left(\eta_{\alpha\beta} dx^\alpha dx^\beta\right)^{1/2} \\
&= \left(dt^2 - \frac{1}{c^2}\left(dx^i\right)^2\right)^{1/2}
\end{aligned}$$

where the Latin index runs over the spatial coordinates so that $dx^i dx^i$ is the usual Euclidean interval. Now we can integrate the infinitesimal proper time along the path to a general point at proper time τ :

$$\begin{aligned}
\tau &= \int d\tau \\
&= \int \sqrt{dt^2 - \frac{1}{c^2}\left(dx^i\right)^2} \\
&= \int dt \sqrt{1 - \frac{1}{c^2}\left(\frac{dx^i}{dt}\right)^2} \\
&= \int dt \sqrt{1 - \frac{\mathbf{v}^2(t)}{c^2}}
\end{aligned}$$

As soon as we know the path $\mathbf{x}(t)$, we can differentiate to find $\mathbf{v}(t)$, integrate to find $\tau(t)$, and invert to find $t(\tau)$. This gives $x^\alpha(\tau) = (t(\tau), \mathbf{x}(\tau))$. Note the useful relationship between infinitesimals,

$$d\tau = dt\sqrt{1 - \frac{\mathbf{v}^2(t)}{c^2}}$$

or

$$\gamma d\tau = dt$$

Once we have the path parameterized in terms of proper time, we can find the tangent to the path simply by differentiating:

$$u^{\beta} = \frac{dx^{\beta}}{d\tau} \tag{1.28}$$

Since τ is Lorentz invariant and the Lorentz transformation matrix is constant (between two given inertial frames), we have

$$\begin{aligned} (u')^{\beta} &= \frac{d(x')^{\beta}}{d\tau'} \\ &= \frac{d\left(\Lambda^{\beta}{}_{\alpha}x^{\alpha}\right)}{d\tau} \\ &= \Lambda^{\beta}{}_{\alpha}u^{\alpha} \end{aligned}$$

so the tangent to the path is a Lorentz vector. It is called the 4 -velocity. It is easy to find the components of the 4-velocity in terms of the usual "3-velocity", $\mathbf{v}$:

$$\begin{aligned} u^{\beta} &= \frac{dx^{\beta}}{d\tau} \\ &= \frac{d}{d\tau}(ct, \mathbf{x}) \\ &= \left(c\frac{dt}{d\tau}, \frac{d\mathbf{x}}{d\tau}\right) \\ &= \frac{dt}{d\tau}\left(c, \frac{d\mathbf{x}}{dt}\right) \\ &= \gamma(c, \mathbf{v}) \end{aligned}$$

Since u^{α} is a 4-vector, its length must be something that is independent of the frame of reference of the observer. Let's compute it to check:

$$\begin{aligned} u^{\alpha}u_{\alpha} &= \gamma(c, \mathbf{v}) \cdot \gamma(c, -\mathbf{v}) \\ &= \gamma^2\left(c^2 - \mathbf{v}^2\right) \\ &= \frac{c^2 - \mathbf{v}^2}{1 - \frac{\mathbf{v}^2}{c^2}} \\ &= c^2 \end{aligned}$$

Indeed, all observers agree on this value!

Now let m be the (Lorentz invariant!) mass of a particle. We define the 4 -momentum,

$$p^{\alpha} = mu^{\alpha}$$

Since u^{α} is a Lorentz vector and m is invariant, p^{α} is a Lorentz vector. Once again, the magnitude is invariant, since $p_{\alpha}p^{\alpha} = m^2u_{\alpha}u^{\alpha} = m^2c^2$. Notice that if m is *not Lorentz invariant*, the 4-momentum is *not* a 4-vector. The components of p^{α} are called the (relativistic) energy and the (relativistic) 3-momentum. Setting

$$p^{\alpha} = (E/c, \mathbf{p})$$

we find the familiar formulas,

$$\begin{aligned} E &= \gamma mc^2 \\ \mathbf{p} &= \gamma m\mathbf{v} \end{aligned}$$

Expanding the γ factor when $\mathbf{v}^2 << c^2$,

$$\begin{aligned}\gamma &= \left(1 - \frac{\mathbf{v}^2}{c^2}\right)^{-1/2} \\ &= 1 + \frac{\mathbf{v}^2}{2c^2} + O\left(\frac{\mathbf{v}^4}{c^4}\right)\end{aligned}$$

we recover the non-relativistic expressions

$$\begin{aligned}E &= m\gamma c^2 \approx mc^2 + \frac{1}{2}m\mathbf{v}^2 \\ \mathbf{p} &= m\gamma\mathbf{v} \approx m\mathbf{v}\end{aligned}$$

We will shortly see other objects with linear, homogeneous transformations under the Lorentz group. Some have multiple indices, $T^{\alpha\beta\cdots\mu}$ and transform linearly on each index,

$$(T')^{\alpha\beta\ldots\mu} = \Lambda^{\alpha}{}_{\rho}\Lambda^{\beta}{}_{\sigma}\Lambda^{\mu}{}_{\nu}T^{\rho\sigma\ldots\nu}$$

The collection of all such objects is called the set of Lorentz tensors. More specifically, we are discussing the group of transformations (Exercise: prove that the Lorentz transformations form a group!) that preserves the matrix $diag(1,-1,-1,-1)$. This group is named $O(1,3)$, meaning the pseudo-orthogonal group that preserves the 4-dimensional metric with 1 plus and 3 minus signs. In general the group of transformations preserving $diag(1,\ldots 1,-1,\ldots -1)$ with p plus signs and q plus signs is named $O(p,q)$. From the definition of $\Lambda^{\alpha}{}_{\mu}$ via

$$\eta_{\mu\nu} = \eta_{\alpha\beta}\Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu} \tag{1.29}$$

or, more concisely $\eta = \Lambda^t\eta\Lambda$ we see that $(\det\Lambda)^2 = 1$. If we restrict to $\det\Lambda = +1$, the corresponding group is called $SO(3,1)$, where the S stands for "special".

1.1.5.3 Lorentz invariant tensors

Notice that the defining property of Lorentz transformations, eq.(1.26) or eq.(1.29), states the invariance of the metric $\eta_{\alpha\beta}$ under Lorentz transformations. This is a very special property – in general, the components of tensors are shuffled linearly by Lorentz transformations.

The Levi-Civita tensor, defined to be the unique, totally antisymmetric rank four tensor $\varepsilon_{\alpha\beta\mu\nu}$ with

$$\begin{aligned}\varepsilon_{0123} &= 1 \\ \varepsilon_{\alpha\beta\mu\nu} &= \varepsilon_{[\alpha\beta\mu\nu]}\end{aligned}$$

is the only other independent tensor which is Lorentz invariant. To see that $\varepsilon_{\alpha\beta\mu\nu}$ is invariant, we first note that it may be used to define determinants. For any matrix $M^{\alpha\beta}$, we may write

$$\det M = \varepsilon_{\alpha\beta\mu\nu}M^{\alpha 0}M^{\beta 1}M^{\mu 2}M^{\nu 3}$$

If we let the fixt indices vary, we may write this in another way

$$\begin{aligned}\det M &= \frac{1}{4!}\varepsilon_{\gamma\delta\rho\sigma}\varepsilon_{\alpha\beta\mu\nu}M^{\alpha\gamma}M^{\beta\delta}M^{\mu\rho}M^{\nu\sigma} \\ &= \frac{1}{4!}\varepsilon^{\gamma\delta\rho\sigma}\varepsilon_{\alpha\beta\mu\nu}M^{\alpha}{}_{\gamma}M^{\beta}{}_{\delta}M^{\mu}{}_{\rho}M^{\nu}{}_{\sigma}\end{aligned}$$

because the required antisymmetrizations are accomplished by the Levi-Civita tensor. An alternative way to write this is

$$(\det M)\,\varepsilon_{\gamma\delta\rho\sigma} = \varepsilon_{\alpha\beta\mu\nu}M^{\alpha}{}_{\gamma}M^{\beta}{}_{\delta}M^{\mu}{}_{\rho}M^{\nu}{}_{\sigma}$$

because the right side is totally antisymmetric on $\gamma\delta\rho\sigma$ and if we set $\gamma\delta\rho\sigma = 0123$ we get our original expression for $\det M$. Since this last expression holds for any matrix $M^{\alpha}{}_{\gamma}$, it holds for the Lorentz transformation matrix, $\Lambda^{\alpha}{}_{\gamma}$:

$$(\det \Lambda)\,\varepsilon_{\gamma\delta\rho\sigma} = \varepsilon_{\alpha\beta\mu\nu}\Lambda^{\alpha}{}_{\gamma}\Lambda^{\beta}{}_{\delta}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}$$

However, since the determinant of a (proper) Lorentz transformation is $+1$, we have the invariance of the Levi-Civita tensor,

$$\varepsilon_{\gamma\delta\rho\sigma} = \varepsilon_{\alpha\beta\mu\nu}\Lambda^{\alpha}{}_{\gamma}\Lambda^{\beta}{}_{\delta}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma} \tag{1.30}$$

This also shows that under spatial inversion, which has $\det \Lambda = -1$, the Levi-Civita tensor changes sign. The presence of an odd number of Levi-Civita tensors in any relativistic expression therefore shows that that expression is odd under parity.

In fact, we need only know this parity argument for a single Levi-Civita tensor, because any pair of them may always be replaced by four antisymmetrized Kronecker deltas using

$$\varepsilon^{\alpha\beta\mu\nu}\varepsilon_{\gamma\delta\rho\sigma} = -4!\delta^{\alpha}_{[\gamma}\delta^{\beta}_{\delta}\delta^{\mu}_{\rho}\delta^{\nu}_{\sigma]}$$

where the square brackets around the indices indicate antisymmetrization over all 24 permutations of $\gamma\delta\rho\sigma$, with the normalization $\frac{1}{4!}$. The minus sign occurs because to raise each of the four indices, with all different, will require three signs from the signs in the inverse metric. By taking one, two, three or four contractions we obtain the following identities:

$$\begin{aligned}
\varepsilon^{\alpha\beta\mu\nu}\varepsilon_{\alpha\delta\rho\sigma} &= -6\delta^{\beta}_{[\delta}\delta^{\mu}_{\rho}\delta^{\nu}_{\sigma]} \\
\varepsilon^{\alpha\beta\mu\nu}\varepsilon_{\alpha\beta\rho\sigma} &= -2\left(\delta^{\mu}_{\rho}\delta^{\nu}_{\sigma} - \delta^{\mu}_{\sigma}\delta^{\nu}_{\rho}\right) \\
\varepsilon^{\alpha\beta\mu\nu}\varepsilon_{\alpha\beta\mu\sigma} &= -6\delta^{\nu}_{\sigma} \\
\varepsilon^{\alpha\beta\mu\nu}\varepsilon_{\alpha\beta\mu\nu} &= -24
\end{aligned}$$

Similar identities hold in every dimension. In n dimensions, the Levi-Civita tensor is of rank n. For example, the Levi-Civita tensor of Euclidean 3-space is ε_{ijk}, where

$$\varepsilon_{123} = 1$$

and all other components follow using the antisymmetry. Along with the metric, $\eta_{ij} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$, ε_{ijk} is invariant under $SO(3)$. It is again odd under parity, and satisfies the following identities

$$\begin{aligned}
\varepsilon^{ijk}\varepsilon_{lmn} &= \delta^{i}_{[l}\delta^{j}_{m}\delta^{k}_{n]} \\
\varepsilon^{ijk}\varepsilon_{imn} &= \delta^{j}_{m}\delta^{k}_{n} - \delta^{j}_{n}\delta^{k}_{m} \\
\varepsilon^{ijk}\varepsilon_{ijn} &= 2\delta^{k}_{n} \\
\varepsilon^{ijk}\varepsilon_{ijk} &= 6
\end{aligned}$$

Here there are no minus signs in the metric. These identities will be useful in our discussion of the rotation group.

1.1.5.4 Discrete Lorentz transformations

In addition to rotations and boosts, there are two additional discrete transformations which preserve τ. Normally these are taken to be parity ($\mathcal{P}$) and time reversal ($\mathcal{T}$). Parity is defined as spatial inversion,

$$\mathcal{P} : (t, \mathbf{x}) \to (t, -\mathbf{x}) \tag{1.31}$$

We do not achieve new symmetries by reflecting only two of the spatial coordinates, e.g., $(t, x, y, z) \to (t, -x, -y, z)$ because this effect is achieved by a rotation by π about the z axis. For the same reason, reflection of a single coordinate is equivalent to reflecting all three. The effect of the parity on energy and momentum follows easily. Since the 4-momentum is defined by

$$p^\beta = m\frac{dx^\beta}{d\tau} \tag{1.32}$$

and because m and τ are Lorentz invariant, we have

$$\begin{aligned}
\mathcal{P}(E/c, \mathbf{p}) &= \mathcal{P}\left(m\frac{d(t, \mathbf{x})}{d\tau}\right) \\
&= m\frac{d}{d\tau}\mathcal{P}(t, \mathbf{x}) \\
&= m\frac{d}{d\tau}(t, -\mathbf{x}) \\
&= (E/c, -\mathbf{p})
\end{aligned}$$

Time reversal is chosen to mimic Newtonian time reversal. In the Newtonian case, time reversal is just the replacement $t \to -t$,

$$\mathcal{T}_N : (t, \mathbf{x}) \to (-t, \mathbf{x})$$

Acting on non-relativistic energy and momentum this gives

$$\begin{aligned}
\mathcal{T}_N E &= \mathcal{T}_N\left(\frac{1}{2}m\left(\frac{d\mathbf{x}}{dt}\right)^2\right) = \frac{1}{2}m\left(\frac{d\mathbf{x}}{d(-t)}\right)^2 = E \\
\mathcal{T}_N \mathbf{p} &= \mathcal{T}_N m\left(\frac{d\mathbf{x}}{dt}\right) = m\frac{d\mathbf{x}}{d(-t)} = -\mathbf{p}
\end{aligned}$$

so that Newtonian time reversal is given by

$$\mathcal{T}_N : (E, \mathbf{p}) \to (E, -\mathbf{p})$$

Define: Relativistic time reversal, $\mathcal{T}$, is the discrete Lorentz transformation which reduces in the non-relativistic limit to Newtonian time reversal, $\mathcal{T}_N$.

An useful mnemonic for the effect of time reversal is to imagine filming some motion, then running the movie backward. The backward running film is the time reversed motion. It follows that:

$$\begin{aligned}
\mathcal{T} &: (t, \mathbf{x}) \to (t, \mathbf{x}) \\
\mathcal{T} &: (E, \mathbf{p}) \to (E, -\mathbf{p})
\end{aligned} \tag{1.33}$$

This transformation is a Lorentz transformation, since it preserves the fundamental invariant, $\tau = (x^\alpha x_\alpha)^{1/2}$. However, the definition means that the 4-momentum is not a proper Lorentz vector, since it does not have the same transformation law as the position vector. Correspondingly, we see that the relativistic norm of $x^\alpha + \beta p^\alpha$ is *not* invariant:

$$(x^\alpha + \beta p^\alpha)(x_\alpha + \beta p_\alpha) = \tau^2 + 2\beta(Et - \mathbf{p}\cdot\mathbf{x}) + m^2$$

but

$$(\mathcal{T}x^\alpha + \beta\mathcal{T}p^\alpha)(\mathcal{T}x_\alpha + \beta\mathcal{T}p_\alpha) = \tau^2 + 2\beta(Et + \mathbf{p}\cdot\mathbf{x}) + m^2$$

In this case we might call the 4-momentum a pseudo-vector or a semi-vector. As with polar vectors in classical mechanics, this distinction causes little confusion. However, there is an alternative definition of time reversal which appears better suited to relativistic problems: chronicity.

We define chronicity as follows.

Define: *Chronicity,* $\times$, is the reversal of the Cartesian time component of 4-vectors

$$\times : (t, \mathbf{x}) \rightarrow (-t, \mathbf{x}) \tag{1.34}$$

This is clearly a Lorentz transformation. Now we *compute* the effect of chronicity on energy and momentum from their definitions in terms of the coordinates:

$$\times (E/c, \mathbf{p}) = \times \left(m \frac{d\,(t, \mathbf{x})}{d\tau} \right) = m \frac{d}{d\tau} \times (t, \mathbf{x}) = m \frac{d}{d\tau} (-t, \mathbf{x}) = (-E/c, \mathbf{p})$$

With this definition of the symmetry, the energy-momentum is once again a proper 4-vector, but the non-relativistic limit is exactly opposite to Newtonian time reversal.

Notice the unexpected role played by the invariance of the proper time. By contrast with Newtonian time reversal, with the invariance of τ and the linearity of both E and $\mathbf{p}$ in τ, only the energy reverses sense. The difference is easy to see in a spacetime diagram, where the old "run the movie backward" prescription is seen to require some fine tuning. In spacetime, the "motion" of the particle is replaced by a world line. Under chronicity, this world line flips into the past light cone. An observer (still moving forward in time in either the Newtonian or the relativistic version) experiences this flipped world line in reverse order, so negative energy appears to depart the endpoint and later arrive at the initial point of the motion. A collision at the endpoint, however, imparts momentum in the same direction regardless of the time orientation (see fig.(1)).

In discussing the inevitable negative energy states that arise in field theories, and their relation to antiparticles, chronicity plays a central role.

The subgroup of Lorentz transformations for which the coordinate system remains right handed is called the *proper Lorentz group*, and the subgroup of Lorentz transformations which maintains the orientation of time is called the *orthochronous Lorentz group.* The simply connected subgroup which maintains both the direction of time and the handedness of the spatial coordinates is the *proper orthochronous Lorentz group.*

1.1.6 Noether's Theorem

While now turn to a proof of Noether's Theorem. This theorem establishes the relationship between symmetry and conserved quantities. This important relationship means that the measurable quantities in physics come from symmetries of the action.

By a symmetry we mean any set of transformations of the fields and/or coordinates that leaves the action invariant. Generally we expect symmetries to form a group. We can argue this as follows. Certainly, if we can transform a field from one value to another, we can transform back to the original field, showing that the set of symmetry transformations include inverses. Also, we can always count the identity transformation, which just leaves the fields alone, as an element of the set. And the set of transformations is closed: transforming a field twice, we still have a field, so the composition of two symmetry transformations defines a third symmetry transformation. The only remaining requirement for the set of transformations to be a group is that the transformations be associative. This is a bit harder to argue qualitatively, so we won't. But it turns out to be the case in all of the symmetries we will consider.

To derive the theorem, suppose we have an action built from some fields ϕ^A, where A is any collection of labels or indices. In this way, ϕ^A can represent any number of scalar, vector and/or other types of fields. Let the transformation

$$\phi^A \rightarrow \tilde{\phi}^A = \phi^A + \Delta^A(\phi^B, x)$$

be a transformation that leaves S invariant, $S\left[\phi^A\right] = S\left[\tilde{\phi}^A\right]$. The function Δ is some *specific* function of the coordinates and fields, not a general variation.

To prove Noether's theorem consider first *consider an arbitrary variation of the action*:

$$\begin{aligned}
\delta S &= \frac{\delta}{\delta \phi^A} \int \mathcal{L} \\
\delta S &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} \delta \phi^A + \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \delta \left(\partial_\mu \phi^A \right) \right) \\
&= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} \delta \phi^A + \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \partial_\mu \left(\delta \phi^A \right) \right) \\
&= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} \delta \phi^A + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \delta \phi^A \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \right) \delta \phi^A \right) \\
&= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \right) \right) \delta \phi^A + \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \delta \phi^A \right)
\end{aligned}$$

Next, we *restrict the variation to the symmetry,* $\delta \phi^A \longrightarrow \Delta \phi^A$. Because $\Delta \phi^A$ is a symmetry variation, δS now vanishes, leaving

$$0 = \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \right) \right) \Delta \phi^A + \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \Delta \phi^A \right)$$

Finally, we *impose the field equation,* $\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \right) = 0$, leaving only the divergence term,

$$0 = \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \Delta \phi^A \right)$$

Since we may integrate over any volume, we may shrink the region to any point to find that the integrand vanishes there. This means that the current defined by

$$J^\mu \equiv \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \Delta^A \tag{1.35}$$

must be conserved,

$$\partial_\mu J^\mu = 0$$

The conserved current J^μ is the *Noether current.*

Now that we see how it works, we can generalize the theorem somewhat. The field equations are unchanged by a variation that changes the action by the integral of a divergence, since such a term contributes only on the boundary. We may therefore allow symmetry variations $\delta_\Delta S$ such that

$$\delta_\Delta S = \int \partial_\mu K^\mu$$

for any K^μ built from the fields and coordinates. Then, starting with the general variation as before, restricting to δ_Δ and imposing the field equations,

$$\begin{aligned}
\delta S &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \right) \right) \delta \phi^A + \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \delta \phi^A \right) \\
\delta_\Delta S &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \right) \right) \delta_\Delta \phi^A + \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \delta_\Delta \phi^A \right) \\
\int \partial_\mu K^\mu &= \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \delta_\Delta \phi^A \right)
\end{aligned}$$

we now conclude that the current is

$$J^\mu \equiv \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \Delta^A - K^\mu$$

is conserved,

$$\partial_\mu J^\mu = 0$$

The most important symmetries of particle physics are the Poincaré transformations, consisting of spacetime translations and Lorentz transformations. We develop the general form for these, then consider scalar and vector fields explicitly.

1.1.6.1 Translation invariance

Consider an action of the form

$$S = \int \mathcal{L}(\phi^A, \partial\phi^A) d^4x$$

Since the integral is over all of spacetime, the value of the integral cannot depend on a translation of the coordinates, either in time or space. If a^μ is an arbitrary constant 4-vector then a *translation*,

$$x^\mu \to x^\mu + a^\mu \tag{1.36}$$

leaves S unchanged. The change in the fields for infinitesimal a^μ is

$$\phi^A(x) \to \phi^A(x+a) \approx \phi^A(x) + \frac{\partial\phi^A}{\partial x^\alpha} a^\alpha$$

so that $\delta_\Delta \phi^A = \frac{\partial\phi^A}{\partial x^\alpha} a^\alpha$. Then

$$\begin{aligned}
\delta_\Delta\left(\partial_\mu\phi^A\right) &= \partial_\mu\left(\delta_\Delta\phi^A\right) \\
&= \partial_\mu\left(\frac{\partial\phi^A}{\partial x^\alpha} a^\alpha\right) \\
&= \frac{\partial^2\phi^A}{\partial x^\mu \partial x^\alpha} a^\alpha
\end{aligned}$$

so that the variation of the Lagrangian density is a total divergence,

$$\begin{aligned}
\delta_\Delta \mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi^A}\delta_\Delta\phi^A + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi^A\right)}\delta_\Delta\left(\partial_\mu\phi^A\right) \\
&= \frac{\partial\mathcal{L}}{\partial\phi^A}\frac{\partial\phi^A}{\partial x^\alpha}a^\alpha + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi^A\right)}\frac{\partial^2\phi^A}{\partial x^\mu\partial x^\alpha}a^\alpha \\
&= \frac{\partial\mathcal{L}}{\partial x^\alpha}a^\alpha \\
&= \frac{\partial}{\partial x^\alpha}\left(\mathcal{L}a^\alpha\right)
\end{aligned}$$

In this case, the general variation of the action is a pure surface term, $\delta S = \int \partial_\alpha\left(\mathcal{L}a^\alpha\right)$. Imposing the Euler-Lagrange field equations we are left with

$$\begin{aligned}
J^\mu &= \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi^A\right)}\delta_\Delta\phi^A - K^\mu \\
&= \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi^A\right)}\frac{\partial\phi^A}{\partial x^\alpha}a^\alpha - \mathcal{L}a^\mu \\
&= \left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi^A\right)}\frac{\partial\phi^A}{\partial x^\alpha} - \mathcal{L}\delta^\mu_\alpha\right)a^\alpha
\end{aligned}$$

as the conserved Noether current. Notice that there is a current for each of the four (3 space and 1 time) translations. For each different translation we get a distinct conserved current. Since a^μ is constant, we can extract it from the derivative

$$0 = a^\alpha \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi^A\right)}\frac{\partial\phi^A}{\partial x^\alpha} - \mathcal{L}\delta^\mu_\alpha\right)$$

and since it is arbitrary we can drop it altogether

$$0 = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \frac{\partial \phi^A}{\partial x^\alpha} - \mathcal{L} \delta^\mu_\alpha \right)$$

We therefore define four independent currents

$$T^\mu{}_\alpha \equiv \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \frac{\partial \phi^A}{\partial x^\alpha} - \mathcal{L} \delta^\mu_\alpha$$

Raising an index with the metric,

$$\begin{aligned} T^{\mu\nu} &\equiv \eta^{\nu\alpha} T^\mu{}_\alpha \\ &= \eta^{\nu\alpha} \frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi^A \right)} \frac{\partial \phi^A}{\partial x^\alpha} - \mathcal{L} \eta^{\mu\nu} \end{aligned} \tag{1.37}$$

The conservation law now takes the form

$$\partial_\mu T^{\mu\nu} = 0 \tag{1.38}$$

The 2^{nd} rank tensor (matrix) $T^{\mu\nu}$ is called the energy-momentum tensor (or sometimes the stress-energy tensor). Although our expression here is not necessarily symmetric ($T^{\mu\nu} \neq T^{\nu\mu}$ in general), we can always add a total vanishing divergence to make it symmetric. It is the symmetric version of the stress-energy tensor that provides the source for curvature in general relativity. Therefore, even though many solutions in general relativity use macroscopic versions of $T^{\mu\nu}$ in which the elements correspond to energy density, pressures and stresses, the field theory approach shows that it is really built from fundamental particle fields. Of course, a statistical average of the fundamental fields gives the pressures and stresses in the macroscopic form, but in a truly fundamental theory $T^{\mu\nu}$ is built purely from quantum fields. For example, researchers studying the early universe will drive the cosmological model by introducing a scalar field, the inflaton, to produce an inflationary phase to the overall cosmological expansion.

We construct conserved charges by integrating the time component of each current over a spatial 3-volume Σ

$$P^\alpha = \int_\Sigma T^{0\alpha} d^3x$$

Then using $0 = \partial_\mu T^{\mu\nu} = \partial_0 T^{0\nu} + \partial_i T^{i\nu}$, and the divergence theorem,

$$\begin{aligned} \frac{dP^\beta}{dt} &= \frac{d}{dt} \int_\Sigma T^{0\beta} d^3x \\ &= \int_\Sigma \frac{\partial}{\partial t} T^{0\beta} d^3x \\ &= -\int_\Sigma \partial_i T^{i\beta} d^3x \\ &= -\int_{\delta\Sigma} T^{i\beta} n_i d^2x \end{aligned}$$

where n_i is normal to the 2-dimensional boundary, $\delta\Sigma$, of the 3-volume, Σ. Therefore, the time rate of change of P^β is given by the rate of flow of $T^{i\beta}$ across the boundary.

These charges are the conserved energy and momentum of the field. It is interesting that conservation of momentum arises from invariance of the action under spatial translations while conservation of energy arises from invariance under displacement in time.

1.1.6.2 Lorentz invariance

We are only interested in relativistic field theories, and therefore demand that the actions we consider must be Lorentz invariant. This requirement also leads to conserved charges. For this example, we will assume that we have made the energy-momentum tensor symmetric. This is always possible.

First, we find the form of an infinitesimal Lorentz transformation. The defining property is

$$\eta_{\mu\nu} = \eta_{\alpha\beta}\Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}$$

We let $\Lambda^{\alpha}{}_{\mu}$ be infinitesimally close to the identity

$$\Lambda^{\alpha}{}_{\mu} = \delta^{\alpha}{}_{\mu} + \varepsilon^{\alpha}{}_{\mu}$$

and expand to first order in epsilon:

$$\begin{aligned}
\eta_{\mu\nu} &= \eta_{\alpha\beta}\Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu} \\
&= \eta_{\alpha\beta}\left(\delta^{\alpha}{}_{\mu} + \varepsilon^{\alpha}{}_{\mu}\right)\left(\delta^{\beta}{}_{\nu} + \varepsilon^{\beta}{}_{\nu}\right) \\
&= \eta_{\alpha\beta}\left(\delta^{\alpha}{}_{\mu}\delta^{\beta}{}_{\nu} + \delta^{\alpha}{}_{\mu}\varepsilon^{\beta}{}_{\nu} + \varepsilon^{\alpha}{}_{\mu}\delta^{\beta}{}_{\nu} + \varepsilon^{\alpha}{}_{\mu}\varepsilon^{\beta}{}_{\nu}\right) \\
&\approx \eta_{\mu\nu} + \eta_{\mu\beta}\varepsilon^{\beta}{}_{\nu} + \eta_{\alpha\nu}\varepsilon^{\alpha}{}_{\mu}
\end{aligned}$$

The $\eta_{\mu\nu}$ terms cancel, leaving

$$\begin{aligned}
0 &= \eta_{\mu\beta}\varepsilon^{\beta}{}_{\nu} + \eta_{\alpha\nu}\varepsilon^{\alpha}{}_{\mu} \\
&= \varepsilon_{\mu\nu} + \varepsilon_{\nu\mu}
\end{aligned}$$

which simply says that $\varepsilon_{\mu\nu}$ is antisymmetric. Since an antisymmetric 4×4 matrix has 6 independent components, we see directly the six independent degrees of freedom of the Lorentz transformations.

Now consider the Noether currents. This time, the infinitesimal transformation of the fields depends not only on the change in the coordinates,

$$\begin{aligned}
x^{\beta} &\rightarrow \Lambda^{\beta}{}_{\nu}x^{\nu} = x^{\beta} + \varepsilon^{\beta}{}_{\nu}x^{\nu} \\
\delta x^{\beta} &= \varepsilon^{\beta}{}_{\nu}x^{\nu}
\end{aligned}$$

but also on what type of field we consider. For example, scalar, contravariant vector fields and covariant vector fields change as

$$\begin{aligned}
\phi(x) &\rightarrow \phi(\Lambda x) = \phi(x) + \frac{\partial\phi}{\partial x^{\alpha}}\delta x^{\alpha} \\
v^{\alpha}(x) &\rightarrow \Lambda^{\alpha}{}_{\mu}v^{\mu}(\Lambda x) = \left(\delta^{\alpha}{}_{\mu} + \varepsilon^{\alpha}{}_{\mu}\right)\left(v^{\mu}(x) + \frac{\partial v^{\mu}}{\partial x^{\beta}}\delta x^{\beta}\right) \\
v_{\alpha}(x) &\rightarrow v_{\mu}(\Lambda x)\left(\Lambda^{-1}\right)^{\mu}{}_{\alpha} = \left(v_{\mu}(x) + \frac{\partial v_{\mu}}{\partial x^{\beta}}\delta x^{\beta}\right)\left(\delta^{\mu}{}_{\alpha} - \varepsilon^{\mu}{}_{\alpha}\right)
\end{aligned}$$

Other types of fields have other transformation properties. Notice the use of the inverse Lorentz transformation for covariant vectors. This follows from the Lorentz invariance of $v^{\alpha}v_{\alpha}$. The infinitesimal expression $\delta^{\mu}{}_{\alpha} - \varepsilon^{\mu}{}_{\alpha}$ is easily shown to be the inverse to $\delta^{\alpha}{}_{\mu} + \varepsilon^{\alpha}{}_{\mu}$ to first order in epsilon.

Example 1: Scalar field

Energy-momentum:

$$\begin{aligned}
T^{\mu\nu} &\equiv \eta^{\nu\alpha}T^{\mu}{}_{\alpha} \\
&= \eta^{\nu\alpha}\frac{\partial\mathcal{L}}{\partial\left(\partial_{\mu}\phi\right)}\frac{\partial\phi}{\partial x^{\alpha}} - \mathcal{L}\eta^{\mu\nu}
\end{aligned}$$

Since

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2}\left(\partial^{\mu}\phi\partial_{\mu}\phi - m^{2}\phi^{2}\right) \\
\frac{\partial\mathcal{L}}{\partial\left(\partial_{\mu}\phi\right)} &= \partial^{\mu}\phi
\end{aligned}$$

we have

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \frac{1}{2}\eta^{\mu\nu}\left(\partial^\alpha\phi\partial_\alpha\phi - m^2\phi^2\right) \tag{1.39}$$

and we check explicitly that the divergence vanishes,

$$\begin{aligned}
\partial_\mu T^{\mu\nu} &= \partial_\mu\partial^\mu\phi\partial^\nu\phi + \partial^\mu\phi\partial_\mu\partial^\nu\phi - \frac{1}{2}\partial^\nu\left(\partial^\alpha\phi\partial_\alpha\phi - m^2\phi^2\right) \\
&= \Box\phi\partial^\nu\phi + \partial^\mu\phi\partial_\mu\partial^\nu\phi - \partial^\nu\partial^\alpha\phi\partial_\alpha\phi + m^2\phi\partial^\nu\phi \\
&= \left(\Box\phi + m^2\phi\right)\partial^\nu\phi \\
&= 0
\end{aligned}$$

Angular momentum: Let's find the general form. First note that the variation of the Lagrangian density is a total divergence,

$$\begin{aligned}
\delta_\Delta\mathcal{L} &= \left(\partial^\mu\phi\partial_\mu\partial_\alpha\phi - m^2\phi\partial_\alpha\phi\right)\ \varepsilon^\alpha{}_\nu x^\nu \\
&= \frac{1}{2}\partial_\alpha\left(\partial^\mu\phi\partial_\mu\phi - m^2\phi^2\right)\ \varepsilon^\alpha{}_\nu x^\nu \\
&= \partial_\alpha\mathcal{L}\ \varepsilon^\alpha{}_\nu x^\nu \\
&= \partial_\alpha\left(\mathcal{L}\ \varepsilon^\alpha{}_\nu x^\nu\right)
\end{aligned}$$

Therefore, the variation of the action restricted to the symmetry is a pure surface term, $\delta_\Delta S = \int \partial_\alpha\left(\mathcal{L}a^\alpha\right)$. Now find the general variation of the action and impose the Euler-Lagrange field equations, which gives a surface term as before,

$$\begin{aligned}
\delta S &= \int d^4x\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\delta_\Delta\phi\right) \\
&= \int d^4x\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\partial_\alpha\phi\ \varepsilon^\alpha{}_\nu x^\nu\right)
\end{aligned}$$

This must equal the integral of the variation of $\mathcal{L}$

$$\int d^4x\partial_\mu\left(\mathcal{L}\ \varepsilon^\mu{}_\nu x^\nu\right) = \int d^4x\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\partial_\alpha\phi\ \varepsilon^\alpha{}_\nu x^\nu\right)$$

so we have the vanishing divergence

$$\begin{aligned}
0 &= \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\partial_\alpha\phi\ \varepsilon^\alpha{}_\nu x^\nu - \mathcal{L}\ \varepsilon^\mu{}_\nu x^\nu\right) \\
&= \partial_\mu\left(\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\partial_\alpha\phi - \mathcal{L}\delta^\mu_\alpha\right)\ \varepsilon^\alpha{}_\nu x^\nu\right) \\
&= \partial_\mu\left(T^\mu{}_\alpha\ \varepsilon^\alpha{}_\nu x^\nu\right) \\
&= \partial_\mu\left(\varepsilon_{\alpha\nu}T^{\mu\alpha}x^\nu\right) \\
&= \partial_\mu\left(\frac{1}{2}\varepsilon_{\alpha\nu}\left(T^{\mu\alpha}x^\nu - T^{\mu\nu}x^\alpha\right)\right)
\end{aligned}$$

Therefore, we define

$$M^{\alpha\beta\nu} = T^{\alpha\beta}\ x^\nu - T^{\alpha\nu}\ x^\beta \tag{1.40}$$

and drop the arbitrary antisymmetric constant $\frac{1}{2}\varepsilon_{\alpha\nu}$, it is conserved:

$$\partial_\alpha M^{\alpha\beta\nu} = 0 \tag{1.41}$$

Notice that if the stress-energy tensor is *not* symmetric, $M^{\alpha\beta\nu}$ is *not* conserved, because then we have

$$\begin{aligned}
\partial_\alpha M^{\alpha\beta\nu} &= \partial_\alpha \left(T^{\alpha\beta}\; x^\nu - T^{\alpha\nu}\; x^\beta\right) \\
&= \partial_\alpha T^{\alpha\beta}\; x^\nu - \partial_\alpha T^{\alpha\nu}\; x^\beta + T^{\alpha\beta}\partial_\alpha\; x^\nu - T^{\alpha\nu}\partial_\alpha\; x^\beta \\
&= T^{\alpha\beta}\delta^\nu_\alpha - T^{\alpha\nu}\delta^\beta_\alpha \\
&= T^{\nu\beta} - T^{\beta\nu}
\end{aligned}$$

Therefore, we return to consider what to do when $T^{\alpha\beta}$ is asymmetric.

1.1.6.3 Asymmetric stress-energy vector field (optional)

We will consider the case of a vector field, which may have an antisymmetric stress-energy tensor. For example, let's figure out the stress-energy tensor for the simplest actoin involving a complex vector field:

$$S = \int d^4x \left(\partial^\alpha \bar{v}^\beta \partial_\alpha v_\beta\right)$$

where $\bar{v}^\beta$ is the complex conjugate of v^β. The stress-energy tensor is then

$$\begin{aligned}
T^{\mu\beta} &= \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi^A\right)}\partial^\beta \phi^A - \mathcal{L}\eta^{\mu\beta} \\
&= \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu v^\alpha\right)}\partial^\beta v^\alpha - \mathcal{L}\eta^{\mu\beta} \\
&= \partial^\mu \bar{v}_\alpha \partial^\beta v^\alpha - \mathcal{L}\eta^{\mu\beta}
\end{aligned}$$

The first term can be antisymmetric:

$$T_{\mu\nu} - T_{\nu\mu} = \partial_\mu \bar{v}_\alpha \partial_\nu v^\alpha - \partial_\nu \bar{v}_\alpha \partial_\mu v^\alpha \neq 0$$

It is easy to write down other asymmetric examples.

To handle this case, we will compute the variations in a slightly different way. For vectors (and other rank tensors) there are two ways to look at Lorentz transformations. First, like the scalar field, we have the coordinate dependence,

$$x^\alpha \to \Lambda^\alpha{}_\beta x^\beta$$

which induces a change in $v^\alpha(x)$. Second, since v^α is a Lorentz vector, the vector itself transforms according to

$$v^\alpha \to \Lambda^\alpha{}_\beta v^\beta$$

This transformation law is the definition of a Lorentz vector; similarly, Lorentz tensors are objects with any number of indices, which transform linearly and homogeneously under Lorentz transformations:

$$T^{\alpha\ldots\beta} \to \Lambda^\alpha{}_\mu \ldots \Lambda^\beta{}_\nu T^{\mu\ldots\nu}$$

Since *covariant* tensors (with lowered indices) transform by $\left(\Lambda^{-1}\right)^\alpha{}_\mu$, it is easy to build actions which are invariant under this second form of transformation simply by making sure that every raised index is contracted with a lowered index, and vice versa. For example, we have

$$\begin{aligned}
v^\alpha v_\alpha &\to \Lambda^\alpha{}_\beta v^\beta v_\mu \left(\Lambda^{-1}\right)^\mu{}_\alpha \\
&= \left(\Lambda^{-1}\right)^\mu{}_\alpha \Lambda^\alpha{}_\beta v^\beta v_\mu \\
&= \delta^\mu{}_\beta v^\beta v_\mu \\
&= v^\beta v_\beta
\end{aligned}$$

and the contraction is invariant.

The separate invariance of the theory under transformations of the fields and transformations of the coordinates makes it possible to consider the two types of transformation independently. This simplifies the calculations.

First, consider the transformation of a vector field without a change of coordinates:

$$\begin{aligned} v^{\alpha}(x) &\rightarrow \Lambda^{\alpha}{}_{\beta} v^{\beta}(x) = v^{\alpha} + \varepsilon^{\alpha}{}_{\beta} v^{\beta} \\ \delta v^{\alpha} &= \varepsilon^{\alpha}{}_{\beta} v^{\beta} \end{aligned}$$

Then for derivatives we have

$$\partial_{\mu}\left(\delta v^{\alpha}\right) = \varepsilon^{\alpha}{}_{\beta} \partial_{\mu} v^{\beta} + \left(\partial_{\beta} v^{\alpha}\right) \varepsilon^{\beta}{}_{\mu}$$

The second term arises because the derivative of v^{α} is a second rank tensor, and each index of a tensor must be transformed. Now the variation of v^{α} under a Lorentz transformation is

$$\begin{aligned} \delta_{\Delta}\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial v^{\alpha}} \delta_{\Delta} v^{\alpha} + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} \delta_{\Delta}\left(\partial_{\mu} v^{\alpha}\right) \\ &= \frac{\partial \mathcal{L}}{\partial v^{\alpha}}\left(\varepsilon^{\alpha}{}_{\beta} v^{\beta}\right) + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)}\left(\varepsilon^{\alpha}{}_{\beta} \partial_{\mu} v^{\beta} + \left(\partial_{\beta} v^{\alpha}\right) \varepsilon^{\beta}{}_{\mu}\right) \\ &= \frac{\partial \mathcal{L}}{\partial v^{\alpha}}\left(\varepsilon^{\alpha}{}_{\beta} v^{\beta}\right) + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} \varepsilon^{\alpha}{}_{\beta} \partial_{\mu} v^{\beta} + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)}\left(\partial_{\beta} v^{\alpha}\right) \varepsilon^{\beta}{}_{\mu} \end{aligned}$$

Because we are only considering the active transformation of the fields and not of the coordinates, the Lagrangian density is invariant. So we can simply set $\delta_{\Delta}\mathcal{L} = 0$:

$$\begin{aligned} 0 &= \delta_{\Delta}\mathcal{L} \\ &= \frac{\partial \mathcal{L}}{\partial v^{\alpha}}\left(\varepsilon^{\alpha}{}_{\beta} v^{\beta}\right) + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} \varepsilon^{\alpha}{}_{\beta} \partial_{\mu} v^{\beta} + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)}\left(\partial_{\beta} v^{\alpha}\right) \varepsilon^{\beta}{}_{\mu} \end{aligned}$$

Now we assume a general variation, so we can use the field equations,

$$0 = \frac{\partial \mathcal{L}}{\partial v^{\alpha}} - \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)}\right)$$

We also use the expression for the stress-energy tensor.

$$T^{\mu}{}_{\alpha} = \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\beta}\right)} \partial_{\alpha} v^{\beta} - \mathcal{L}\delta^{\mu}_{\alpha}$$

Then, combining these with the vanishing symmetry variation,

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial v^{\alpha}}\left(\varepsilon^{\alpha}{}_{\beta} v^{\beta}\right) + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} \varepsilon^{\alpha}{}_{\beta} \partial_{\mu} v^{\beta} + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)}\left(\partial_{\beta} v^{\alpha}\right) \varepsilon^{\beta}{}_{\mu} \\ &= \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)}\right)\left(\varepsilon^{\alpha}{}_{\beta} v^{\beta}\right) + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} \varepsilon^{\alpha}{}_{\beta} \partial_{\mu} v^{\beta} + \left(T^{\mu}{}_{\beta} + \mathcal{L}\delta^{\mu}_{\beta}\right) \varepsilon^{\beta}{}_{\mu} \\ &= \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)}\right)\left(\varepsilon^{\alpha}{}_{\beta} v^{\beta}\right) + \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} \varepsilon^{\alpha}{}_{\beta} \partial_{\mu} v^{\beta} + T^{\mu}{}_{\beta} \varepsilon^{\beta}{}_{\mu} \\ &= \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} v^{\beta}\right) \varepsilon^{\alpha}{}_{\beta} + T^{\mu\beta} \varepsilon_{\beta\mu} \\ &= \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v^{\alpha}\right)} v^{\beta}\right) \varepsilon^{\alpha}{}_{\beta} + \frac{1}{2}\left(T^{\mu\beta} - T^{\beta\mu}\right) \varepsilon_{\beta\mu} \end{aligned}$$

where we used $\delta^{\mu}_{\beta}\varepsilon^{\beta}{}_{\mu} = \varepsilon^{\beta}{}_{\beta} = 0$. Notice the explicit appearance of the antisymmetric part of the stress-energy tensor. Extract the arbitrary matrix $\varepsilon_{\beta\mu}$:

$$0 = \frac{1}{2}\left(\partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v_{\alpha}\right)} v^{\beta}\right) - \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu} v_{\beta}\right)} v^{\alpha}\right) - \left(T^{\alpha\beta} - T^{\beta\alpha}\right)\right) \varepsilon_{\alpha\beta}$$

Since the expression contracted with $\varepsilon_{\alpha\beta}$ is now explicitly antisymmetric we can drop the $\varepsilon_{\alpha\beta}$.

$$0 = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu v_\alpha)} v^\beta \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu v_\beta)} v^\alpha \right) - \left(T^{\alpha\beta} - T^{\beta\alpha} \right) \tag{1.42}$$

This is our first result.

Eq.(1.42) gives us the tool we need to construct a new, symmetric form of the stress energy tensor. To see why, suppose we have any tensor $\Sigma^{\mu\alpha\beta}$ which is antisymmetric on the first two indices,

$$\Sigma^{\mu\alpha\beta} = -\Sigma^{\alpha\mu\beta}$$

Then its divergence $\partial_\mu \Sigma^{\mu\alpha\beta}$ is automatically divergence free:

$$\partial_\alpha \partial_\mu \Sigma^{\mu\alpha\beta} = 0$$

This follows because the mixed partials are symmetric on $\mu\alpha$ while sigma is antisymmetric. Therefore,

$$\Theta^{\alpha\beta} = T^{\alpha\beta} + \partial_\mu \Sigma^{\mu\alpha\beta}$$

is conserved as long as $T^{\alpha\beta}$ is. In addition, $\Theta^{\alpha\beta}$ will be symmetric provided

$$\begin{aligned} 0 &= \Theta^{\alpha\beta} - \Theta^{\beta\alpha} \\ &= T^{\alpha\beta} + \partial_\mu \Sigma^{\mu\alpha\beta} - T^{\beta\alpha} - \partial_\mu \Sigma^{\mu\beta\alpha} \end{aligned}$$

Let's find what $\Sigma^{\mu\alpha\beta}$ must be. If we define

$$\lambda^{\mu\alpha\beta} = \frac{\partial \mathcal{L}}{\partial (\partial_\mu v_\alpha)} v^\beta$$

then the condition of Lorentz symmetry, eq.(1.42), may be written more compactly:

$$T^{\alpha\beta} - T^{\beta\alpha} = \partial_\mu \lambda^{\mu\alpha\beta} - \partial_\mu \lambda^{\mu\beta\alpha}$$

Therefore, we have two conditions on $\Sigma^{\mu\beta\alpha}$:

$$\begin{aligned} \partial_\mu \Sigma^{\mu\beta\alpha} - \partial_\mu \Sigma^{\mu\alpha\beta} &= T^{\alpha\beta} - T^{\beta\alpha} \\ &= \partial_\mu \lambda^{\mu\alpha\beta} - \partial_\mu \lambda^{\mu\beta\alpha} \end{aligned}$$

and

$$\Sigma^{\mu\alpha\beta} = -\Sigma^{\alpha\mu\beta}$$

It is sufficient (but not necessary) to drop the divergence on each term of the first equation. Then

$$\begin{aligned} \Sigma^{\mu\beta\alpha} - \Sigma^{\mu\alpha\beta} &= \lambda^{\mu\alpha\beta} - \lambda^{\mu\beta\alpha} \\ \Sigma^{\mu\alpha\beta} &= -\Sigma^{\alpha\mu\beta} \end{aligned}$$

This is not hard to sort out if you know the trick. Write the first equation three times, permuting the indices each time:

$$\begin{aligned} \Sigma^{\mu\beta\alpha} - \Sigma^{\mu\alpha\beta} &= \lambda^{\mu\alpha\beta} - \lambda^{\mu\beta\alpha} \\ \Sigma^{\beta\alpha\mu} - \Sigma^{\beta\mu\alpha} &= \lambda^{\beta\mu\alpha} - \lambda^{\beta\alpha\mu} \\ \Sigma^{\alpha\mu\beta} - \Sigma^{\alpha\beta\mu} &= \lambda^{\alpha\beta\mu} - \lambda^{\alpha\mu\beta} \end{aligned}$$

Each of these is a correct equation, so we can combine them freely. The trick is to add the first two equations and subtract the third. For the left side this gives

$$\begin{aligned} LHS &= \Sigma^{\mu\beta\alpha} - \Sigma^{\mu\alpha\beta} + \Sigma^{\beta\alpha\mu} - \Sigma^{\beta\mu\alpha} - \Sigma^{\alpha\mu\beta} + \Sigma^{\alpha\beta\mu} \\ &= \left(\Sigma^{\mu\beta\alpha} - \Sigma^{\beta\mu\alpha}\right) - \left(\Sigma^{\mu\alpha\beta} + \Sigma^{\alpha\mu\beta}\right) + \left(\Sigma^{\beta\alpha\mu} + \Sigma^{\alpha\beta\mu}\right) \\ &= 2\Sigma^{\mu\beta\alpha} \end{aligned}$$

Where we use our second condition, the antisymmetry of sigma on the first two indices. Since the right hand side is just

$$RHS = \lambda^{\beta\mu\alpha} - \lambda^{\beta\alpha\mu} + \lambda^{\mu\alpha\beta} - \lambda^{\mu\beta\alpha} - \lambda^{\alpha\beta\mu} + \lambda^{\alpha\mu\beta}$$

we have solved for the required form of sigma:

$$\Sigma^{\mu\beta\alpha} = \frac{1}{2}\left(\lambda^{\beta\mu\alpha} - \lambda^{\beta\alpha\mu} + \lambda^{\mu\alpha\beta} - \lambda^{\mu\beta\alpha} - \lambda^{\alpha\beta\mu} + \lambda^{\alpha\mu\beta}\right)$$

Therefore, the symmetric form of the stress energy is (interchanging α and β to get the right form):

$$\begin{aligned} \Theta^{\alpha\beta} &= T^{\alpha\beta} + \partial_\mu \Sigma^{\mu\alpha\beta} \\ &= T^{\alpha\beta} + \frac{1}{2}\partial_\mu\left(\lambda^{\alpha\mu\beta} - \lambda^{\alpha\beta\mu} + \lambda^{\mu\beta\alpha} - \lambda^{\mu\alpha\beta} - \lambda^{\beta\alpha\mu} + \lambda^{\beta\mu\alpha}\right) \end{aligned}$$

where

$$\lambda^{\mu\alpha\beta} = \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu v_\alpha\right)} v^\beta$$

This object is called the Belinfante tensor (see Weinberg, vol I, p. 316; ref to Belinfante, Physica 6, 887 (1939)).

If we substitute this expression for $T^{\alpha\beta}$ in the equation for Lorentz invariance we now get zero automatically:

$$\begin{aligned} \partial_\mu \lambda^{\mu\alpha\beta} - \partial_\mu \lambda^{\mu\beta\alpha} - \left(T^{\alpha\beta} - T^{\beta\alpha}\right) &= \partial_\mu \lambda^{\mu\alpha\beta} - \partial_\mu \lambda^{\mu\beta\alpha} + \partial_\mu \Sigma^{\mu\alpha\beta} - \partial_\mu \Sigma^{\mu\beta\alpha} \\ &= 0 \end{aligned}$$

What has happened to the conservation law? We replaced $T^{\alpha\beta}$ by $\Theta^{\alpha\beta}$ and we still have $\partial_\alpha \Theta^{\alpha\beta} = 0$ for translation invariance, but what about Lorentz invariance? The answer lies in the remaining part of the calculation, namely, the coordinate transformations. We considered only the invariance of the Lagrangian density under Lorentz transformations of the *fields*, but not under transformations of the *coordinates*. We can demand both. Therefore, we now consider what happens when we let

$$\delta x^\beta = \varepsilon^\beta{}_\nu x^\nu$$

as we did for the scalar field.

The symmetry variation of the Lagrangian for this case is simply

$$\delta_\Delta \mathcal{L} = \frac{\partial\mathcal{L}}{\partial x^\alpha}\delta_\Delta x^\alpha = \frac{\partial\mathcal{L}}{\partial x^\alpha}\varepsilon^\alpha{}_\beta x^\beta$$

Now, quite generally, Lagrangian densities depend directly on the coordinates *only* in the volume density, and it is not hard to show that the volume density is Lorentz invariant. Any other dependence is through the fields using the chain rule

$$\begin{aligned} \delta_\Delta \mathcal{L} &\sim \frac{\partial\mathcal{L}}{\partial v^\alpha}\frac{\partial v^\alpha}{\partial x^\mu}\delta x^\mu + \frac{\partial\mathcal{L}}{\partial\left(\partial_\beta v^\alpha\right)}\partial_\mu\partial_\beta v^\alpha \delta x^\mu \\ &= \frac{\partial\mathcal{L}}{\partial v^\alpha}\delta v^\alpha + \frac{\partial\mathcal{L}}{\partial\left(\partial_\beta v^\alpha\right)}\delta\left(\partial_\beta v^\alpha\right) \end{aligned}$$

and these have already been set to zero. Therefore, $\frac{\partial \ \mathcal{L}}{\partial x^\alpha} = 0$ and the variation gives zero, $\delta_\Delta \mathcal{L} = 0$. Therefore

$$(\partial_\mu \mathcal{L}) \varepsilon^\mu{}_\beta x^\beta = 0$$

We expand this expression and use the field equations

$$\begin{aligned}
0 &= \partial_\mu \mathcal{L} \varepsilon^\mu{}_\beta x^\beta \\
&= \left(\frac{\partial \mathcal{L}}{\partial v^\alpha} \partial_\mu v^\alpha + \frac{\partial \mathcal{L}}{\partial \left(\partial_\beta v^\alpha\right)} \partial_\mu \partial_\beta v^\alpha \right) \varepsilon^\mu{}_\nu x^\nu \\
&= \left(\partial_\beta \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\beta v^\alpha\right)} \right) \partial_\mu v^\alpha + \frac{\partial \mathcal{L}}{\partial \left(\partial_\beta v^\alpha\right)} \partial_\mu \partial_\beta v^\alpha \right) \varepsilon^\mu{}_\nu x^\nu \\
&= \left(\partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\alpha v^\nu\right)} \partial^\beta v^\nu \right) \right) \varepsilon_{\beta\rho} x^\rho
\end{aligned}$$

Next, let's use the definition of the *symmetric* stress-energy tensor,

$$\begin{aligned}
\Theta^{\alpha\beta} &= T^{\alpha\beta} + \partial_\mu \Sigma^{\mu\alpha\beta} \\
&= \frac{\partial \mathcal{L}}{\partial \left(\partial_\alpha v^\nu\right)} \partial^\beta v^\nu - \mathcal{L} \eta^{\alpha\beta} + \partial_\mu \Sigma^{\mu\alpha\beta}
\end{aligned}$$

or

$$\frac{\partial \mathcal{L}}{\partial \left(\partial_\alpha v^\nu\right)} \partial^\beta v^\nu = \Theta^{\alpha\beta} + \mathcal{L} \eta^{\alpha\beta} - \partial_\mu \Sigma^{\mu\alpha\beta}$$

to replace this term. Substituting, we find

$$\begin{aligned}
0 &= \left(\partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\alpha v^\nu\right)} \partial^\beta v^\nu \right) \right) \varepsilon_{\beta\rho} x^\rho \\
&= \left(\partial_\alpha \left(\Theta^{\alpha\beta} + \mathcal{L} \eta^{\alpha\beta} - \partial_\mu \Sigma^{\mu\alpha\beta} \right) \right) \varepsilon_{\beta\rho} x^\rho \\
&= \left(\partial_\alpha \Theta^{\alpha\beta} + \partial_\alpha \mathcal{L} \eta^{\alpha\beta} - \partial_\alpha \partial_\mu \Sigma^{\mu\alpha\beta} \right) \varepsilon_{\beta\rho} x^\rho \\
&= \left(\partial_\alpha \Theta^{\alpha\beta} + \partial_\alpha \mathcal{L} \eta^{\alpha\beta} \right) \varepsilon_{\beta\rho} x^\rho \\
&= \partial_\alpha \Theta^{\alpha\beta} \varepsilon_{\beta\rho} x^\rho + \partial_\alpha \left(\mathcal{L} \varepsilon^\alpha{}_\rho x^\rho \right)
\end{aligned}$$

But the second term vanishes:

$$\begin{aligned}
\partial_\alpha \left(\mathcal{L} \varepsilon^\alpha{}_\rho x^\rho \right) &= (\partial_\alpha \mathcal{L}) \varepsilon^\alpha{}_\beta x^\beta + \mathcal{L} \varepsilon^\alpha{}_\beta \frac{\partial x^\beta}{\partial x^\alpha} \\
&= \mathcal{L} \varepsilon^\beta{}_\beta = 0
\end{aligned}$$

so we are left with

$$\begin{aligned}
0 &= \left(\partial_\alpha \Theta^{\alpha\beta} \right) \varepsilon_{\beta\rho} x^\rho \\
&= \partial_\alpha \left(\Theta^{\alpha\beta} \varepsilon_{\beta\rho} x^\rho \right) - \left(\Theta^{\alpha\beta} \varepsilon_{\beta\rho} \partial_\alpha x^\rho \right) \\
&= \partial_\alpha \left(\Theta^{\alpha\beta} \varepsilon_{\beta\rho} x^\rho \right) - \Theta^{\alpha\beta} \varepsilon_{\beta\alpha}
\end{aligned}$$

Now, since $\Theta^{\alpha\beta}$ is symmetric by construction, the second term is zero, leaving

$$\begin{aligned}
0 &= \partial_\alpha \left(\Theta^{\alpha\beta} \varepsilon_{\beta\rho} x^\rho \right) \\
&= \varepsilon_{\beta\rho} \partial_\alpha \left(\Theta^{\alpha\beta} x^\rho \right) \\
&= \frac{1}{2} \varepsilon_{\beta\rho} \partial_\alpha \left(\Theta^{\alpha\beta} x^\rho - \Theta^{\alpha\rho} x^\beta \right)
\end{aligned}$$

and therefore arrive at our conservation law:

$$\begin{aligned} M^{\alpha\beta\rho} &= \Theta^{\alpha\beta}x^{\rho} - \Theta^{\alpha\rho}x^{\beta} \\ \partial_{\alpha}M^{\alpha\beta\rho} &= 0 \end{aligned}$$

Finally, consider the possible conserved currents. If we integrate $M^{0\alpha\beta}$ as usual, we get

$$\begin{aligned} J^{\alpha\beta} &= \int M^{0\alpha\beta}d^3x \\ &= \int \left(\Theta^{0\alpha}x^{\beta} - \Theta^{0\beta}x^{\alpha}\right) d^3x \end{aligned}$$

There are six independent components here, since $M^{0\alpha\beta}$, and therefore $J^{\alpha\beta}$, is antisymmetric under interchange of α and β. These correspond to the three rotations and three boosts of the Lorentz transformations. The rotations are the spatial components, $(i, j = 1, 2, 3)$,

$$\begin{aligned} J^{ij} &= \int M^{0ij}d^3x \\ &= \int \left(\Theta^{0i}x^{j} - \Theta^{0j}x^{i}\right) d^3x \end{aligned}$$

Notice that these do not depend explicitly on the time coordinate and that the components Θ^{0i} of the stress-energy generate momentum. The expression is much like the usual $\mathbf{r} \times \mathbf{p}$ form of angular momentum. The remaining independent charges are

$$\begin{aligned} J^{0i} &= \int M^{00i}d^3x \\ &= \int \left(\Theta^{00}x^{i} - \Theta^{0i}x^{0}\right) d^3x \end{aligned}$$

These depend on energy and time, and generate boosts.

Chapter 2

Group theory

2.1 Groups

Nearly all of the central symmetries of modern physics are group symmetries, for simple a reason. If we imagine a transformation of our fields or coordinates, we can look at linear versions of those transformations. Such linear transformations may be represented by matrices, and therefore (as we shall see) even finite transformations may be given a matrix representation. But matrix multiplication has an important property: associativity. We get a group if we couple this property with three further simple observations: (1) we expect two transformations to combine in such a way as to give another allowed transformation, (2) the identity may always be regarded as a null transformation, and (3) any transformation that we can do we can also undo. These four properties (associativity, closure, identity, and inverses) are the defining properties of a *group*.

We begin with a very basic sketch of groups and Lie groups. The remainder of the chapter, of great importance in field theory, is the development of *spinors*.

Define: A *group* is a pair $G = \{S, \star\}$ where S is a set and $\star$ is an operation mapping pairs of elements in S to elements in S (i.e., $\star : S \star S \to S$. This implies closure) and satisfying the following conditions:

1. Existence of an identity: $\exists\ e \in S$ such that $e \star a = a \star e = a$, $\forall a \in S$.
2. Existence of inverses: $\forall\ a \in S$, $\exists\ a^{-1} \in S$ such that $a \star a^{-1} = a^{-1} \star a = e$.
3. Associativity: $\forall\ a, b, c \in S$, $a \star (b \star c) = (a \star b) \star c = a \star b \star c$

We consider several examples of groups.

1. The simplest group is the familiar boolean one with two elements $S = \{0, 1\}$ where the operation $\star$ is addition modulo two. Then the "multiplication" table is simply

$[+_2]$	0	1
0	0	1
1	1	0

The element 0 is the identity, and each element is its own inverse. This is, in fact, the *only* two element group, for suppose we pick any set with two elements, $S = \{a, b\}$. The multiplication table is of the form

$\star$	a	b
a		
b		

One of these must be the identity; without loss of generality we choose $a = e$. Then

$\star$	a	b
a	a	b
b	b	

Finally, since b must have an inverse, and its inverse cannot be a(for then $ab = a$ implies $a(ab) = (aa)b = b$ so that $a = b$) we must fill in the final spot with the identity, thereby making b its own inverse:

$\star$	a	b
a	a	b
b	b	a

Comparing to the boolean table, we see that a simple renaming, $a \to 0, b \to 1$ reproduces the boolean group. Such a one-to-one mapping between groups that preserves the group product is called an isomorphism.

2. Let $G = \{Z, +\}$, the integers under addition. For all integers a, b, c we have $a + b \in R$ (closure); $0 + a = a + 0 = a$ (identity); $a + (-a) = 0$ (inverse); $a + (b + c) = (a + b) + c$ (associativity). Therefore, G is a group. The integers also form a group under addition mod p, where p is any integer (Recall that $a = b \, mod \, p$ if there exists an integer n such that $a = b + np$).

3. Let $G = \{R, +\}$, the real numbers under addition. For all real numbers a, b, c we have $a + b \in R$ (closure); $0 + a = a + 0 = a$ (identity); $a + (-a) = 0$ (inverse); $a + (b + c) = (a + b) + c$ (associativity). Therefore, G is a group. Notice that the rationals, Q, do not form a group under addition because they do not close under addition:

$$\pi = 3 + .1 + .04 + .001 + .0005 + .00009 + \ldots$$

Exercise: Find all groups (up to isomorphism) with three elements. Find all groups (up to isomorphism) with four elements.

2.2 Lie groups

Of course, the real numbers form a much nicer object than a group. The form a complete Archimedean field. But for our purposes, they form one of the easiest examples of yet another object: a Lie group. The following definition is sufficient for our purposes.

Define: A *Lie group* is a group which is also a manifold. Essentially, this means that a Lie group is a group in which the elements can be labeled by a finite set of continuous labels. Qualitatively, a manifold is a space that is smooth enough that if we look at any sufficiently small region, it looks just like a small region of R^n; the dimension n is fixed over the entire manifold. We will not go into the details of manifolds here, but instead will look at enough examples to get across the general idea.

Define: A *representation* of a Lie group is a space on which the group acts. We are particularly interested in *linear representations*, for which this space is a vector space. A linear representation, in order to map vectors to vectors, may therefore be written as a matrix.

The real numbers form a Lie group because each element of R provides its own label! Since only one label is required, R is a 1-dimensional Lie group. The way to think of R as a manifold is to picture the real line. Some examples:

1. The vector space R^n under vector addition is an n-dim Lie group, since each element of the group may be labeled by n real numbers.

2. Let's move to something more interesting. The set of non-degenerate linear transformations of a real, n-dimensional vector space forms a Lie group. This one is important enough to have its own name: $GL(n;R)$, or more simply, $GL(n)$ where the field (usually $\mathbb{R}$ or $\mathbb{C}$) is unambiguous. The GL stands for General Linear. The transformations may be represented by $n\times n$ matrices with nonzero determinant. Since for any $A\in GL(n;R)$ we have $\det A\neq 0$, the matrix A is invertible. The identity is the identity matrix, and it is not too hard to prove that matrix multiplication is always associative. Since each A can be written in terms of n^2 real numbers, $GL(n)$ has dimension n^2. $GL(n)$ is an example of a Lie group with more than one *connected component*. We can imagine starting with the identity element and smoothly varying the parameters that define the group elements, thereby sweeping out curves in the space of all group elements. If such continuous variation can take us to every group element, we say the group is *connected*. If there remain elements that cannot be connected to the identity by such a continuous variation (actually a curve in the group manifold), then the group has more than one *component*. $GL(n)$ is of this form because as we vary the parameters to move from element to element of the group, the determinant of those elements also varies smoothly. But since the determinant of the identity is 1 and no element can have determinant zero, we can never get to an element that has negative determinant. Any elements of $GL(n)$ with negative determinant are related to those of positive determinant by a discrete transformation: if we pick any element of $GL(n)$ with negative determinant, and multiply it by each element of $GL(n)$ with positive determinant, we get a new element of negative determinant. This shows that the two components of $GL(n)$ are in 1 to 1 correspondence. In odd dimensions, a suitable 1 to 1 mapping is given by $-\mathbf{1}$, which is called the *parity* transformation.

3. We will be concerned with Lie groups that have *linear representations*. This means that each group element may be written as a matrix and the group multiplication is correctly given by the usual form of matrix multiplication. Since $GL(n)$ is the set of *all* linear, invertible transformations in n-dimensions, all Lie groups with linear representations must be subgroups of $GL(n)$. We now look at two principled ways of constructing such subgroups. The simplest subgroup of $GL(n)$ removes the second component to give a connected Lie group. In fact, it is useful to factor out the determinant entirely, because the operation of multiplying by a constant commutes with every other transformation of the group. In this way, we arrive at a *simple group*, one in which each transformation has nontrivial effect on some other transformations. For a general matrix $A\in GL(n)$ with positive determinant, let

$$A=(\det A)^{\frac{1}{n}}\ \hat{A}$$

Then $\det\hat{A}=1$. Since

$$\det\left(\hat{A}\hat{B}\right)=\det\hat{A}\det\hat{B}=1$$

the set of all $\hat{A}$ closes under matrix multiplication. We also have $\det\hat{A}^{-1}=1$, and $\det 1=1$, so the set of all matrices with unit determinant, $\left\{\hat{A}\middle|\det\hat{A}=1\right\}$ forms a Lie group. This group is called the Special Linear group, $SL(n)$.

Frequently, the most useful way to characterize a group is by a set of objects that group transformations leave invariant. In this way, we produce the orthogonal, unitary and symplectic groups:

Theorem: Consider the subset of $GL(n;R)$ that leaves a fixed matrix, M, invariant under a similarity transformation:

$$H=\left\{A|A\in GL(n),AMA^t=M\right\}$$

Then H is also a Lie group.

Proof: First, H is closed, since if both

$$\begin{aligned}AMA^t&=M\\BMB^t&=M\end{aligned}$$

then the product AB is also in H because

$$\begin{aligned} (AB)\,M(AB)^t &= (AB)\,M(B^tA^t) \\ &= A\left(BMB^t\right)A^t \\ &= AMA^t \\ &= M \end{aligned}$$

The identity is present because $IMI^t = M$ and if any A leaves M invariant then so does A^{-1}. To see this, notice that $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$ because the transpose of

$$(A)^{-1}A = I$$

is

$$A^t\left((A)^{-1}\right)^t = I$$

Since it is easy to show (exercise!) that inverses are unique, this shows that $\left((A)^{-1}\right)^t$ must be the inverse of A^t. Using this, we start with

$$AMA^t = M$$

and multiply on the left by A^{-1} and on the right by $\left(A^t\right)^{-1}$:

$$\begin{aligned} A^{-1}AMA^t\left(A^t\right)^{-1} &= A^{-1}M\left(A^t\right)^{-1} \\ M &= A^{-1}M\left(A^t\right)^{-1} \\ M &= A^{-1}M\left(A^{-1}\right)^t \end{aligned}$$

The last line is the statement that A^{-1} leaves M invariant, and is therefore in H. Finally, we still have the associative matrix product, so H is a group, concluding our proof.

Now, fix a (nondegenerate) matrix M and consider the group that leaves M invariant. Suppose M is asymmetrical, so it has both symmetric and antisymmetric parts:

$$\begin{aligned} M &= \frac{1}{2}\left(M + M^t\right) + \frac{1}{2}\left(M - M^t\right) \\ &\equiv M_s + M_a \end{aligned}$$

Then, for any A in H,

$$AMA^t = M$$

implies

$$A\left(M_s + M_a\right)A^t = \left(M_s + M_a\right) \tag{2.1}$$

The transpose of this equation must also hold,

$$A\left(M_s - M_a\right)A^t = \left(M_s - M_a\right) \tag{2.2}$$

so adding and subtracting eqs.(2.1) and (2.2) gives two independent constraints on A :

$$\begin{aligned} AM_sA^t &= M_s \\ AM_aA^t &= M_a \end{aligned}$$

Imposing both of these together is a stronger constraint than each separately, so the largest subgroups H_s and H_a of G that we can form in this way are found by demanding that M be either symmetric or antisymmetric.

If M is symmetric, then we can always choose a basis for the vector space on which the transformations act (the representation) such that M is diagonal; indeed we can go further, for rescaling the basis we can make every diagonal element into +1 or −1. Therefore, any non-degenerate, symmetric M may be put in the form

$$\eta_{ij} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix} \tag{2.3}$$

where there are p terms +1 and q terms −1. We can use η as a pseudo-metric; in components, for any vector v^i,

$$\langle v, v\rangle = \eta_{ij} v^i v^j = \sum_{i=1}^{p} \left(v^i\right)^2 - \sum_{i=p+1}^{p+q} \left(v^i\right)^2$$

Notice that this includes the $O(1,3)$ Lorentz metric of the previous section, as well as the $O(3)$ case of Euclidean 3-space. In general, the subgroup of $GL(n)$ leaving $M_{p,q}$ invariant is termed $O(p,q)$, the pseudo-orthogonal group in $n = p+q$ dimensions. The *signature* of η is $s = p - q$ or simply (p,q).

Now suppose M is antisymmetric. This case arises in classical Hamiltonian dynamics, where we have canonically conjugate variables satisfying fundamental Poisson bracket relations.

$$\begin{aligned} \{q_i, q_j\}_{x\pi} &= \{p_i, p_j\}_{x\pi} = 0 \\ \{p_i, q_j\}_{x\pi} &= -\{q_i, p_j\}_{x\pi} = \delta_{ij} \end{aligned}$$

If we define a single set of coordinates including both p_i and q_i,

$$\xi^a = (q^i, p_j)$$

where if $i, j = 1, 2, \ldots, n$ then $a = 1, 2, \ldots, 2n$, then the fundamental brackets may be written in terms of an antisymmetric matrix Ω^{ab} as

$$\{\xi^a, \xi^b\} = \Omega^{ab}$$

where

$$\Omega^{ab} = \begin{pmatrix} 0 & -\delta^{ij} \\ \delta^{ij} & 0 \end{pmatrix} = -\Omega^{ba} \tag{2.4}$$

Since canonical transformations are precisely the coordinate transformations that preserve the fundamental brackets, we can define a group of symplectic transformations which preserve Ω^{ab}. Then canonical transformations are local symplectic transformations. In general, the subgroup of $GL(n)$ preserving an antisymmetric matrix is called the *symplectic group*. We have a similar result here as for the (pseudo-) orthogonal groups – we can always choose a basis for the vector space that puts the invariant matrix Ω^{ab} in the form given in eq.(2.4). From the form of eq.(2.4) we suspect, correctly, that the symplectic group is always even dimensional (the determinant of an antisymmetric matrix in odd dimensions is always zero, so we cannot have a non-degenerate, odd-dimensioned, antisymmetric matrix). The notation for the symplectic groups is therefore $Sp(2n)$.

For either the orthogonal or symplectic groups, we can consider the unit determinant subgroups. Especially important are the resulting *Special Orthogonal* groups, $SO(p,q)$.

We give one particular example that will be useful to illustrate Lie algebras in the next section. The very simplest case of an orthogonal group is $O(2)$, leaving

$$\eta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

invariant. Equivalently, $O(2)$ leaves the Euclidean norm

$$\langle \mathbf{x}, \mathbf{x}\rangle = M_{ij} x^i x^j = x^2 + y^2$$

invariant. The form of $O(2)$ transformations is the familiar set of rotation matrices,

$$A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and we see that every group element is labeled by a continuous parameter θ lying in the range $\theta \in [0, 2\pi)$. The *group manifold* is the set of all of the group elements regarded as a geometric object. From the range of θ we see that there is one group element for every point on a circle – the group manifold of $O(2)$ is the circle. Note the inverse of $A(\theta)$ is just $A(-\theta)$ and the identity is $A(0)$. All of the transformations of $O(2)$ already have unit determinant, so that $SO(2)$ and $O(2)$ are isomorphic.

2.3 Lie algebras

If we want to work with more complicated Lie groups, working directly with the transformation matrices becomes prohibitively difficult. Instead, most of the information we need to know about the group is already present in the infinitesimal transformations. Unlike the group multiplication, the combination of the infinitesimal transformations is usually fairly simple. This is why, in the previous section, we worked with infinitesimal Lorentz transformations. Here we start with the simpler case of $O(2)$ to develop some of the ideas further.

2.3.1 The simplest example: O(2)

Consider those transformations of $O(2)$ that are close to the identity. Since the identity is $A(0)$, these will be the transformations $A(\varepsilon)$ with $\varepsilon << 1$. Expanding in a Taylor series, we keep only terms to first order:

$$\begin{aligned} A(\varepsilon) &= \begin{pmatrix} \cos\varepsilon & -\sin\varepsilon \\ \sin\varepsilon & \cos\varepsilon \end{pmatrix} \approx \begin{pmatrix} 1 & -\varepsilon \\ \varepsilon & 1 \end{pmatrix} \\ &= \mathbf{1} + \varepsilon \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

The only information here besides the identity is the matrix

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

but remarkably, this is enough to recover the whole group! The matrix G is the single *generator* of $O(2)$. For general Lie groups, we get one generator for each continuous parameter labeling the group elements. The set of all linear combinations of these generators is a vector space called the *Lie algebra* of the group, so that the Lie algebra of $O(2)$ is $\{\theta G |\, \theta \in [0, 2\pi).\}$. We give the full defining set of properties of a Lie algebra below.

Imagine iterating this infinitesimal group element many times. Applying $A(\varepsilon)$ n times rotates the plane by an angle $n\varepsilon$:

$$A(n\varepsilon) = (A(\varepsilon))^n = \left(\mathbf{1} + \varepsilon \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right)^n$$

Expanding the power on the right using the binomial expansion,

$$A(n\varepsilon) \approx \sum_{k=0}^{n} \binom{n}{k} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \varepsilon^k \mathbf{1}^{n-k}$$

To make the equality rigorous, we must take the limit as $\varepsilon \to 0$ and $n \to \infty$, holding the product $n\varepsilon = \theta$

finite. Then:

$$\begin{aligned}
A(\theta) &= \lim_{\varepsilon\to 0, n\varepsilon\to\theta} \sum_{k=0}^{n} \binom{n}{k} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \varepsilon^k \\
&= \lim_{\varepsilon\to 0} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \varepsilon^k \\
&= \lim_{\varepsilon\to 0} \sum_{k=0}^{n} \frac{n\,(n-1)\cdots(n-k+1)}{k!} \varepsilon^k \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \\
&= \lim_{\varepsilon\to 0} \sum_{k=0}^{n} \frac{1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{k!} (n\varepsilon)^k \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}\theta^k \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \\
&\equiv \exp\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\theta\right)
\end{aligned}$$

where in the last step we define the exponential of a matrix to be the power series in the second to last line. Quite generally, since we know how to take powers of matrices, we can define the exponential of any matrix, N, by its power series:

$$\exp N \equiv \sum_{k=0}^{\infty} \frac{1}{k!} N^k$$

Thus, every element $A(\theta) \in O(2)$ is the exponential of an element of the Lie algebra,

$$A(\theta) = e^{\theta G} \tag{2.5}$$

This turns out to be a general property of Lie groups and Lie algebras, though the relationship is not quite one to one.

Next, we check that the exponential form of $A(\theta)$ actually is the original class of transformations. To do this we first examine powers of $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$:

$$\begin{aligned}
\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 &= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathbf{1} \\
\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^3 &= -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^3 &= \mathbf{1}
\end{aligned}$$

The even terms are plus or minus the identity, while the odd terms are always proportional to the generator, $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Therefore, we divide the power series into even and odd parts, and remove the matrices from the sums:

$$\begin{aligned}
A(\theta) &= \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \theta^k \\
&= \sum_{m=0}^{\infty} \frac{1}{(2m)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{2m} \theta^{2m} + \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{2m+1} \theta^{2m+1} \\
&= \mathbf{1} \left(\sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} \theta^{2m} \right) + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} \theta^{2m+1} \\
&= \mathbf{1} \cos\theta + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sin\theta \\
&= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}
\end{aligned}$$

The generator has given us the whole group back.

2.3.2 A non-abelian example: SO(3)

To begin to see the power of this technique, let's look at $SO(3)$, the subgroup of elements of $O(3)$ with unit determinant.

The generators of $SO(3)$ may be found from the property of leaving the matrix (Euclidean metric),

$$g_{ij} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$$

invariant:

$$g_{ij} A^i{}_m A^j{}_n = g_{mn}$$

Just as in the Lorentz case in the previous chapter, this is equivalent to preserving the proper length of vectors. Thus, the transformation

$$y^i = A^i{}_m x^m$$

is a rotation if it preserves the length-squared:

$$g_{ij} y^i y^j = g_{ij} x^i x^j$$

Substituting, we get

$$\begin{aligned}
g_{mn} x^m x^n &= g_{ij} \left(A^i{}_m x^m \right) \left(A^j{}_n x^n \right) \\
&= \left(g_{ij} A^i{}_m A^j{}_n \right) x^m x^n
\end{aligned}$$

Since x^m is arbitrary, we can turn this into a relation between the transformations and the metric, g_{mn}, but we have to be careful with the symmetry since $x^m x^n = x^n x^m$. It isn't a problem here because both sets of coefficients are also symmetric:

$$\begin{aligned}
g_{mn} &= g_{nm} \\
g_{ij} A^i{}_m A^j{}_n &= g_{ji} A^j{}_m A^i{}_n \\
&= g_{ji} A^i{}_n A^j{}_m \\
&= g_{ij} A^i{}_n A^j{}_m
\end{aligned}$$

Therefore, we can strip off the xs and write

$$g_{mn} = g_{ij} A^i{}_m A^j{}_n$$

This is the most convenient form of the definition of the group to use in finding the Lie algebra. For future reference, we note that the inverse to g_{ij} is written as g^{ij}; it is also the identity matrix.

If we write the defining relationship as normal matrix multiplication by setting $A^i{}_m = [A^t]_m{}^i$,

$$g_{mn} = \left[A^t\right]_m{}^i g_{ij} A^j{}_n$$

we may treat the objects as matrices, $g = A^t g A$. Furthermore, since in this case g_{mn} is the identity matrix, we may simply write

$$A^t A = 1$$

Now we have

$$\begin{aligned} 1 &= \det(1) \\ &= \det\left(A^t\right)\det(A) \\ &= (\det(A))^2 \end{aligned}$$

so either $\det A = 1$ or $\det A = -1$. Defining the parity transformation to be the single operator

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

then every element O of $O(3)$ is of the form $O = (\det O)^{\frac{1}{n}} A$ or $(\det O)^{\frac{1}{n}} PA$, where A is in $SO(3)$.

As in the 2-dimensional case, we look at transformations close to the identity. Let

$$A^i{}_j = \delta^i_j + \varepsilon^i{}_j$$

where all components of $\varepsilon^i{}_m$ are small. Then

$$\begin{aligned} g_{mn} &= g_{ij}\left(\delta^i_m + \varepsilon^i{}_m\right)\left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= \left(g_{ij}\delta^i_m + g_{ij}\varepsilon^i{}_m\right)\left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= \left(g_{mj} + g_{ji}\varepsilon^i{}_m\right)\left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= \left(g_{mj} + \varepsilon_{jm}\right)\left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= g_{mj}\delta^j_n + \varepsilon_{jm}\delta^j_n + g_{mj}\varepsilon^j{}_n + \varepsilon_{jm}\varepsilon^j{}_n \end{aligned}$$

where we use the metric to lower an index on $\varepsilon^i{}_m$. Cancelling g_{mn} each side and dropping the term quadratic in ε, we are left with

$$0 = \varepsilon_{nm} + \varepsilon_{mn} \tag{2.6}$$

This shows that the the generators ε_{mn} must be antisymmetric:

$$\varepsilon_{nm} = -\varepsilon_{mn} \tag{2.7}$$

We are dealing with 3×3 matrices here, but note the power of index notation! There is actually nothing in the preceeding calculation that is specific to $n = 3$, and we could draw all the same conclusions up to this point for $SO(p, q)$). For the 3×3 case, every antisymmetric matrix is of the form Next, we write the most general antisymmetric 3×3 matrix as a linear combination of a convenient basis,

$$\begin{aligned} \varepsilon &= w^i J_i \\ &= w^1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + w^2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + w^3 \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & w^2 & -w^3 \\ -w^2 & 0 & w^1 \\ w^3 & -w^1 & 0 \end{pmatrix} \end{aligned}$$

and therefore is a linear combination of the three generators

$$\begin{aligned}
J_1 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \\
J_2 &= \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\
J_3 &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad (2.8)
\end{aligned}$$

Notice that any three independent, antisymmetric matrices could serve as the generators. This is why the Lie algebra is defined as the entire vector space

$$v^1 J_1 + v^2 J_2 + v^3 J_3 = \mathbf{v} \cdot \mathbf{J}$$

As we found for $O(2)$, every element of $SO(3)$ is of the form $\exp(\mathbf{v} \cdot \mathbf{J})$.

The three generators, J_i have been chosen so that their components are given by the Levi-Civita tensor,

$$[J_i]_{jk} = \varepsilon_{ijk}$$

where ε_{ijk} is the totally antisymmetric Levi-Civita tensor. For example, $[J_1]_{ij} = \varepsilon_{1ij} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$.

Again we may use the generators to recover an arbitrary rotation. Starting with

$$O = 1 + \varepsilon$$

we may apply O repeatedly, taking the limit

$$\begin{aligned}
O(\theta) &= \lim_{n \longrightarrow \infty} O^n \\
&= \lim_{n \longrightarrow \infty} (1+\varepsilon)^n \\
&= \lim_{n \longrightarrow \infty} \left(1 + w^i J_i\right)^n
\end{aligned}$$

Let ε be the length of the infinitesmal vector $\mathbf{w}$, so that $\mathbf{w} = \varepsilon \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is a unit vector. Then the limit is taken in such a way that

$$\lim_{n \longrightarrow \infty} n\varepsilon = \varphi$$

where φ is finite. Using the binomial expansion, $(a+b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^{n-k} b^k$ and following the combinatoric argument that led to Eq.(2.5) we have

$$\begin{aligned}
\lim_{n \longrightarrow \infty} [O(\varepsilon)]^n &= \sum_{k=0}^{\infty} \frac{1}{k!} (\varphi \mathbf{n} \cdot \mathbf{J})^k \\
&\equiv \exp(\varphi \mathbf{n} \cdot \mathbf{J})
\end{aligned}$$

We now show that this is the matrix for a rotation through an angle φ around an axis in the direction of $\mathbf{n}$.

To find the detailed form of a general rotation, we now need to find powers of $\mathbf{n}\cdot\mathbf{J}$. This turns out to be straightforward when we use the Levi-Civiat tensor:

$$\begin{aligned}
[\mathbf{n}\cdot\mathbf{J}]^j_{\ k} &= n^i \varepsilon_i^{\ j}{}_{k} \\
\left[(\mathbf{n}\cdot\mathbf{J})^2\right]^m_{\ n} &= \left(n^i \varepsilon_i^{\ m}{}_{k}\right)\left(n^j \varepsilon_j^{\ k}{}_{n}\right) \\
&= n^i n^j \varepsilon_i^{\ m}{}_{k} \varepsilon_j^{\ k}{}_{n} \\
&= -n^i n^j \varepsilon_i^{\ m}{}_{k} \varepsilon_{jn}^{\ \ \ k} \\
&= -n^i n^j \left(\delta_{ij}\delta^m_n - \delta_{in}\delta^m_j\right) \\
&= -\left(\delta_{ij} n^i n^j\right)\delta^m_n + \delta_{in} n^i \delta^m_j n^j \\
&= -\delta^m_n + n^m n_n \\
&= -\left(\delta^m_n - n^m n_n\right) \\
\left[(\mathbf{n}\cdot\mathbf{J})^3\right]^m_{\ n} &= -\left(\delta^m_k - n^m n_k\right) n^i \varepsilon_{i\ n}^{\ k} \\
&= -\delta^m_k n^i \varepsilon_{i\ n}^{\ k} + n^m n_k n^i \varepsilon_{i\ n}^{\ k} \\
&= -n^i \varepsilon_{i\ n}^{\ m} + n^m n^k n^i \varepsilon_{ikn} \\
&= -\left[(\mathbf{n}\cdot\mathbf{J})\right]^m_{\ n}
\end{aligned}$$

where $n^m n^k n^i \varepsilon_{ikn} = n^m (\mathbf{n}\times\mathbf{n}) = 0$. Notice that $(\delta^m_n - n^m n_n)$ is a projection operator, since it is idempotent,

$$\left(\delta^m_n - n^m n_n\right)\left(\delta^n_k - n^n n_k\right) = \left(\delta^m_k - n^m n_k - n^m n_k + n^m n_k\right) = \left(\delta^m_k - n^m n_k\right)$$

Noting that $(\delta^m_n - n^m n_n)\, n^n = 0$, we infer that it projects out vectors orthogonal to $\mathbf{n}$.

The powers come back to $\mathbf{n}\cdot\mathbf{J}$ with only a sign change, so we can divide the series into even and odd powers. For all $k>0$,

$$\begin{aligned}
\left[(\mathbf{n}\cdot\mathbf{J})^{2k}\right]^m_{\ n} &= (-1)^k\left(\delta^m_n - n^m n_n\right) \\
\left[(\mathbf{n}\cdot\mathbf{J})^{2k+1}\right]^m_{\ n} &= (-1)^k\left[\mathbf{n}\cdot\mathbf{J}\right]^m_{\ n}
\end{aligned}$$

For $k=0$ we have the identity, $\left[(\mathbf{n}\cdot\mathbf{J})^0\right]^m_{\ n} = \delta^m_n$.

We can now compute the exponential explicitly:

$$\begin{aligned}
\left[O(\varphi,\hat{\mathbf{n}})\right]^m_{\ n} &= \left[\exp(\varphi\mathbf{n}\cdot\mathbf{J})\right]^m_{\ n} \\
&= \left[\sum_{k=0}^{\infty}\frac{1}{k!}\varphi^k(\mathbf{n}\cdot\mathbf{J})^k\right]^m_{\ n} \\
&= \left[\sum_{l=0}^{\infty}\frac{1}{(2l)!}\varphi^{2l}(\mathbf{n}\cdot\mathbf{J})^{2l}\right]^m_{\ n} + \left[\sum_{l=0}^{\infty}\frac{1}{(2l+1)!}\varphi^{2l+1}(\mathbf{n}\cdot\mathbf{J})^{2l+1}\right]^m_{\ n} \\
&= \delta^m_n + \left(\delta^m_n - n^m n_n\right)\sum_{l=1}^{\infty}\frac{(-1)^l}{(2l)!}\varphi^{2l} + \left[\mathbf{n}\cdot\mathbf{J}\right]^m_{\ n}\sum_{l=0}^{\infty}\frac{(-1)^l}{(2l+1)!}\varphi^{2l+1} \\
&= \delta^m_n + \left(\delta^m_n - n^m n_n\right)(\cos\varphi - 1) + n^i \varepsilon_i^{\ m}{}_{n}\sin\varphi
\end{aligned}$$

where we get $(\cos\varphi - 1)$ because the $l=0$ term is missing from the sum.

To see what this means, let O act on an arbitrary vector $\mathbf{v}$, and write the result in normal vector notation,

$$\begin{aligned}
\left[O(\varphi,\hat{\mathbf{n}})\right]^m_{\ n} v^n &= \delta^m_n v^n + (\cos\varphi - 1)\left(\delta^m_n - n^m n_n\right) v^n + n^i \varepsilon_{i\ n}^{\ m} v^n \sin\varphi \\
&= v^m + (\cos\varphi - 1)\left(v^m - n^m n_n v^n\right) - n^i v^n \varepsilon_{in}^{\ \ \ m}\sin\varphi \\
&= v^m + (\cos\varphi - 1)\left(v^m - (\mathbf{n}\cdot\mathbf{v})\, n^m\right) - [\mathbf{n}\times\mathbf{v}]^m\sin\varphi
\end{aligned}$$

Going fully to vector notation,

$$O(\theta, \hat{\mathbf{n}})\,\mathbf{v} = \mathbf{v} + (\cos\theta - 1)(\mathbf{v} - (\mathbf{n}\cdot\mathbf{v})\,\mathbf{n}) - (\mathbf{n}\times\mathbf{v})\sin\theta$$

Finally, define the components of $\mathbf{v}$ parallel and perpendicular to the unit vector $\mathbf{n}$:

$$\begin{aligned} \mathbf{v}_{\parallel} &= (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n} \\ \mathbf{v}_{\perp} &= \mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n} \end{aligned}$$

In terms of these, the rotated vector becomes,

$$\begin{aligned} O(\theta, \hat{\mathbf{n}})\,\mathbf{v} &= \mathbf{v} - (\mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n}) + (\mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n})\cos\theta - \sin\theta\,(\mathbf{n}\times\mathbf{v}) \\ &= \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}\cos\theta - \sin\theta\,(\mathbf{n}\times\mathbf{v}) \end{aligned} \quad (2.9)$$

This expresses the rotated vector in terms of three mutually perpendicular vectors, $\mathbf{v}_{\parallel}, \mathbf{v}_{\perp}, (\mathbf{n}\times\mathbf{v})$. The direction $\mathbf{n}$ is the axis of the rotation. The part of $\mathbf{v}$ parallel to $\mathbf{n}$ is therefore unchanged. The rotation takes place in the plane perpendicular to $\mathbf{n}$, and this plane is spanned by $\mathbf{v}_{\perp}, (\mathbf{n}\times\mathbf{v})$. The rotation in this plane takes $\mathbf{v}_{\perp}$ into the linear combination $\mathbf{v}_{\perp}\cos\theta - (\mathbf{n}\times\mathbf{v})\sin\theta$, which is exactly what we expect for a rotation of $\mathbf{v}_{\perp}$ through an angle θ. The rotation $O(\theta, \hat{\mathbf{n}})$ is therefore a rotation by θ around the axis $\hat{\mathbf{n}}$.

2.3.3 Definition of a Lie algebra

A Lie algebra has three defining properties.

Define: A *Lie algebra* is a finite dimensional vector space V together with a bilinear, antisymmetric (commutator) product satisfying

1. For all $u, v \in V$, the product $[u, v] = -[v, u] = w$ is in V.
2. All $u, v, w \in V$ satisfy the Jacobi identity

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0 \quad (2.10)$$

These properties may be expressed in terms of a basis. Let $\{J_a | a = 1, \ldots, n\}$ be a vector basis for V. Then we may compute the commutators of the basis,

$$[J_a, J_b] = w_{ab}$$

where for each a and each b, w_{ab} is some vector in V. We may expand each w_{ab} in the basis as well,

$$w_{ab} = c_{ab}{}^{c} J_c$$

for some constants $c_{ab}{}^{c}$. The $c_{ab}{}^{c} = -c_{ba}{}^{c}$ are called the *Lie structure constants*. The basis then satisfies,

$$[J_a, J_b] = c_{ab}{}^{c} J_c \quad (2.11)$$

which is sufficient, using linearity, to determine the commutators of all elements of the algebra:

$$\begin{aligned} [u, v] &= \left[u^a J_a, v^b J_b\right] \\ &= u^a v^b [J_a, J_b] \\ &= u^a v^b c_{ab}{}^{c} J_c \\ &= w^c J_c \\ &= w \end{aligned}$$

Exercise: Show that the commutation relations of the three $SO(3)$ generators, J_i, given in Eq.(2.8) are given by

$$[J_i, J_j] = \varepsilon_{ij}{}^{k} J_k \tag{2.12}$$

where $\varepsilon_{ij}{}^{k} = g^{km}\varepsilon_{ijm}$, and ε_{ijm} is the 3-dimensional version of the totally antisymmetric *Levi-Civita* tensor,

$$\begin{aligned}\varepsilon_{123} &= \varepsilon_{231} = \varepsilon_{312} = 1 \\ \varepsilon_{132} &= \varepsilon_{321} = \varepsilon_{213} = -1\end{aligned}$$

with all other components vanishing. See our discussion of invariant tensors in the section on special relativity for further properties of the Levi-Civita tensors. In particular, you will need

$$\varepsilon^{ijk}\varepsilon_{imn} = \delta^j_m\delta^k_n - \delta^j_n\delta^k_m$$

Notice that most of the calculations above for $SO(3)$ actually apply to any of the pseudo-orthogonal groups $SO(p,q)$. In the general case, the form of the generators is still given by the antisymmetry constraint of Eq.(2.7), with g_{mn} replaced by η_{mn} of eq.(2.3).

$$\begin{aligned}\eta_{mn} &= \eta_{ij}\left(\delta^i_m + \varepsilon^i{}_m\right)\left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= \eta_{mn} + \eta_{ni}\varepsilon^i{}_m + \eta_{mj}\varepsilon^j{}_n\end{aligned}$$

and using the metric to lower the index, $\varepsilon_{nm} = \eta_{ni}\varepsilon^i{}_m$, leading to

$$\varepsilon_{nm} = -\varepsilon_{mn}$$

so that the doubly covariant generators are still antisymmetric. The only difference is that the indices are lowered with M_{mn} instead of g_{mn}. Another difference occurs when we compute the Lie algebra because in n-dimensions we no longer have the convenient form, ε_{ijm}, for the Levi-Civita tensor. The Levi-Civita tensor in n-dimensions has n indices, and does not simplify the Lie algebra expressions. Instead, we choose the following set of antisymmetric matrices as generators:

$$\left[\varepsilon^{(rs)}\right]_{mn} = \left(\delta^r_m\delta^s_n - \delta^r_n\delta^s_m\right) \tag{2.13}$$

The (rs) indices tell us *which* generator we are talking about, while the m and n indices are the matrix components. To compute the Lie algebra, we need the mixed form of the generators,

$$\begin{aligned}\left[\varepsilon^{(rs)}\right]^m{}_n &= \eta^{mk}\left[\varepsilon^{(rs)}\right]_{kn} = \eta^{mk}\delta^r_k\delta^s_n - \eta^{mk}\delta^r_n\delta^s_k \\ &= \eta^{mr}\delta^s_n - \eta^{ms}\delta^r_n\end{aligned}$$

We can now calculate the Lie algebra for *any* $SO(p,q)$:

$$\begin{aligned}\left[\left[\varepsilon^{(uv)}\right],\left[\varepsilon^{(rs)}\right]\right]^m{}_n &= \left[\varepsilon^{(uv)}\right]^m{}_k\left[\varepsilon^{(rs)}\right]^k{}_n - \left[\varepsilon^{(rs)}\right]^m{}_k\left[\varepsilon^{(uv)}\right]^k{}_n \\ &= \left(\eta^{mu}\delta^v_k - \eta^{mv}\delta^u_k\right)\left(\eta^{kr}\delta^s_n - \eta^{ks}\delta^r_n\right) \\ &\quad - \left(\eta^{mr}\delta^s_k - \eta^{ms}\delta^r_k\right)\left(\eta^{ku}\delta^v_n - \eta^{kv}\delta^u_n\right) \\ &= \eta^{mu}\eta^{vr}\delta^s_n - \eta^{mu}\eta^{vs}\delta^r_n - \eta^{mv}\eta^{ur}\delta^s_n + \eta^{mv}\eta^{us}\delta^r_n \\ &\quad -\eta^{mr}\eta^{su}\delta^v_n + \eta^{ms}\eta^{ru}\delta^v_n + \eta^{mr}\eta^{sv}\delta^u_n - \eta^{ms}\eta^{rv}\delta^u_n\end{aligned}$$

Rearranging to collect the terms as generators, note that each generator must have the free m and n indices, while the structure constants depend only on u, v, r, s. Separating with this in mind, we get

$$\begin{aligned}\left[\left[\varepsilon^{(uv)}\right],\left[\varepsilon^{(rs)}\right]\right]^m{}_n &= \eta^{vr}\left(\eta^{mu}\delta^s_n - \eta^{ms}\delta^u_n\right) - \eta^{vs}\left(\eta^{mu}\delta^r_n - \eta^{mr}\delta^u_n\right) \\ &\quad -\eta^{ur}\left(\eta^{mv}\delta^s_n - \eta^{ms}\delta^v_n\right) + \eta^{us}\left(\eta^{mv}\delta^r_n - \eta^{mr}\delta^v_n\right) \\ &= \eta^{vr}\left[\varepsilon^{(us)}\right]^m{}_n - \eta^{vs}\left[\varepsilon^{(ur)}\right]^m{}_n - \eta^{ur}\left[\varepsilon^{(vs)}\right]^m{}_n + \eta^{us}\left[\varepsilon^{(vr)}\right]^m{}_n\end{aligned}$$

Finally, we can drop the matrix indices. It is important that we can do this, because it demonstrates that the Lie algebra is a relationship among the different generators that and not dependent on whether the operators are written as matrices or not. The result, valid for any $SO(p,q)$, is

$$\left[\varepsilon^{(uv)}, \varepsilon^{(rs)}\right] = \eta^{vr}\varepsilon^{(us)} - \eta^{vs}\varepsilon^{(ur)} - \eta^{ur}\varepsilon^{(vs)} + \eta^{us}\varepsilon^{(vr)} \tag{2.14}$$

We will need this result when we study the Dirac matrices.

Exercise: Show that the $SO(p,q)$ Lie algebra in eq.(2.14) reduces to the $SO(3)$ Lie algebra in eq.(2.12) when $(p,q) = (3,0)$. (Hint: go back to Eq.(2.14) and multiply the whole equation by $\varepsilon_{uvw}\varepsilon_{rst}$. Notice that η_{mn} is just the identity g_{mn} and that $J_i = \frac{1}{2}\varepsilon_{ijk}\varepsilon^{(jk)}$).

2.3.4 From Lie groups to Lie algebras (optional)

Every Lie group gives rise to a Lie algebra as its vector space of infinitesimal generators. Conversely, the properties of a Lie algebra guarantee that exponentiating the algebra gives a Lie group. The relationship is not quite one-to-one, since any quotient of the exponentiated Lie algebra by a discrete subgroup gives a Lie group with the same Lie algebra.

To see the relationship betwen the properties of Lie groups and those of Lie algebras, let's work from the group side. We have group elements that depend on continuous parameters, so we can expand $g(a, b, \ldots, c)$ near the identity in a Taylor series:

$$\begin{aligned}g(x^1, \ldots, x^n) &= 1 + \frac{\partial g}{\partial x^a}x^a + \frac{1}{2}\frac{\partial^2 g}{\partial x^a x^b}x^a x^b + \ldots \\ &\equiv 1 + J_a x^a + \frac{1}{2}K_{ab}x^a x^b + \ldots\end{aligned}$$

Now consider the consequences of the properties of the group on the infinitesimal generators, J_a. First, there exists a group product, which must close:

$$\begin{aligned}g(x_1^a)g\left(x_2^b\right) &= g(x_3^a) \\ \left(1 + J_a x_1^a + \ldots\right)\left(1 + J_a x_2^a + \ldots\right) &= 1 + J_a x_3^a + \ldots \\ 1 + J_a x_1^a + J_a x_2^a + \ldots &= 1 + J_a x_3^a + \ldots\end{aligned}$$

so that at linear order,

$$J_a x_1^a + J_a x_2^a = J_a x_3^a$$

This requires the generators to combine linearly under addition and scalar multiplication. Next, we require an identity operator. This just means that the zero vector lies in the space of generators, since $g(0, \ldots, 0) = 1 = 1 + J_a 0^a$. For inverses, we have

$$\begin{aligned}g(x_1^a)g^{-1}\left(x_2^b\right) &= 1 \\ \left(1 + J_a x_1^a + \ldots\right)\left(1 + J_a x_2^a + \ldots\right) &= 1 \\ 1 + J_a x_1^a + J_a x_2^a &= 1\end{aligned}$$

so that $x_2^a = -x_1^a$, guarantees an additive inverse in the space of generators. These properties together make the set $\{x^a J_a\}$ a vector space.

Now we need the commutator product. For this, consider the (closed!) product of group elements

$$g_1 g_2 g_1^{-1} g_2^{-1} = g_3$$

We need to compute this in a Taylor series to second order, so we need the inverse to second order.

Exercise: Show to second order that the inverse of

$$g \equiv 1 + J_a x^a + \frac{1}{2} K_{ab} x^a x^b + \dots$$

is

$$g^{-1} \equiv 1 - J_b x^b + \frac{1}{2} \left(J_a J_b + J_b J_a - K_{ab} \right) x^a x^b + \dots$$

Now, expanding to second order in the Taylor series,

$$\begin{aligned}
g_3 &= 1 + J_a z^a(x,y) + \frac{1}{2} K_{ab} z^a(x,y) z^b(x,y) \\
&= \left(1 + J_a x^a + \frac{1}{2} K_{ab} x^a x^b\right)\left(1 + J_b\ y^b + \frac{1}{2} K_{bc} y^b y^c\right) \\
&\quad \times \left(1 - J_c x^c + \left(J_c J_d - \frac{1}{2} K_{cd}\right) x^c x^d\right)\left(1 - J_d y^d + \left(J_d J_e - \frac{1}{2} K_{de}\right) y^d y^e\right) \\
&= \left(1 + J_b x^b + J_b y^b + J_a J_b x^a y^b + \frac{1}{2} K_{bc} y^b y^c + \frac{1}{2} K_{ab} x^a x^b\right) \\
&\quad \times \left(1 - J_d x^d - J_d y^d + J_d J_e y^d y^e + J_c J_d x^c y^d + J_c J_d x^c x^d - \frac{1}{2} K_{de} y^d y^e - \frac{1}{2} K_{cd} x^c x^d\right) \\
&= 1 - J_d x^d - J_d y^d + J_d J_e y^d y^e + J_c J_d x^c y^d + J_c J_d x^c x^d \\
&\quad - \frac{1}{2} K_{de} y^d y^e - \frac{1}{2} K_{cd} x^c x^d + \left(J_b x^b + J_b y^b\right)\left(1 - J_d x^d - J_d y^d\right) \\
&\quad + J_a J_b x^a y^b + \frac{1}{2} K_{bc} y^b y^c + \frac{1}{2} K_{ab} x^a x^b
\end{aligned}$$

Collecting terms,

$$\begin{aligned}
g_3 &= 1 + J_a z^a(x,y) + \cdots \\
&= 1 - J_d x^d - J_d y^d + J_b x^b + J_b y^b + J_d J_e y^d y^e + J_c J_d x^c y^d + J_c J_d x^c x^d - J_b J_d x^b x^d - J_b J_d y^b x^d \\
&\quad - J_b J_d x^b y^d - J_b J_d y^b y^d + J_a J_b x^a y^b + \frac{1}{2} K_{bc} y^b y^c + \frac{1}{2} K_{ab} x^a x^b - \frac{1}{2} K_{de} y^d y^e - \frac{1}{2} K_{cd} x^c x^d \\
&= 1 + J_c J_d x^c y^d - J_b J_d y^b x^d \\
&= 1 + J_c J_d x^c y^d - J_d J_c x^c y^d \\
&= 1 + \left[J_c, J_d\right] x^c y^d
\end{aligned}$$

Equating the expansion of g_3 to the collected terms we see that we must have z^a such that

$$\left[J_c, J_d\right] x^c y^d = J_a z^a(x,y)$$

Since x^c and y^d are arbitrary, z^a must be bilinear in them:

$$z^a = x^c y^d c_{cd}^{\ \ a}$$

and we have derived the need for a commutator product for the Lie algebra,

$$[J_c, J_d] = c_{cd}{}^a J_a$$

Finally, the Lie group is associative: if we have three group elements, g_1, g_2 and g_3, then

$$g_1 (g_2 g_3) = (g_1 g_2) g_3$$

This is not easy to expand, since the relationship we seek among the generators must be third order. We would need to expand each group element to third order to find the necessary terms. However, since we are principally interested in linear representations, the generators will be matrices and therefore associative. This is already a stronger condition than the Jacobi identity.

Exercise: To show the sufficiency (but not the necessity!) of associative generators, assume the generators are associative and confirm that the Jacobi identity,

$$0 = [J_a, [J_b, J_c]] + [J_b, [J_c, J_a]] + [J_c, [J_a, J_b]]$$

holds identically.

To carry out a proof that the Jacobi identity is the necessary and sufficient condition for the Lie algebra to extend to a Lie group, it is easiest to rewrite the Lie commutator using differential forms dual to the generators. This gives a differential equation, the Maurer-Cartan equation for the Lie group, and the Jacobi identity is easily shown to be the integrability condition for the Maurer-Cartan equation.

With this, the definition of a Lie algebra is a necessary consequence of being built from the infinitesimal generators of a Lie group. The conditions are also sufficient, though we won't give the proof here.

The correspondence between Lie groups and Lie algebras is not one to one, because in general several Lie groups may share the same Lie algebra. However, groups with the same Lie algebra are related in a simple way. Our example above of the relationship between $O(3)$ and $SO(3)$ is typical – these two groups are related by a discrete symmetry. Since discrete symmetries do not participate in the computation of infinitesimal generators, they do not change the Lie algebra. The central result is this: for every Lie algebra there is a unique maximal Lie group called the *covering group* such that every Lie group sharing the same Lie algebra is the group quotient of the covering group by a discrete symmetry group. This result suggests that when examining a group symmetry of nature, we should always look at the covering group in order to extract the greatest possible symmetry. Following this suggestion for Euclidean 3-space and for Minkowski space leads us directly to the use of *spinors*.

In the next section, we discuss spinors in three way *spinors*. The first two make use of convenient tricks that work in low dimensions (2, 3 and 4), and provide easy ways to handle rotations and Lorentz transformations. The third treatment is begins with Dirac's development of the Dirac equation, which leads us to the introduction of Clifford algebras.

2.4 Spinors for rotations

When we work with linear representations of Lie groups and Lie algebras, it is important to keep track of the representation – the vector space on which the group acts. In the case of $O(3)$, the vector space is Euclidean 3-space, while for Lorentz transformations the vector space is spacetime. As we shall see in this section, the covering groups of these same symmetries act on other, more abstract, complex vector spaces. The elements of these complex vector spaces are called spinors.

2.4.1 A complex representation for real 3-vectors

Let's start with $SO(3)$, the group which preserves the lengths, $\mathbf{x}^2 = x^2 + y^2 + z^2 = g_{ij} x^i x^j$ of Euclidean 3-vectors. We can encode this length as the determinant of a matrix. Let the real 3-vector $\mathbf{x} = (x, y, z)$ be represented as the matrix

$$X \equiv \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \qquad (2.15)$$

Then the determinant of X is

$$\det X = -\left(x^2 + y^2 + z^2\right)$$

This fact is useful because matrices of this type are easy to characterize. Let

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

be any matrix with complex entries, and demand hermiticity, $M = M^\dagger$:

$$\begin{aligned} M &= M^\dagger \\ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} &= \begin{pmatrix} \alpha^* & \gamma^* \\ \beta^* & \delta^* \end{pmatrix} \end{aligned}$$

Then $\alpha \to a$ is real, $\delta \to d$ is real, and $\beta = \gamma^*$. Only $\gamma = b + ic$ remains arbitrary, so that

$$M = \begin{pmatrix} a & b - ic \\ b + ic & d \end{pmatrix} \qquad (2.16)$$

If we further require M to be traceless, then M reduces to

$$M = \begin{pmatrix} a & b - ic \\ b + ic & -a \end{pmatrix}$$

just the same as X. Therefore, X is a general traceless, Hermitian 2×2 matrix.

2.4.2 SU(2)

Rotations may now be characterized as the set of transformations of X preserving the following properties of X :

1. $\det X = -|\mathbf{x}|^2$
2. $X^\dagger = X$
3. $tr\,(X) = 0$

To find the set of such transformations, recall that matrices transform by a similarity transformation

$$X \to X' = AXA^\dagger$$

Here we use the adjoint instead of the inverse because we imagine X as doubly covariant, X_{ij}. For the mixed form, $X^i{}_j$ we would write $X \to AXA^{-1}$. We use the adjoiont instead of the transpose because we allow X to be complex.

From this form, we have:

$$\begin{aligned} \det X' &= \det\left(AXA^\dagger\right) \\ &= (\det A)(\det X)\left(\det A^\dagger\right) \end{aligned}$$

so in order to have $\det X' = \det X$ we demand $|\det A|^2 = 1$ so that for some real φ,

$$\det A = e^{i\varphi}$$

We can constrain this further, because if we write

$$A = e^{i\varphi/2}U$$

where $\det U = 1$ then

$$\begin{aligned} X' &= AXA^{\dagger} \\ &= e^{i\varphi/2}UXe^{-i\varphi/2}U^{\dagger} \\ &= UXU^{\dagger} \end{aligned}$$

That is, without loss of generality, we can take the determinant to be one because an overall phase has no effect on X.

Next, notice that hermiticity is automatic. Whenever X is hermitian we have

$$\begin{aligned} (X')^{\dagger} &= \left(AXA^{\dagger}\right)^{\dagger} \\ &= A^{\dagger\dagger}X^{\dagger}A^{\dagger} \\ &= AXA^{\dagger} \\ &= X' \end{aligned}$$

so X' is also hermitian.

Finally, we impose the trace condition. Suppose $tr(X) = 0$. Then

$$tr(X') = tr(AXA^{\dagger}) = tr(A^{\dagger}AX)$$

where we use the cyclic property of the trace. For the final expression to reduce to $tr(X)$ for all X, we must have $A^{\dagger}A = 1$. Therefore, $A^{\dagger} = A^{-1}$ and the transformations must be *unitary*. Using the unit determinant unitary matrices, U, we see that the group is $SU(2)$. This shows that $SU(2)$ can be used to write 3-dimensional rotations. In fact, we will see that $SU(2)$ includes two transformations corresponding to each element of $SO(3)$.

The exponential of any anti-hermitian matrix is unitary matrix because if $U = \exp(iH)$ with $H^{\dagger} = H$, then

$$U^{\dagger} = \exp\left(-iH^{\dagger}\right) = \exp(-iH) = U^{-1}$$

Conversely, any unitary matrix may be written this way. Moreover, since

$$\det A = e^{tr(\ln A)}$$

the transformation $U = \exp(iH)$ has unit determinant whenever H is traceless. Since every traceless, hermitian matrix is a linear combination of the Pauli matrices,

$$\sigma_m = \left(\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} & -i \\ i & \end{pmatrix}, \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \right) \tag{2.17}$$

we may write every element of $SU(2)$ as the exponential

$$U(\mathbf{w}) = e^{i\mathbf{w}\cdot\boldsymbol{\sigma}}$$

where the dot product in the exponent is the matrix $\mathbf{w}\cdot\boldsymbol{\sigma} = w^m\sigma_m$ and the three parameters w^m are real. The Pauli matrices are mixed type tensors, $\sigma_m = [\sigma_m]^a{}_b$, because U is a transformation matrix. It is most convenient to write the vector $\mathbf{w}$ as a magnitude $\frac{\varphi}{2}$ times a unit vector, $\mathbf{w} = \frac{\varphi}{2}\hat{\mathbf{n}}$, so that φ is then the angle through which a real 3-vector is rotated. Then the elements of $SU(2)$ are then written as

$$U(\varphi, \hat{\mathbf{n}}) = \exp\left(\frac{i\varphi}{2}\hat{\mathbf{n}}\cdot\sigma\right) \tag{2.18}$$

This transformation performs a rotation through an angle φ about the $\hat{\mathbf{n}}$ direction.

Exercise: Show the the product of any two Pauli matrices may be written as

$$\sigma_m \sigma_n = \delta_{mn} \mathbf{1} + i\varepsilon_{mnk}\sigma_k \tag{2.19}$$

Exercise: Show that $U(2\pi, \hat{\mathbf{n}}) = -1$ and $U(4\pi, \hat{\mathbf{n}}) = 1$ for any unit vector, $\hat{\mathbf{n}}$. From this, show that $U(2\pi, \hat{\mathbf{n}})$ nonetheless gives $X' = X$.

Exercise: By expanding the exponential in a power series and working out the powers of $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$ for a general unit vector $\hat{\mathbf{n}}$, prove the identity

$$\exp\left(\frac{i\varphi}{2}\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}\right) = \mathbf{1}\cos\frac{\varphi}{2} + i\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}\ \sin\frac{\varphi}{2} \tag{2.20}$$

A general 3-vector may be written using the Pauli matrices as $X = \mathbf{x} \cdot \boldsymbol{\sigma}$, so the correspondence is one-to-one. This allows us to carry out rotations on 3-vectors using $SU(2)$.

Exercise: Reproduce Eq.(2.9) by transforming $X = \mathbf{x} \cdot \boldsymbol{\sigma}$ using $SU(2)$ to find $X' = \begin{pmatrix} z' & x' - iy' \\ x' + iy' & -z' \end{pmatrix} = UXU^\dagger$. Since we know that X' must also be traceless and Hermitian we may write

$$\mathbf{x}' \cdot \boldsymbol{\sigma} = \left(\exp\left(\frac{i\varphi}{2}\hat{\mathbf{n}} \cdot \sigma\right)\right)(\mathbf{x} \cdot \boldsymbol{\sigma})\left(\exp\left(-\frac{i\varphi}{2}\hat{\mathbf{n}} \cdot \sigma\right)\right)$$

Use Eq.(2.20) to write the exponentials. This shows that $\exp\left(\frac{i\varphi}{2}\hat{\mathbf{n}} \cdot \sigma\right)$ describes a rotation by φ about the axis through $\hat{\mathbf{n}}$.

2.4.3 The representation for $SU(2)$

Now consider what vector space $SU(2)$ acts on. We have used a similarity transformation on matrices to show how it acts on a 3-dimensional subspace of the 8-dimensional space of 2×2 complex matrices. But more basically, $SU(2)$ acts the vector space of complex, two component *spinors*:

$$\chi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$
$$\chi' = U\chi$$

Exercise: Using Eq.(2.20) from the previous exercise, find the most general action of $SU(2)$ on χ. As special cases, show that the periodicity of the mapping is 4π, that is, that

$$U(4\pi m, \hat{\mathbf{n}})\chi = \chi$$

for all integers m, while

$$U(2\pi m, \hat{\mathbf{n}})\chi = -\chi \neq \chi$$

for odd m.

The vector space of spinors χ is the simplest set of objects that Euclidean rotations act nontrivially on. These objects are familiar from quantum mechanics as the spin-up and spin-down states of spin-1/2 fermions. It is interesting to observe that spin is a perfectly classical property arising from symmetry. It was not necessary to discover quantum mechanics in order to discover spin. Apparently, the reason that "classical spin" was not discovered first is that its magnitude is microscopic. Indeed, with the advent of supersymmetry, there has been some interest in classical supersymmetry – supersymmetric classical theories whose quantization leads to now-familiar quantum field theories.

2.5 Spinors for the Lorentz group: SL(2,C)

Next, we extend this same approach to the Lorentz group. Recall that we defined Lorentz transformations as those preserving the Minkowski 4-vector length,

$$s^2 = t^2 - (x^2 + y^2 + z^2) \tag{2.21}$$

or equivalently, those transformations leaving the Minkowski metric $\eta_{\mu\nu}$ invariant. Once again, we write a matrix that contains the invariant information in its determinant. Let

$$[X]_{AB} \equiv \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix}$$

noting from Eq.(2.16) that X is now the most general hermitian 2×2 matrix, $X^\dagger = X$, without any constraint on the trace. The determinant is now

$$\det X = t^2 - x^2 - y^2 - z^2 = s^2$$

and we only need to preserve two properties: hermiticity and the determinant.

Let $[X']_{AB} = [X']_{CD}\, A^C{}_A A^D{}_B$, or in matrix notation,

$$X' = AXA^\dagger$$

Then hermiticity is again automatic and all we need is $|\det A|^2 = 1$ to preserve the determinant, $\det X' = \det X$. As before, an overall phase does not affect X, so we can choose $\det A = 1$. There is no further constraint needed, so Lorentz transformations are given by the special linear group in two complex dimensions, $SL\,(2,\mathbb{C})$.

It is easy to find a set of generators for the general linear group, because every non-degenerate matrix is allowed. Expanding a general group element g infinitesimally about the identity gives

$$\begin{aligned} g &= \begin{pmatrix} \mu & \nu \\ \rho & \sigma \end{pmatrix} \\ &= 1 + \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \\ &= 1 + \begin{pmatrix} a & b \\ c & d \end{pmatrix} + i \begin{pmatrix} e & f \\ g & h \end{pmatrix} \end{aligned}$$

for complex numbers $\alpha, \beta, \gamma, \delta$ and small real parameters $a, \ldots, h$. Since the deviation from the identity is small, the determinant will be close to one, hence nonzero. We recover the whole group by exponentiation,

$$G \ni g = \exp \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

The unit determinant is achieved by making the generators traceless, setting $\delta = -\alpha$. A complete set of generators for $SL(2,C)$ is therefore

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

Because any six independent linear combinations of these are an equivalently good basis, we choose instead the set

$$J_m = i\sigma_m \tag{2.22}$$
$$K_m = \sigma_m \tag{2.23}$$

which have the advantage of being hermitian and anti-hermitian, respectively.

When we exponentiate J_m and K_m (with real parameters) to recover the various types of Lorentz transformation, the anti-hermitian generators J_m give $SU(2)$ as before. We already know that these preserve lengths of spatial 3-vectors, so we see again that the 3-dimensional rotations are part of the Lorentz group. Since the generators K_m are hermitian, the corresponding group elements are not unitary. The corresponding transformations are hyperbolic rather than circular, corresponding to boosts.

Exercise: Recalling the Taylor series

$$\begin{aligned}\sinh\lambda &= \sum_{k=0}^{\infty}\frac{\lambda^{2k+1}}{(2k+1)!}\\ \cosh\lambda &= \sum_{k=0}^{\infty}\frac{\lambda^{2k}}{(2k)!}\end{aligned}$$

show that $K_1 = \begin{pmatrix} & 1\\ 1 & \end{pmatrix}$ generates a boost in spacetime.

The Lie algebra of $SL(2,C)$ is now easy to calculate using the products of the Pauli matrices given by Eq.(2.19). Using Eq.(2.19) that the commutators are $[\sigma_m, \sigma_n] = 2i\varepsilon_{mnk}\sigma_k$, and from this that the Lie algebra $sl\,(2,\mathbb{C})$ of the Lorentz group $SL\,(2,\mathbb{C})$ may be written as

$$[J_m, J_n] = [i\sigma_m, i\sigma_n] = -2i\varepsilon_{mnk}\sigma_k = -2\varepsilon_{mnk}J_k \tag{2.24}$$

$$[J_m, K_n] = [i\sigma_m, \sigma_n] = -2\varepsilon_{mnk}\sigma_k = -2\varepsilon_{mnk}K_k \tag{2.25}$$

$$[K_m, K_n] = [\sigma_m, \sigma_n] = -2i\varepsilon_{mnk}\sigma_k = -2\varepsilon_{mnk}J_k \tag{2.26}$$

This is an important result. It shows that while the rotations form a subgroup of the Lorentz group (because the J_m commutators close into themselves), the boosts do not. Instead, two boosts applied in succession produce a rotation as well as a change of relative velocity. This is the source of a noted correction to Thomas precession (see Jackson, pp. 556-560; indeed, see Jackson's chapters 11 and 12 for a good discussion of special relativity in a context with real examples)

There is another convenient basis for the Lorentz Lie algebra. Consider the six generators

$$L_m = \frac{1}{2}(J_m + K_m) \tag{2.27}$$

$$M_m = \frac{1}{2}(J_m - K_m) \tag{2.28}$$

These satisfy

$$\begin{aligned}[L_m, L_n] &= \frac{1}{4}[J_m + K_m, J_n + K_n]\\ &= \frac{1}{4}(-2\varepsilon_{mnk}J_k - 2\varepsilon_{mnk}K_k - 2\varepsilon_{mnk}K_k - 2\varepsilon_{mnk}J_k)\\ &= -2\varepsilon_{mnk}L_k\\ [L_m, M_n] &= \frac{1}{4}(-2\varepsilon_{mnk}J_k + 2\varepsilon_{mnk}K_k - 2\varepsilon_{mnk}K_k + 2\varepsilon_{mnk}J_k)\\ &= 0\\ [M_m, M_n] &= \frac{1}{4}(-2\varepsilon_{mnk}J_k + 2\varepsilon_{mnk}K_k + 2\varepsilon_{mnk}K_k - 2\varepsilon_{mnk}J_k)\\ &= -2\varepsilon_{mnk}M_k\end{aligned}$$

showing that the Lorentz group actually decouples into two commuting copies of $SU(2)$. Extensive use of this fact is made in general relativity (see, eg., Penrose and Rindler, Wald). In particular, we can use this decomposition of the Lie algebra $sl(2,C)$ to introduce *two* sets of 2-component spinors, called *Weyl spinors.*

$$\chi^A, \bar{\chi}^{\dot{A}} \tag{2.29}$$

with the first set transforming under the action of $\exp(u^m L_m)$ and the second set under $\exp(v^m M_m)$. For our study of field theory, however, we will be more interested in a different set of spinors: the 4-component *Dirac spinors*.

2.6 Dirac spinors and the Dirac equation

Before we develop spinor representations systematically for any pseudo-orthogonal group, $SO(p,q)$, we describe how Dirac arrived at the $SO(1,3)$ spacetime representation when he sought a relativistic form for quantum theory. Without looking at the full details of Dirac's original approach, we use a similar construction. Dirac wanted to build a relativistic quantum theory, and recognizing that relativity requires space and time variables to enter on the same footing, sought an equation *linear* in both space and time derivatives:

$$i\frac{\partial\psi}{\partial t} = \left(-i\alpha^i\partial_i + m\beta\right)\psi \tag{2.30}$$

where the α^i and β are constant. The Klein-Gordon equation,

$$\Box\phi = -m^2\phi \tag{2.31}$$

had already been tried and discarded by Schrödinger because the second order equation requires two initial conditions and the uncertainty principle allows us only one. To determine the coefficients, Dirac demanded that the linear equation should imply the Klein-Gordon equation. Acting on our version of Dirac's equation with the same operator again,

$$\begin{aligned} -\frac{\partial^2\psi}{\partial t^2} &= \left(-i\alpha^i\partial_i + m\beta\right)\left(-i\alpha^i\partial_i + m\beta\right)\psi \\ &= \left(-\alpha^i\alpha^j\partial_i\partial_j - im\alpha^i\beta\partial_i - im\beta\alpha^i\partial_i + m^2\beta^2\right)\psi \end{aligned}$$

we reproduce the Klein-Gordon equation provided

$$\begin{aligned} -\alpha^i\alpha^j\partial_i\partial_j &= -\nabla^2 \\ m\left(\alpha^i\beta + \beta\alpha^i\right)\partial_i &= 0 \\ m^2\beta^2 &= m^2 \end{aligned}$$

or equivalently,

$$\begin{aligned} \alpha^i\alpha^j + \alpha^i\alpha^i &= 2\delta^{ij} \\ \alpha^i\beta + \beta\alpha^i &= 0 \\ \beta^2 &= 1 \end{aligned} \tag{2.32}$$

We can put these conditions into a more systematic form by defining

$$\gamma^\mu = \left(\beta, \beta\alpha^i\right) \tag{2.33}$$

Then the constraints on γ^μ may be neatly expressed as

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} \tag{2.34}$$

where the curly brackets denote the anti-commutator. This relationship is impossible to achieve with vectors. To see this, suppose γ^μ is a 4-vector and note that we can always perform a Lorentz transformation that brings a 4-vector γ^μ to one of the forms

$$\begin{aligned} \gamma^\mu &= (\alpha, 0, 0, 0) \\ \gamma^\mu &= (\alpha, \alpha, 0, 0) \\ \gamma^\mu &= (0, \alpha, 0, 0) \end{aligned}$$

depending on whether γ^μ is timelike, null or spacelike. Then, since $\eta^{\mu\nu}$ is Lorentz invariant, we have the possibilities:

$$\{\gamma^\mu, \gamma^\nu\} = \begin{pmatrix} \alpha^2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

$$\{\gamma^\mu, \gamma^\nu\} = \begin{pmatrix} \alpha^2 & \alpha^2 & & \\ \alpha^2 & \alpha^2 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

$$\{\gamma^\mu, \gamma^\nu\} = \begin{pmatrix} 0 & & & \\ & \alpha^2 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

none of which equals $\eta^{\mu\nu}$. Therefore, γ^μ must be a more general kind of object. It *is* sufficient to let γ^μ be a set of four, 4×4 matrices, and it is not hard to show that this is the smallest size matrix that works.

Exercise: Show that there do not exist four, 2×2 matrices satisfying $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$

Here is a convenient choice for the *Dirac matrices*, or *gamma matrices*:

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \\ & -\mathbf{1} \end{pmatrix} \\ \gamma^i &= \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \end{aligned} \quad (2.35)$$

where the σ^i are the usual 2×2 Pauli matrices.

Exercise: Show that these matrices satisfy $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$.

Substituting γ^μ into eq.(2.30), we have the *Dirac equation*,

$$(i\gamma^\mu \partial_\mu - m)\,\psi = 0 \quad (2.36)$$

and the required *Clifford algebra*,

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$$

This equation gives us more than we bargained for. Since the γ^μ are 4×4 Dirac matrices, the object ψ that they act on must also be a 4-component vector. We now show that ψ is a spinor by showing that they transform as a 4-dimensional complex representation, $Spin(p,q)$ of the Lorentz group.

We give the proof for the general case of $Spin(p,q)$ rather than just $Spin(1,3)$, since the development is essentially the same in all cases. In the process, we will not only see that the object ψ is a spinor, but also find the form for Lorentz transformations.

2.7 The Lie algebra of Spin(p,q)

Let the $O(p,q)$ metric be as in eq.(2.3), $\eta_{ij} = diag(1, \ldots, 1, -1, \ldots, -1)$, with $+1$ occurring p times and -1 occurring q times, and the indices i, j run from 1 to $n = p + q$. Let its inverse be written as η^{ij}. We first define n distinct matrices by

$$\left\{\gamma^i, \gamma^j\right\} = 2\eta^{ij}\mathbf{1} \quad (2.37)$$

This is the defining relationship of a *Clifford algebra.* Notice that there are two matrices on the right side. The metric η^{ij} is just a set of coefficients telling us whether the right side is zero or not for any given pair of gamma matrices. The identity matrix occurs because the γ^i are matrices (with components $\left[\gamma^i\right]^A_{\ \ B}$) and therefore their anticommutator must also be a matrix. Often the identity matrix is suppressed for brevity. It is always possible to choose the gamma matrices so that $\left(\gamma^i\right)^\dagger = \eta^{ii}\left(\gamma^i\right)$, making some hermitian and the rest anti-hermitian. Now, from these gamma matrices we construct the commutator,

$$\sigma^{ij} = \frac{1}{4}\left[\gamma^i, \gamma^j\right] \tag{2.38}$$

Exercise: Show that for spacetime, with $\left(\gamma^0\right)^\dagger = \gamma^0$ and $\left(\gamma^i\right)^\dagger = -\gamma^i$, $\sigma^{\mu\nu}$ has the following hermiticity relations:

$$\begin{aligned} \left(\sigma^{0i}\right)^\dagger &= \sigma^{0i} \\ \left(\sigma^{ij}\right)^\dagger &= -\sigma^{ij} \end{aligned}$$

Next, we show that these commutators satisfy the Lie algebra of $SO(n)$. We first use the anticommutator relation to rearrange the terms. Since the anticommutator relation gives $\gamma^j\gamma^i = -\gamma^i\gamma^j + 2\eta^{ij}$ we can rewrite the commutator as

$$\begin{aligned} \sigma^{ij} &= \frac{1}{4}\left(\gamma^i\gamma^j - \gamma^j\gamma^i\right) \\ &= \frac{1}{4}\left(\gamma^i\gamma^j + \gamma^i\gamma^j - 2\eta^{ij}\mathbf{1}\right) \\ &= \frac{1}{2}\left(\gamma^i\gamma^j - \eta^{ij}\mathbf{1}\right) \end{aligned}$$

Using this relation, the commutator of two sigmas is:

$$\begin{aligned} \left[\sigma^{ij}, \sigma^{kl}\right] &= -\frac{1}{4}\left[\gamma^i\gamma^j - \eta^{ij}\mathbf{1}, \gamma^k\gamma^l - \eta^{kl}\mathbf{1}\right] \\ &= -\frac{1}{4}\left[\gamma^i\gamma^j, \gamma^k\gamma^l\right] \\ &= -\frac{1}{4}\gamma^i\gamma^j\gamma^k\gamma^l + \frac{1}{4}\gamma^k\gamma^l\gamma^i\gamma^j \end{aligned}$$

Now we just rearrange the order of gamma matrices in the second term until it matches the first term. Interchanging gamma matrices one pair at a time,

$$\begin{aligned} \gamma^k\gamma^l\gamma^i\gamma^j &= \gamma^k\left(-\gamma^i\gamma^l + 2\eta^{il}\right)\gamma^j \\ &= -\gamma^k\gamma^i\gamma^l\gamma^j + 2\eta^{il}\gamma^k\gamma^j \\ &= -\gamma^k\gamma^i\left(-\gamma^j\gamma^l + 2\eta^{jl}\right) + 2\eta^{il}\gamma^k\gamma^j \\ &= \gamma^k\gamma^i\gamma^j\gamma^l - 2\eta^{jl}\gamma^k\gamma^i + 2\eta^{il}\gamma^k\gamma^j \\ &= \left(-\gamma^i\gamma^k + 2\eta^{ik}\right)\gamma^j\gamma^l - 2\eta^{jl}\gamma^k\gamma^i + 2\eta^{il}\gamma^k\gamma^j \\ &= -\gamma^i\gamma^k\gamma^j\gamma^l + 2\eta^{ik}\gamma^j\gamma^l - 2\eta^{jl}\gamma^k\gamma^i + 2\eta^{il}\gamma^k\gamma^j \\ &= -\gamma^i\left(-\gamma^j\gamma^k + 2\eta^{jk}\right)\gamma^l + 2\eta^{ik}\gamma^j\gamma^l - 2\eta^{jl}\gamma^k\gamma^i + 2\eta^{il}\gamma^k\gamma^j \\ &= \gamma^i\gamma^j\gamma^k\gamma^l - 2\eta^{jk}\gamma^i\gamma^l + 2\eta^{ik}\gamma^j\gamma^l - 2\eta^{jl}\gamma^k\gamma^i + 2\eta^{il}\gamma^k\gamma^j \end{aligned}$$

Finally, using

$$\begin{aligned} \gamma^i\gamma^j &= \frac{1}{2}\left\{\gamma^i, \gamma^j\right\} + \frac{1}{2}\left[\gamma^i, \gamma^j\right] \\ &= \eta^{ij} - 2\sigma^{ij} \end{aligned}$$

we have

$$
\begin{aligned}
\left[\sigma^{ij},\sigma^{kl}\right] &= -\frac{1}{4}\gamma^i\gamma^j\gamma^k\gamma^l+\frac{1}{4}\gamma^i\gamma^j\gamma^k\gamma^l-\frac{1}{2}\left(M^{jk}\gamma^i\gamma^l-M^{ik}\gamma^j\gamma^l+M^{jl}\gamma^k\gamma^i-M^{il}\gamma^k\gamma^j\right)\\
&= -\frac{1}{2}\left(\eta^{jk}\left(\eta^{il}-2\sigma^{il}\right)-\eta^{ik}\left(\eta^{jl}-2\sigma^{jl}\right)+\eta^{jl}\left(\eta^{ki}-2\sigma^{ki}\right)-\eta^{il}\left(\eta^{kj}-2\sigma^{kj}\right)\right)\\
&= -\frac{1}{2}\left(\eta^{jk}\eta^{il}-\eta^{ik}\eta^{jl}+\eta^{jl}\eta^{ki}-\eta^{il}\eta^{kj}\right)\\
&\quad +\eta^{jk}\sigma^{il}-\eta^{ik}\sigma^{jl}+\eta^{jl}\sigma^{ki}-\eta^{il}\sigma^{kj}
\end{aligned}
$$

The first four terms cancel, leaving the $spin(p,q)$ Lie algebra:

$$
\left[\sigma^{ij},\sigma^{kl}\right]=\eta^{jk}\sigma^{il}-\eta^{ik}\sigma^{jl}+\eta^{jl}\sigma^{ki}-\eta^{il}\sigma^{kj} \tag{2.39}
$$

This is the same algebra we found in Eq.(2.14) for $so\,(p,q)$, but it acts on complex vectors (spinors) instead of real vectors. This shows us why Dirac's wave function ψ is a spinor. Using infinitesimal, real parameters, ε_{rs}, can use linear combinations of the σ^{rs} matrices to generate an infinitesimal Lorentz transformation,

$$
\Lambda^A{}_B = \delta^A{}_B+\frac{1}{2}\varepsilon_{rs}\left[\sigma^{rs}\right]^A{}_B
$$

which act on spinors according to

$$
\left[\psi'\right]^A = \Lambda^A{}_B\left[\psi\right]^B \tag{2.40}
$$

We assume that $\varepsilon_{rs}=-\varepsilon_{sr}$, so the factor of $\frac{1}{2}$ avoids double counting.

To see that ψ is really a spinor, we use them to construct vectors. Let the spinor space have an hermitian metric, h_{AB}, so that we can form inner products of spinors

$$
\langle\chi,\psi\rangle=\left[\chi^\dagger\right]^A h_{AB}\left[\psi\right]^B \tag{2.41}
$$

We require h_{AB} to be invariant under Lorentz transformations, in the sense that

$$
\left[\Lambda^\dagger\right]_C{}^A h_{AB}\Lambda^B{}_D = h_{CD}
$$

For infinitesimal transformations, this means that

$$
\begin{aligned}
h_{CD} &= \left(\delta_C{}^A+\frac{1}{2}\varepsilon_{rs}\left[\left(\sigma^{rs}\right)^\dagger\right]_C{}^A\right)h_{AB}\left(\delta^B{}_D+\frac{1}{2}\varepsilon_{uv}\left[\sigma^{uv}\right]^B{}_D\right)\\
&= h_{CD}+\frac{1}{2}\varepsilon_{rs}h_{CB}\left[\sigma^{rs}\right]^B{}_D+\frac{1}{2}\varepsilon_{rs}\left[\left(\sigma^{rs}\right)^\dagger\right]_C{}^A h_{AD}
\end{aligned}
$$

where we drop the negligible quadratic term. Cancelling the common h_{CD} term we are left with

$$
0 = h_{CB}\left[\sigma^{rs}\right]^B{}_D+\left[\left(\sigma^{rs}\right)^\dagger\right]_C{}^A h_{AD} \tag{2.42}
$$

Now we can build the real, n-dimensional vector

$$
v^i = \left[\psi^\dagger\right]^B h_{BC}\left[\gamma^i\right]^C{}_D\left[\psi\right]^D \tag{2.43}
$$

Suppose we transform ψ according to Eq.(2.40). Then

$$
\left[\psi'^\dagger\right]^A = \left[\psi^\dagger\right]^B\left[\Lambda^\dagger\right]_B{}^A
$$

so v^i changes to

$$\begin{aligned}
[v']^i &= \left[\psi'^\dagger\right]^B h_{BC} \left[\gamma^i\right]^C{}_D \left[\psi'\right]^D \\
&= \left[\psi^\dagger\right]^A \left[\Lambda^\dagger\right]_A{}^B h_{BC} \left[\gamma^i\right]^C{}_D \Lambda^D{}_E \left[\psi\right]^E
\end{aligned}$$

For an infinitesimal transformation the matrix product is

$$\begin{aligned}
\Lambda^\dagger h \gamma^i \Lambda &= \left[\Lambda^\dagger\right]_A{}^B h_{BC} \left[\gamma^i\right]^C{}_D \Lambda^D{}_E \\
&= \left(\delta_A^B + \frac{1}{2}\varepsilon_{rs}\left[(\sigma^{rs})^\dagger\right]_A{}^B\right) h_{BC} \left[\gamma^i\right]^C{}_D \left(\delta^D{}_E + \frac{1}{2}\varepsilon_{uv}\left[\sigma^{uv}\right]^D{}_E\right) \\
&= h_{AC}\left[\gamma^i\right]^C{}_E + \frac{1}{2}\varepsilon_{uv} h_{AC}\left[\gamma^i\right]^C{}_D \left[\sigma^{uv}\right]^D{}_E - \frac{1}{2}\varepsilon_{rs}\left[(\sigma^{rs})^\dagger\right]_A{}^B h_{BC}\left[\gamma^i\right]^C{}_E
\end{aligned}$$

Writing $[v']^i = [v]^i + [\delta v]^i$ and using the Lorentz invariance of h_{AB}, eq.(2.42), we see that the change in v^i is given by

$$\begin{aligned}
[\delta v]^i &= \frac{1}{2}\left[\psi^\dagger\right]^A h_{AB}\varepsilon_{rs}\left(\left[\gamma^i\right]^B{}_D \left[\sigma^{rs}\right]^D{}_E - \left[\sigma^{rs}\right]^B{}_C \left[\gamma^i\right]^C{}_E\right)\left[\psi\right]^E \\
&= \frac{1}{2}\left[\psi^\dagger\right]^A h_{AB}\varepsilon_{rs}\left[\gamma^i, \sigma^{rs}\right]^B{}_E \left[\psi\right]^E
\end{aligned}$$

Computing the resulting commutator,

$$\begin{aligned}
\left[\gamma^i, \sigma^{rs}\right] &= \left[\gamma^i, \frac{1}{2}\left(\gamma^r\gamma^s - \eta^{rs}\right)\right] \\
&= \frac{1}{2}\left[\gamma^i, \gamma^r\gamma^s\right] \\
&= \frac{1}{2}\left(\gamma^i\gamma^r\gamma^s - \gamma^r\gamma^s\gamma^i\right) \\
&= \frac{1}{2}\left(\gamma^i\gamma^r\gamma^s + \gamma^r\gamma^i\gamma^s - 2\eta^{is}\gamma^r\right) \\
&= \eta^{ir}\gamma^s - \eta^{is}\gamma^r
\end{aligned}$$

we substitute into δv^i,

$$\begin{aligned}
[\delta v]^i &= \frac{1}{2}\left[\psi^\dagger\right]^A h_{AB}\varepsilon_{rs}\left[\gamma^i, \sigma^{rs}\right]^B{}_E \left[\psi\right]^E \\
&= \frac{1}{2}\left[\psi^\dagger\right]^A h_{AB}\varepsilon_{rs}\left(\eta^{ir}\left[\gamma^s\right]^B{}_E - \eta^{is}\left[\gamma^r\right]^B{}_E\right)\left[\psi\right]^E \\
&= \frac{1}{2}\eta^{ir}\varepsilon_{rs}\left[\psi^\dagger\right]^A h_{AB}\left[\gamma^s\right]^B{}_E \left[\psi\right]^E - \frac{1}{2}\eta^{is}\varepsilon_{rs}\left[\psi^\dagger\right]^A h_{AB}\left[\gamma^r\right]^B{}_E \left[\psi\right]^E \\
&= \frac{1}{2}\left(\eta^{ir}\varepsilon_{rs}v^s - \eta^{is}\varepsilon_{rs}v^r\right) \\
&= \varepsilon^i{}_s v^s
\end{aligned}$$

But $\varepsilon^i{}_s$ is just an arbitrary antisymmetric matrix, ε_{is}, with one index raised using the $O(p,q)$ metric, η^{is}, and is therefore an infinitesimal Lorentz transformation. This means that the real n-vector v^i defined from a spinor according to Eq.(2.43) transforms correctly under an infinitesimal $SO(p,q)$ transformation. Since we may exponentiate the infinitesimal transformation to get any finite one, v^i is an $SO(p,q)$ vector.

Now we see why ψ is a spinor. If we think of v^i as a bi-spinor,

$$[v]^i \rightarrow \left[v^i\gamma_i\right]^A{}_B$$

then we can write the infinitesimal transformation laws as

$$\begin{aligned}
[\delta\psi]^A &= \frac{1}{2}\varepsilon_{uv}\left[\sigma^{uv}\right]^A{}_B\,[\delta\psi]^B \\
\left[\delta v_i\gamma^i\right]^A{}_B &= \frac{1}{2}\varepsilon_{rs}\left[v_i\gamma^i\right]^B{}_D\left[\sigma^{rs}\right]^D{}_E - \frac{1}{2}\varepsilon_{rs}\left[\sigma^{rs}\right]^B{}_C\left[v_i\gamma^i\right]^C{}_E
\end{aligned}$$

The real vector transforms as a $\begin{pmatrix}1\\1\end{pmatrix}$ tensor under $\mathrm{Spin}(p,q)$, while a spinor transforms as a $\begin{pmatrix}1\\0\end{pmatrix}$ tensor – a vector – under $\mathrm{Spin}(p,q)$. If we rotate ψ by an angle $\frac{\varphi}{2}$, the same transformation will rotate v^i by φ. This is the characteristic property of a spinor.

To find the number of components ψ^A has in general, we find the minimum size for the gamma matrices. We can do this by finding out how many independent matrices we can build from the gamma matrices. We can always remove symmetric parts of products of gamma matrices, but the antisymmetric parts remain independent. Let

$$\Gamma^{ij\ldots k} = \gamma^{[i}\gamma^j\ldots\gamma^{k]} \tag{2.44}$$

where the bracket on the indices means to take the antisymmetric part. If there are n distinct γ^i, there will be $\binom{n}{m}$ different matrices $\Gamma^{i_1\cdots i_m}$ having m indices. Each of these can be shown to be independent, for all m, so we have $\sum_{m=0}^{n}\binom{n}{m} = 2^n$ independent matrices constructible from the γ^i. The linear combinations of these 2^n matrices form the *Clifford algebra* associated with $O(p,q)$. The minimum dimension having 2^n independent matrices is $2^{n/2}$ (or $2^{(n+1)/2}$ if n is odd) since a $2^{n/2}\times 2^{n/2}$ matrix has 2^n components. It is not too difficult to show that a satisfactory set of matrices of this dimension always exists. Therefore, spinors in n dimensions will have $2^{n/2}$ components (n even), and this agrees with the the 4-component spinors found by Dirac.

We still need to know what the metric h_{AB} is for the Dirac case. It must satisfy the invariance condition of eq.(2.42), which in 4 dimensions reduces to

$$\begin{aligned}
0 &= h\sigma^{0i} + \left(\sigma^{0i}\right)^\dagger h = h\sigma^{0i} + \sigma^{0i}h \\
0 &= h\sigma^{ij} + \left(\sigma^{ij}\right)^\dagger h = h\sigma^{ij} - \sigma^{ij}h
\end{aligned}$$

These relations are satisfied if we define the metric to be

$$h_{AB} \equiv \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{2.45}$$

The choice of h_{AB} as the metric is fixed by its rotational invariance under the $\sigma^{\mu\nu}$ which we easily check as follows. If we momentarily ignore index positions, we see that h_{AB} has the same form as γ^0, and we can use the properties of γ^0 to compute its effects. Thus for the $0i$ components,

$$\begin{aligned}
h\sigma^{0i} + \sigma^{0i}h &\sim \gamma^0\sigma^{0i} + \sigma^{0i}\gamma^0 \\
&= \frac{1}{4}\left(\gamma^0\gamma^0\gamma^i - \gamma^0\gamma^i\gamma^0 + \gamma^0\gamma^i\gamma^0 - \gamma^i\gamma^0\gamma^0\right) \\
&= \frac{1}{4}\left(\gamma^i + \gamma^i\gamma^0\gamma^0 - \gamma^i\gamma^0\gamma^0 - \gamma^i\right) \\
&= 0
\end{aligned}$$

while for the ij components,

$$\begin{aligned}
h\sigma^{ij} - \sigma^{ij}h &\sim \gamma^0\sigma^{ij} - \sigma^{ij}\gamma^0 \\
&= \frac{1}{4}\left(\gamma^0\gamma^i\gamma^j - \gamma^0\gamma^j\gamma^i - \gamma^i\gamma^j\gamma^0 + \gamma^j\gamma^i\gamma^0\right) \\
&= \frac{1}{2}\left(\gamma^0\gamma^i\gamma^j - \gamma^0\gamma^j\gamma^i - \gamma^0\gamma^i\gamma^j + \gamma^0\gamma^j\gamma^i\right) \\
&= 0
\end{aligned}$$

as required. Therefore, Dirac spinors have the Lorentz-invariant inner product

$$\langle\psi,\psi\rangle = \left[\psi^\dagger\right]^A h_{AB}\psi^B$$

It is convenient to define $\bar{\psi} \equiv \psi^\dagger h$, with components

$$\bar{\psi}_B \equiv \left[\psi^\dagger\right]^A h_{AB} \tag{2.46}$$

Then we may write the inner product simply as

$$\langle\psi,\psi\rangle = \bar{\psi}\psi$$

The simplicity of using γ^0 to compute inner products has led many authors of field theory texts to actually write γ^0 for h, as in $\bar{\psi}\psi = \psi^\dagger\gamma^0\psi$, but the difference between a metric and a transformation is important. Indeed, the index structure is clearly wrong in the latter expression – we would have. $\psi^\dagger\gamma^0\psi = \left[\psi^\dagger\right]^A \left[\gamma^0\right]^A_{\ \ B} [\psi]^B$.

2.8 Working with the gamma matrices and the Dirac equation

Returning to our original goal, we now have the Dirac equation, Eq.(2.36) and the Clifford algebra, Eq.(2.37). Because spinors rotate by $\frac{\varphi}{2}$ when a vector is rotated by φ, we say spinors have spin $\frac{1}{2}$. Eq.(2.36) is therefore the field equation for a spin-$\frac{1}{2}$ field. Since we have an invariant inner product, we can write a Lorentz invariant action as

$$S = \int d^4x\ \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi \tag{2.47}$$

The action is to be varied with respect to ψ and $\bar{\psi}$ independently

$$0 = \delta S = \int d^4x\ \left(\delta\bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi + \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\delta\psi\right)$$

The $\bar{\psi}$ variation immediately yields the Dirac equation, $\left(i\gamma^\mu\partial_\mu - m\right)\psi = 0$ while the $\delta\psi$ required integration by parts:

$$\begin{aligned} 0 &= \int d^4x\ \bar{\psi}\left(i\gamma^\mu\partial_\mu\delta\psi - m\delta\psi\right) \\ &= \int d^4x\ \left(-i\partial_\mu\bar{\psi}\gamma^\mu\delta\psi - \bar{\psi}m\right)\delta\psi \end{aligned}$$

Thus

$$i\partial_\mu\bar{\psi}\gamma^\mu + m\bar{\psi} = 0 \tag{2.48}$$

which is sometimes written as

$$\bar{\psi}\left(i\gamma^\mu\overleftarrow{\partial}_\mu + m\right) = 0 \tag{2.49}$$

2.8.1 Some further properties of the gamma matrices

In four dimensions, there are 16 independent matrices that we can construct from the Dirac matrices. We have already encountered eleven of them:

$$\mathbf{1}, \gamma^\mu, \sigma^{\mu\nu}$$

The remaining five are most readily expressed in terms of

$$\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{2.50}$$

Exercise: Prove that γ_5 is hermitian.

Exercise: Prove that $\{\gamma_5, \gamma^\mu\} = 0$.

Exercise: Prove that $\gamma_5\gamma_5 = \mathbf{1}$.

Then the remaining five matrices may be taken as

$$\gamma_5, \gamma_5\gamma^\mu$$

Any complex 4×4 matrix can be expressed as linear combination of these 16 matrices, $\Gamma^a = \{\mathbf{1}, \gamma^\mu, \sigma^{\mu\nu}, \gamma_5, \gamma_5\gamma^\mu\}$. We will need several other properties of these matrices. First, if we contract the product of pair of gammas, we get 4:

$$\gamma^\mu\gamma_\mu = \eta_{\mu\nu}\gamma^\mu\gamma^\nu = \frac{1}{2}\eta_{\mu\nu}\{\gamma^\mu, \gamma^\nu\} = \eta_{\mu\nu}\eta^{\mu\nu} = 4$$

We will need various traces. For any product of an odd number of gamma matrices we have

$$\begin{aligned} tr\left(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_{2k+1}}\right) &= tr\left((\gamma_5)^2\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_{2k+1}}\right) \\ &= (-1)^{2k+1}\ tr\left(\gamma_5\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_{2k+1}}\gamma_5\right) \end{aligned}$$

using the fact that γ_5 squares to $\mathbf{1}$ and commutes with any of the γ^μ. Now, using the cyclic property of the trace

$$tr\left(A\ldots BC\right) = tr\left(CA\ldots B\right)$$

we cycle γ_5 back to the front:

$$\begin{aligned} tr\left(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_{2k+1}}\right) &= (-1)^{2k+1}\ tr\left(\gamma_5\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_{2k+1}}\gamma_5\right) \\ &= (-1)^{2k+1}\ tr\left(\gamma_5\gamma_5\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_{2k+1}}\right) \\ &= -\ tr\left(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_{2k+1}}\right) \\ &= 0 \end{aligned}$$

Thus, the trace of the product of any *odd* number of gamma matrices vanishes.

Traces of even numbers are trickier. For two:

$$\begin{aligned} tr\left(\gamma^\mu\gamma^\nu\right) &= tr\left(-\gamma^\nu\gamma^\mu + 2\eta^{\mu\nu}\mathbf{1}\right) \\ &= -tr\left(\gamma^\nu\gamma^\mu\right) + 2\eta^{\mu\nu}tr\ \mathbf{1} \end{aligned}$$

or, since $tr\mathbf{1} = 4$,

$$tr\left(\gamma^\mu\gamma^\nu\right) = 4\eta^{\mu\nu} \tag{2.51}$$

Exercise: Prove that

$$tr\left(\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu\right) = 4\left(\eta^{\alpha\beta}\eta^{\mu\nu} - \eta^{\alpha\mu}\eta^{\beta\nu} + \eta^{\alpha\nu}\eta^{\beta\mu}\right)$$

Exercise: Prove that

$$\gamma^\mu\gamma^\alpha\gamma_\mu = -2\gamma^\alpha$$

and

$$\gamma^\mu\gamma^\alpha\gamma^\beta\gamma_\mu = 4\eta^{\alpha\beta}$$

2.8.2 Casimir Operators

For any Lie algebra, $\mathcal{G}$, with generators G_a and commutators

$$[G_a, G_b] = c_{ab}{}^c G_c$$

we can consider composite operators found by multiplying together two or more generators,

$$G_1G_2, G_3G_9G_{17}, \ldots$$

and taking linear combinations,

$$A = \alpha G_1 G_2 + \beta G_3 G_9 G_{17} + \dots$$

The set of all such linear combinations of products is called the *free algebra* of $\mathcal{G}$. Among the elements of the free algebra are a very few special cases called *Casmir operators*, which have the special property of commuting with all of the generators. For example; the generators J_i of $SO(3)$ may be combined into the combination

$$R = \delta^{ij} J_i J_j = \sum (J_i)^2 \tag{2.52}$$

We can compute

$$\begin{aligned} [J_i, R] &= \left[J_i, \sum (J_j)^2\right] \\ &= J_j [J_i, J_j] + [J_i, J_j] J_j \\ &= J_j \varepsilon_{ijk} J_k + \varepsilon_{ijk} J_k J_j \\ &= \varepsilon_{ijk} (J_j J_k + J_k J_j) \\ &= 0 \end{aligned}$$

where, in the last step, we used the fact that ε_{ijk} is antisymmetric on jk, while the expression $J_j J_k + J_k J_k$ is explicitly symmetric. Since R commutes with the three J_i, it also commutes with all products of them. R is therefore a Casimir operator for $O(3)$. For this reason, Casimir operators become extremely important in quantum physics. Because the symmetries of our system are group symmetries, the set of all Casimir operators gives us a list of the conserved quantities. Generically, elements of a Lie group take us from one set of fields to a physically equivalent set *covariantly*. Since the Casimir operators are left *invariant*, we can use eigenvalues of the Casimir operators to classify the possible distinct physical states of the system.

Let's look at the Casimir operators that are most imporant for particle physics – those of the Poincaré group. The Poincaré group is the set of transformations leaving the *infinitesimal line element*

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \tag{2.53}$$

invariant. It clearly includes Lorentz transformations, $[dx']^\alpha = \Lambda^\alpha{}_\beta dx^\beta$ but now also includes translations:

$$\begin{aligned} [x']^\alpha &= x^\alpha + a^\alpha \\ &\Rightarrow [dx']^\alpha = dx^\beta \end{aligned}$$

Since there are 4 translations and 6 Lorentz transformations, there are a total of 10 Poincaré symmetries. There are several ways to write a set of generators for these transformations. One common one is to let

$$\begin{aligned} M^\alpha{}_\beta &= x^\alpha \partial_\beta - x_\beta \partial^\alpha \\ P_\alpha &= \partial_\alpha \end{aligned} \tag{2.54}$$

Then it is easy to show that

$$\begin{aligned} \left[M^\alpha{}_\beta, M^\mu{}_\nu\right] &= [x^\alpha \partial_\beta - x_\beta \partial^\alpha, x^\mu \partial_\nu - x_\nu \partial^\mu] \\ &= x^\alpha \partial_\beta (x^\mu \partial_\nu - x_\nu \partial^\mu) - x_\beta \partial^\alpha (x^\mu \partial_\nu - x_\nu \partial^\mu) \\ &\quad - x^\mu \partial_\nu (x^\alpha \partial_\beta - x_\beta \partial^\alpha) + x_\nu \partial^\mu (x^\alpha \partial_\beta - x_\beta \partial^\alpha) \\ &= x^\alpha \delta^\mu_\beta \partial_\nu - x^\alpha \eta_{\beta\nu} \partial^\mu - x_\beta \eta^{\alpha\mu} \partial_\nu + x_\beta \delta^\alpha_\nu \partial^\mu \\ &\quad - x^\mu \delta^\alpha_\nu \partial_\beta + x^\mu \eta_{\nu\beta} \partial^\alpha + x_\nu \eta^{\mu\alpha} \partial_\beta - x_\nu \delta^\mu_\beta \partial^\alpha \\ &= \delta^\mu_\beta M^\alpha{}_\nu - \eta_{\beta\nu} M^{\alpha\mu} - \eta^{\alpha\mu} M_{\beta\nu} + \delta^\alpha_\nu M_\beta{}^\mu \end{aligned} \tag{2.55}$$

To compute these, we imagine the derivatives acting on a function to the right of the commutator, $\left[M^\alpha{}_\beta, M^\mu{}_\nu\right] f(x)$. Then all of the derivatives of f cancel when we antisymmetrize. With suitable adjustments of the index positions, we see the result above is equivalent to the Lorentz ($o(3,1)$) case of eq.(2.14). Two similar but shorter calculations show that

$$\left[M^{\alpha}{}_{\beta}, P_{\nu}\right] = \eta_{\nu\beta}P^{\alpha} - \delta^{\alpha}_{\nu}P_{\beta} \tag{2.56}$$

$$\left[P_{\alpha}, P_{\beta}\right] = 0 \tag{2.57}$$

Eqs.(2.55-2.57) comprise the Lie algebra of the Poincaré group.

Exercise: Prove eq.(2.56) and eq.(2.57) using eqs.(2.54).

Now we can write the Casimir operators of the Poincaré group. There are two:

$$P^2 = \eta^{\alpha\beta}P_{\alpha}P_{\beta} \tag{2.58}$$

$$W^2 = \eta_{\alpha\beta}W^{\alpha}W^{\beta} \tag{2.59}$$

where

$$W^{\mu} \equiv \frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}P_{\nu}M_{\alpha\beta} \tag{2.60}$$

and $\varepsilon^{\mu\nu\alpha\beta}$ is the spacetime Levi-Civita tensor. To see what these correspond to, recall from our discussion of Noether currents that the conservation of 4-momentum is associated with translation invariance, and P_{α} is the generator of translations. In fact, $P_{\alpha} = i\partial_{\alpha}$, the Hermitian form of the translation generator, is the usual energy-momentum operator of quantum mechanics. We directly interpret eigenvectors of P_{α} as energy and momentum. Thus, we expect that eigenvalues of P^2 will be the mass, $p_{\alpha}p^{\alpha} = m^2$.

Similarly, W^2 is built from the rotation generators. To see this, notice that we expect the momentum, p^{α}, to be a timelike vector. This means that there exists a frame of reference in which $p^{\alpha} = (mc, 0, 0, 0)$. In this frame, W^{α} becomes

$$\begin{aligned} W^{\mu} &= \frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}P_{\nu}M_{\alpha\beta} \\ &= \frac{1}{2}mc\varepsilon^{\mu 0\alpha\beta}M_{\alpha\beta} \end{aligned}$$

Therefore, $W^0 = 0$, and using $\varepsilon^{k0ij} = -\varepsilon^{k0ij} = \varepsilon^{ijk}$, for the spatial components,

$$\begin{aligned} W^{k} &= \frac{1}{2}mc\varepsilon^{ijk}M_{ij} \\ &= mcJ^{k} \end{aligned}$$

Since m is separately conserved, this shows that the magnitude of the angular momentum J^2 is also conserved.

Exercise: Using the Lie algebra of the Poincaré group, eqs.(2.55-2.57), prove that P^2 and W^2 commute with $M_{\alpha\beta}$ and P_{α}. (Warning! The proof for W^2 is a bit tricky!) Notice that the proof requires only the Lie algebra relations for the Poincaré group, and *not* the specific representation of the operators given in eqs.(2.54).

Since the Casimir operators of the Poincaré group correspond to mass and spin, we will be able to classify states of quantum fields by mass and spin. We will extend this list when we introduce additional symmetry groups.

Chapter 3

The classification of particles

Using the Casimir operators we can give a complete classification of fundamental particles, that is, representations of the Poincaré group. We may specify particle states by their eigenvalues under the Casimir operators. Notice that since Casimir operators are left invariant Since these correspond to mass and spin, we write

$$\begin{aligned} P^2 |M, j\rangle &= M^2 |M, j\rangle \\ J^2 |M, j\rangle &= j(j+1)\hbar^2 |M, j\rangle \end{aligned}$$

Depending on the model, there may be a continuous or discrete spectrum of particle masses.

We now may find a basis for states of angular momentum, that is, all finite-dimensional representations for the three operators $\hat{J}_i$. All results follow from the fundamental commutation relation for hermitian rotational generators,

$$\left[\hat{J}_i, \hat{J}_j\right] = i\hbar\varepsilon_{ijk}\hat{J}_k$$

where i, j, k each take values $1, 2, 3$ and we sum on k. These results are often developed for classification of spin states in quantum mechanics, but we will be interested in their use in classifying fundamental particles, not multi-electron atoms.

As we show in the remainder of this Chapter, the eigenvalues, j, for spin include every non-negative integer and half-integer value, $0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$, with spin 0 fields called *scalars*, spin-$\frac{1}{2}$ fields *spinors*, and spin-1 fields *vectors*. Within the Standard Model, the Higgs particle is the only scalar. Each of the quarks and leptons is a spinor particle, while the gauge bosons, that is, the photon, $W^{\pm}$, Z^0, and the gluons, are spin-1. Higher spin fundamental particles have not been observed, but quantization of gravity is expected to include a spin-2 quantum called the graviton, and supersymmetry predicts a fermionic, spin-$\frac{3}{2}$ partner to the graviton called a gravitino.

3.1 A maximal set of commuting observables

To begin, we ask how many mutually commuting operators we can build from $\hat{J}_i$. We can diagonalize any one of $\hat{J}_1, \hat{J}_2, \hat{J}_3$, but since none commute with either or the others, we cannot diagonalize more than one. We choose $\hat{J}_3$ diagonal. There is one further commuting combination – since rotations preserve lengths, the length of $\hat{J}_i$ itself is preserved by rotations,

$$\begin{aligned}
\left[\hat{J}_i, \hat{\mathbf{J}}^2\right] &= \left[\hat{J}_i, \hat{J}_k\hat{J}_k\right] \\
&= \hat{J}_k\left[\hat{J}_i, \hat{J}_k\right] + \left[\hat{J}_i, \hat{J}_k\right]\hat{J}_k \\
&= \hat{J}_k i\hbar\varepsilon_{ikm}\hat{J}_m + i\hbar\varepsilon_{ikm}\hat{J}_m\hat{J}_k \\
&= i\hbar\varepsilon_{ikm}\left(\hat{J}_k\hat{J}_m + \hat{J}_m\hat{J}_k\right) \\
&= 0
\end{aligned}$$

where the last step follows because $\hat{J}_k\hat{J}_m + \hat{J}_m\hat{J}_k$ is symmetric in mk while ε_{ikm} is antisymmetric. In particular, we have

$$\left[\hat{J}_3, \hat{\mathbf{J}}^2\right] = 0$$

so these may be simultaneously diagonalized. Since we already know that the Pauli matrices give a 2-dimensional example for the generators, there cannot be more than two indepedent diagonal combinations.

Having found a maximal set of commuting observables, we may use their eigenvalues to label their simultaneous eigenkets. Let

$$\begin{aligned}
\hat{\mathbf{J}}^2\left|\alpha,\beta\right\rangle &= \alpha^2\hbar^2\left|\alpha,\beta\right\rangle \\
\hat{J}_3\left|\alpha,\beta\right\rangle &= \beta\hbar\left|\alpha,\beta\right\rangle
\end{aligned}$$

We take these kets to be orthonormal and seek all allowed values of the real eigenvalues, α, β.

3.2 Raising and lowering operators

We combine the remaing two generators in the useful combinations,

$$\hat{J}_\pm \equiv \hat{J}_1 \pm i\hat{J}_2$$

where we note that $\hat{J}_+^\dagger = \hat{J}_-$. These satisfy:

$$\begin{aligned}
\left[\hat{J}_+, \hat{J}_-\right] &= \left[\hat{J}_1 + i\hat{J}_2, \hat{J}_1 - i\hat{J}_2\right] \\
&= -i\left[\hat{J}_1, \hat{J}_2\right] + i\left[\hat{J}_2, \hat{J}_1\right] \\
&= 2\hbar\hat{J}_3
\end{aligned}$$

and

$$\begin{aligned}
\left[\hat{J}_3, \hat{J}_\pm\right] &= \left[\hat{J}_3, \hat{J}_1 \pm i\hat{J}_2\right] \\
&= \left[\hat{J}_3, \hat{J}_1\right] \pm i\left[\hat{J}_3, \hat{J}_2\right] \\
&= i\hbar\hat{J}_2 \pm i\left(-i\hbar\hat{J}_1\right) \\
&= \pm\hbar\hat{J}_\pm
\end{aligned}$$

as well as commuting with the length,

$$\left[\hat{J}_\pm, \hat{\mathbf{J}}^2\right] = 0$$

Consider the actions of $\hat{\mathbf{J}}^2$ and $\hat{J}_3$ on the state $\hat{J}_+\left|\alpha,\beta\right\rangle$,

$$\begin{aligned}
\hat{\mathbf{J}}^2\hat{J}_+\left|\alpha,\beta\right\rangle &= \hat{J}_+\hat{\mathbf{J}}^2\left|\alpha,\beta\right\rangle \\
&= \alpha^2\hbar^2\hat{J}_+\left|\alpha,\beta\right\rangle
\end{aligned}$$

so this state is also an eigenstate of $\hat{\mathbf{J}}^2$ with the eigenvalue α, while

$$\begin{aligned}\hat{J}_3\hat{J}_+ |\alpha,\beta\rangle &= \left(\left[\hat{J}_3,\hat{J}_+\right] + \hat{J}_+\hat{J}_3\right)|\alpha,\beta\rangle \\ &= \hbar\hat{J}_+|\alpha,\beta\rangle + \hat{J}_+\hat{J}_3|\alpha,\beta\rangle \\ &= (\beta+1)\,\hbar\hat{J}_+|\alpha,\beta\rangle\end{aligned}$$

We once again have an eigenstate, but the eigenvalue β has increased by $\hbar$. Up to an overall constant λ we have

$$\hat{J}_+|\alpha,\beta\rangle = \lambda|\alpha,\beta+1\rangle$$

Exercise: Show that $\hat{J}_+|\alpha,\beta\rangle = \lambda|\alpha,\beta-1\rangle$ for some constant λ.

3.3 Limits on eigenvalues

3.3.1 Inequalities on the eigenvalues

Products of $\hat{J}_+$ and $\hat{J}_-$ may be expressed in term of our diagonal operators. For the product $\hat{J}_+\hat{J}_-$:

$$\begin{aligned}\hat{J}_+\hat{J}_- &= \left(\hat{J}_1 + i\hat{J}_2\right)\left(\hat{J}_1 - i\hat{J}_2\right) \\ &= \hat{J}_1^2 + \hat{J}_2^2 - i\hat{J}_1\hat{J}_2 + i\hat{J}_2\hat{J}_1 \\ &= \hat{\mathbf{J}}^2 - \hat{J}_3^2 - i\left[\hat{J}_1,\hat{J}_2\right] \\ &= \hat{\mathbf{J}}^2 - \hat{J}_3^2 + \hbar\hat{J}_3\end{aligned}$$

Exercise: Show that $\hat{J}_-\hat{J}_+ = \hat{\mathbf{J}}^2 - \hat{J}_3^2 - \hbar\hat{J}_3$

Since $\hat{J}_+^\dagger = \hat{J}_-$ and $\hat{J}_-^\dagger = \hat{J}_+$ we have inequalities from the norms of $\hat{J}_+|\alpha,\beta\rangle$ and $\hat{J}_-|\alpha,\beta\rangle$:

$$\begin{aligned}\left[\langle\alpha,\beta|\,\hat{J}_+^\dagger\right]\left[\hat{J}_+|\alpha,\beta\rangle\right] = \langle\alpha,\beta|\,\hat{J}_-\hat{J}_+|\alpha,\beta\rangle &\geq 0 \\ \left[\langle\alpha,\beta|\,\hat{J}_-^\dagger\right]\left[\hat{J}_-|\alpha,\beta\rangle\right] = \langle\alpha,\beta|\,\hat{J}_=\hat{J}_-|\alpha,\beta\rangle &\geq 0\end{aligned}$$

These give, respectively,

$$\begin{aligned}0 &\leq \langle\alpha,\beta|\,\hat{J}_-\hat{J}_+|\alpha,\beta\rangle \\ &= \langle\alpha,\beta|\left(\hat{\mathbf{J}}^2 - \hat{J}_3^2 - \hbar\hat{J}_3\right)|\alpha,\beta\rangle \\ &= \left(\alpha^2 - \beta^2 - \beta\right)\hbar^2\langle\alpha,\beta\,|\alpha,\beta\rangle \\ &= \left(\alpha^2 - \beta^2 - \beta\right)\hbar^2\end{aligned}$$

and

$$\begin{aligned}0 &\leq \langle\alpha,\beta|\,\hat{J}_+\hat{J}_-|\alpha,\beta\rangle \\ &= \langle\alpha,\beta|\left(\hat{\mathbf{J}}^2 - \hat{J}_3^2 + \hbar\hat{J}_3\right)|\alpha,\beta\rangle \\ &= \left(\alpha^2 - \beta^2 + \beta\right)\hbar^2\end{aligned}$$

so two distinct inequalities must hold:

$$\beta^2 + \beta \leq \alpha^2 \tag{3.1}$$
$$\beta^2 - \beta \leq \alpha^2 \tag{3.2}$$

3.3.2 The eigenvalues

Now, just as we did for the simple harmonic oscillator, we start with any eigenstate and lower the eigenvalue k times,

$$\hat{J}_3 \left(\hat{J}_-\right)^k |\alpha, \beta\rangle = \lambda_{\beta-k} (\beta - k) \hbar |\alpha, \beta - k\rangle$$

for some normalization constant, $\lambda_{\beta-k}$. However, this series must terminate, since eq.(3.1) for the state $|\alpha, \beta - k\rangle$ leads to

$$\begin{aligned} (\beta-k)^2 + (\beta-k) &\leq \alpha^2 \\ k^2 - 2\beta k - k + \beta^2 + \beta &\leq \alpha^2 \end{aligned}$$

Regardless of the value of α and β, there is some value of k which is sufficiently large to violate this inequality. Therefore, there must exist some β_{min} such that

$$\hat{J}_- |\alpha, \beta_{min}\rangle = 0$$

Since $\beta = 0$ satisfies both inequalities we must have $\beta_{min} < 0$, and therefore

$$\beta_{min}^2 - \beta_{min} \leq \alpha^2$$

gives the strongest constraint on β_{min}.

Now we apply $\hat{J}_+$ to $|\alpha, \beta_{min}\rangle$ to produce eigenkets of larger and larger β,

$$\hat{J}_+^k |\alpha, \beta_{min}\rangle = \lambda_{\beta_{min}+k} |\alpha, \beta_{min} + k\rangle$$

Once again we eventually reach a value of k which violates one of the inequalities, so there exists some positive, maximum β_{max}, satisfying both inequalities. The strongest constraint is

$$\beta_{max}^2 + \beta_{max} \leq \alpha^2$$

Notice that if $\beta_{min} = -\beta_{max} = -m$ then both inequalities give

$$m(m+1) \leq \alpha^2$$

Now acting on the highest state, $|\alpha, \beta_{max}\rangle$ with $\hat{J}_+$, or acting on the lowest state, $|\alpha, \beta_{min}\rangle$, with $\hat{J}_-$ must give zero

$$\begin{aligned} \hat{J}_+ |\alpha, \beta_{max}\rangle &= 0 \\ \hat{J}_- |\alpha, \beta_{min}\rangle &= 0 \end{aligned}$$

and therefore, acting on the first with $\hat{J}_-$ and the second with $\hat{J}_+$

$$\begin{aligned} 0 &= \hat{J}_- \hat{J}_+ |\alpha, \beta_{max}\rangle \\ &= \left(\hat{\mathbf{J}}^2 - \hat{J}_3^2 - \hbar \hat{J}_3\right) |\alpha, \beta_{max}\rangle \\ &= \left(\alpha^2 - \beta_{max}^2 - \beta_{max}\right) \hbar^2 |\alpha, \beta_{max}\rangle \end{aligned}$$

and

$$\begin{aligned} 0 &= \hat{J}_+ \hat{J}_- |\alpha, \beta_{min}\rangle \\ &= \left(\hat{\mathbf{J}}^2 - \hat{J}_3^2 + \hbar \hat{J}_3\right) |\alpha, \beta_{min}\rangle \\ &= \left(\alpha^2 - \beta_{min}^2 + \beta_{min}\right) \hbar^2 |\alpha, \beta_{min}\rangle \end{aligned}$$

giving us two equalities for the maximum and minimum values:

$$\begin{aligned}
\alpha^2 &= \beta_{max}^2 + \beta_{max} \\
\alpha^2 &= \beta_{min}^2 - \beta_{min}
\end{aligned}$$

We also know that $\beta_{max} - \beta_{min} = k$ for some non-negative integer, k. Setting $\beta_{max} = \beta_{min} + k$ and equating the two expressions,

$$\begin{aligned}
\beta_{min}^2 - \beta_{min} &= \beta_{max}^2 + \beta_{max} \\
&= (\beta_{min} + k)^2 + (\beta_{min} + k) \\
\beta_{min}^2 - \beta_{min} &= \beta_{min}^2 + 2\beta_{min}k + k^2 + \beta_{min} + k \\
0 &= (k+1)\, 2\beta_{min} + k\,(k+1) \\
0 &= 2\beta_{min} + k \\
\beta_{min} &= -\frac{k}{2}
\end{aligned}$$

so that β_{min} is some negative integer or half-integer we will call $-j$:

$$\beta_{min} = -j \in \left\{0, -\frac{1}{2}, -1, -\frac{3}{2}, -2, \ldots\right\}$$

The maximum value $\beta_{max} = \beta_{min} + k = +\frac{k}{2} = +j$, and the remaining eigenvalue is

$$\begin{aligned}
\alpha^2 &= \frac{k}{2}\left(\frac{k}{2} + 1\right) \\
&= j\,(j+1)
\end{aligned}$$

The labeling of our states is complete. Letting $\beta = m$, the complete set of possible states for any fixed half-integer j is given by the $2j+1$ states,

$$|\alpha, \beta\rangle = \{|j, m\rangle \mid m = -j, -j+1, \ldots, j+1, j\}$$

and we have one such set for every choice of $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$. The eigenvalues of these states are given by

$$\hat{\mathbf{J}}^2 |j, m\rangle = j\,(j+1)\,\hbar^2 |j, m\rangle \tag{3.3}$$

$$\hat{J}_3 |j, m\rangle = m\hbar\, |j, m\rangle \tag{3.4}$$

These states will be referred to as "spin-j" representations.

3.3.3 Normalization of raising and lowering

We define these eigenstates to be normalized, and since they are nondegenerate, they are orthonormal,

$$\langle j_1, m_1 | j_2, m_2\rangle = \delta_{j_1 j_2}\delta_{m_1 m_2}$$

However, we need to know the effect of the raising and lowering operators. We already know that

$$\hat{J}_\pm |j, m\rangle = \lambda_{m\pm1} |j, m \pm 1\rangle$$

for some constants $\lambda_{m\pm1}$. To find $\lambda_{m\pm1}$, look again at the norm

$$\begin{aligned}
\langle j, m| \hat{J}_- \hat{J}_+ |j, m\rangle &= \langle j, m| \left(\hat{\mathbf{J}}^2 - \hat{J}_3^2 - \hbar\hat{J}_3\right) |j, m\rangle \\
|\lambda_{m+1}|^2 &= (j\,(j+1) - m\,(m+1))\,\hbar^2 \\
\lambda_{m+1} &= \sqrt{j\,(j+1) - m\,(m+1)}\hbar
\end{aligned}$$

where we choose the phase so that λ_{m+1} is real. For $\hat{J}_-$ we have

$$\begin{aligned}
\langle j,m| \hat{J}_+\hat{J}_- |j,m\rangle &= \langle j,m| \left(\hat{\mathbf{J}}^2 - \hat{J}_3^2 + \hbar\hat{J}_3\right) |j,m\rangle \\
|\lambda_{m-1}|^2 &= \left(j\,(j+1) - m\,(m-1)\right)\hbar^2 \\
\lambda_{m-1} &= \sqrt{j\,(j+1) - m\,(m-1)}\hbar
\end{aligned}$$

Therefore, the action of the raising and lowering operators is

$$\hat{J}_\pm |j,m\rangle = \sqrt{j\,(j+1) - m\,(m\pm 1)}\hbar\, |j,m\pm 1\rangle \tag{3.5}$$

3.4 Examples of representations

3.4.1 Spin 0

For $j = 0$, we only have the single allowed value $m = 0$ and there is only one state,

$$|j,m\rangle = |0,0\rangle$$

These are scalars. We may find the expectation value of any component of angular momentum using

$$\begin{aligned}
J_1 &= \frac{1}{2}\left(\hat{J}_+ + \hat{J}_-\right) \\
J_2 &= \frac{1}{2i}\left(\hat{J}_+ - \hat{J}_-\right)
\end{aligned}$$

Since $m = 0 = \beta_{min} = \beta_{max}$, both $\hat{J}_+$ and $\hat{J}_-$ must give zero:

$$\hat{J}_\pm |0,0\rangle = 0$$

and we have

$$\begin{aligned}
\hat{J}_1 |0,0\rangle &= 0 \\
\hat{J}_2 |0,0\rangle &= 0 \\
\hat{J}_3 |0,0\rangle &= 0
\end{aligned}$$

so the action of all generators is zero. Furthermore,

$$\begin{aligned}
\langle 0,0| \hat{J}_x |0,0\rangle &= 0 \\
\langle 0,0| \hat{J}_y |0,0\rangle &= 0 \\
\langle 0,0| \hat{J}_z |0,0\rangle &= 0
\end{aligned}$$

so every component of angular momentum has zero expectation value.

The effect of a general infinitesimal rotation on a scalar state is given by

$$\begin{aligned}
\mathscr{D}(\mathbf{n},\varphi)\,|0,0\rangle &= \left(\hat{1} - \frac{i\varphi}{\hbar}\mathbf{n}\cdot\hat{\boldsymbol{J}}\right)|0,0\rangle \\
&= |0,0\rangle
\end{aligned}$$

so scalars are unaffected by any rotation.

3.4.2 Spin 1/2 particles

The simplest spin half particles are the wave functions of quantum mechanics, extended by including the spin up and spin down states. These are $j = \frac{1}{2}$ representations. For $j = \frac{1}{2}$ we have our familiar algebra of Pauli matrices, but we now have a more systematic labelling for the states. When we wish to be explicit about the value of j, we will write

$$\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$$

instead of $|\pm\rangle$. Notice that in all cases here we are taking $\hat{J}_3$ diagonal. We already know the expectation values of $\hat{J}_i$ in these states. For $\hat{\mathbf{J}}^2$ and $\hat{J}_\pm$ we have

$$\begin{aligned}
\hat{\mathbf{J}}^2 \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle &= \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2 \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle \\
&= \frac{3}{4}\hbar^2 \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle
\end{aligned}$$

and

$$\begin{aligned}
\hat{J}_+ \left|\frac{1}{2}, \frac{1}{2}\right\rangle &= 0 \\
\hat{J}_- \left|\frac{1}{2}, \frac{1}{2}\right\rangle &= \sqrt{j(j+1) - m(m-1)}\hbar \left|\frac{1}{2}, \frac{1}{2} - 1\right\rangle \\
&= \sqrt{\frac{1}{2}\left(\frac{3}{2}\right) - \frac{1}{2}\left(-\frac{1}{2}\right)}\hbar \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \\
&= \hbar \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \\
\hat{J}_- \left|\frac{1}{2}, -\frac{1}{2}\right\rangle &= \sqrt{\frac{1}{2}\left(\frac{3}{2}\right) - \left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)} \left|\frac{1}{2}, -\frac{1}{2} - 1\right\rangle \\
&= 0 \\
\hat{J}_+ \left|\frac{1}{2}, -\frac{1}{2}\right\rangle &= \sqrt{j(j+1) - m(m+1)}\hbar \left|\frac{1}{2}, \frac{1}{2}\right\rangle \\
&= \sqrt{\frac{3}{4} - \left(-\frac{1}{2}\right)\left(\frac{1}{2}\right)}\hbar \left|\frac{1}{2}, \frac{1}{2}\right\rangle \\
&= \hbar \left|\frac{1}{2}, \frac{1}{2}\right\rangle
\end{aligned}$$

The spin-$\frac{1}{2}$ states forem a 2-dimensional representation, so the generators are the Pauli matrices. Writing the raising and lowering operators in matrix notation,

$$\begin{aligned}
\hat{J}_+ &= \hat{J}_x + i\hat{J}_y \\
&= \frac{\hbar}{2}(\sigma_x + i\sigma_y) \\
&= \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\
\hat{J}_- &= \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}
\end{aligned}$$

so that

$$\begin{aligned}
\hat{J}_+ \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= 0 \\
\hat{J}_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \hbar \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\hat{J}_- \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= 0 \\
\hat{J}_+ \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \hbar \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}$$

Quite generally, the components of the raising and lowering operators are unit off-diagonal matrices.

In field theory, the Dirac spinors represent a pair spin-$\frac{1}{2}$ states, where the pair are the positive and negative energy solutions to the Dirac equation.

3.4.3 Spin 1

We have a total of three $j = 1$ states,

$$|j, m\rangle = |1, 1\rangle, |1, 0\rangle, |1, -1\rangle$$

related by

$$\begin{aligned}
\hat{J}_- |1, 1\rangle &= \sqrt{1(1+1) - 1(1-1)}\hbar |1, 1-1\rangle \\
&= \sqrt{2}\hbar |1, 0\rangle
\end{aligned}$$

and

$$\begin{aligned}
\hat{J}_- |1, 0\rangle &= \sqrt{1(1+1) - 0(0-1)}\hbar |1, 0-1\rangle \\
&= \sqrt{2}\hbar |1, -1\rangle
\end{aligned}$$

with similar relations for the raising operator. The eigenvalue of $\hat{\mathbf{J}}^2$ is $j(j+1)\hbar^2 = 2\hbar^2$.

The photon is the most familiar spin-1 particle, but there are also the weak intermediate vector bosons $\left(W^+, W^-, Z^0\right)$ and eight differnt gluons.

3.4.4 Spin 3/2

We have $2j + 1 = 4$ states,

$$|j, m\rangle = \left|\frac{3}{2}, \frac{3}{2}\right\rangle, \left|\frac{3}{2}, \frac{1}{2}\right\rangle, \left|\frac{3}{2}, -\frac{1}{2}\right\rangle, \left|\frac{3}{2}, -\frac{3}{2}\right\rangle$$

related by

$$
\begin{aligned}
\hat{J}_- \left|\frac{3}{2},\frac{3}{2}\right\rangle &= \sqrt{\frac{3}{2}\left(\frac{3}{2}+1\right)-\frac{3}{2}\left(\frac{3}{2}-1\right)}\hbar\left|\frac{3}{2},\frac{1}{2}\right\rangle \\
&= \sqrt{3}\hbar\left|\frac{3}{2},\frac{1}{2}\right\rangle \\
\hat{J}_- \left|\frac{3}{2},\frac{1}{2}\right\rangle &= \sqrt{\frac{3}{2}\left(\frac{3}{2}+1\right)-\frac{1}{2}\left(\frac{1}{2}-1\right)}\hbar\left|\frac{3}{2},-\frac{1}{2}\right\rangle \\
&= 2\hbar\left|\frac{3}{2},-\frac{1}{2}\right\rangle \\
\hat{J}_- \left|\frac{3}{2},-\frac{1}{2}\right\rangle &= \sqrt{\frac{3}{2}\left(\frac{3}{2}+1\right)-\left(-\frac{1}{2}\right)\left(-\frac{1}{2}-1\right)}\hbar\left|\frac{3}{2},-\frac{3}{2}\right\rangle \\
&= \sqrt{3}\hbar\left|\frac{3}{2},-\frac{3}{2}\right\rangle \\
\hat{J}_- \left|\frac{3}{2},-\frac{3}{2}\right\rangle &= 0
\end{aligned}
$$

with similar relations for the raising operator. The eigenvalue of $\hat{\mathbf{J}}^2$ is $j(j+1)\hbar^2 = \frac{15}{4}\hbar^2$.

We have not observed fundamental spin-$\frac{3}{2}$ particles, but supergravity and superstring theories include them as the fermionic superpartners to the graviton. They are called Rarita-Schwinger particles. The conjectured graviton, the quantum state of the gravitational field, should have spin-2.

3.4.5 Spin j

We summarize here the general results we have shown above.

For spin-j, where $j = \frac{n}{2}$ is any integer or half-integer there are $2j+1 = n+1$ orthonormal states labeled $|j,m\rangle$, where m ranges over all $2j+1$ values from $-j$ to $+j$. The actions of $\hat{\mathbf{J}}^2$, $\hat{J}_3$, $\hat{J}_\pm$ on these are given by

$$
\begin{aligned}
\hat{\mathbf{J}}^2 |j,m\rangle &= j(j+1)\hbar^2 |j,m\rangle \\
\hat{J}_3 |j,m\rangle &= m\hbar |j,m\rangle \\
\hat{J}_+ |j,m\rangle &= \sqrt{j(j+1)-m(m+1)}\hbar |j,m+1\rangle \\
\hat{J}_- |j,m\rangle &= \sqrt{j(j+1)-m(m-1)}\hbar |j,m-1\rangle
\end{aligned}
$$

while the actions of $\hat{J}_1, \hat{J}_2$ may be found using

$$
\begin{aligned}
\hat{J}_1 &= \frac{1}{2}\left(\hat{J}_+ + \hat{J}_-\right) \\
\hat{J}_2 &= \frac{1}{2i}\left(\hat{J}_+ - \hat{J}_-\right)
\end{aligned}
$$

There is a vector space of every positive integer dimension spanned by $|j,m\rangle$ for some j. Taken together, these give all of the irreducible representations of the 3-dimensional rotation group. This means that any tensor, i.e., any object that the 3-dim rotation group acts on multi-linearly and homogeneously, may be decomposed into some combination of the $|j,m\rangle$ vector space.

Exercise: Find all spin-2 states by acting repeatedly with $\hat{J}_-$ on the highest state $|2,2\rangle$, including showing that $\hat{J}_-|2,-2\rangle = 0$.

Exercise: Study the effect of infinitesimal rotations on spin-1 states. Consider rotations of each of the three states about the z-axis by arbitrary amounts, and about the x-axis until you can describe what is happening clearly.

3.5 Decomposition of tensors

We have observed previously that a matrix can be decomposed into its trace, its antisymmetric part, and its traceless symmetric part:

$$\begin{aligned} M_{ij} &= \frac{1}{2}\delta_{ij} trM + \frac{1}{2}\left(M_{ij} - M_{ji}\right) + \frac{1}{2}\left(M_{ij} + M_{ji} - \frac{2}{3} trM\right) \\ &= T_{ij} + A_{ij} + S_{ij} \end{aligned}$$

When we rotate M_{ij} with an orthogonal transformation,

$$\begin{aligned} \tilde{M}_{ij} &= O_i^{\ m} O_j^{\ n} M_{mn} \\ &= O_i^{\ m} M_{mn} \left[O^t\right]_{\ i}^{m} \\ \tilde{M} &= OMO^{-1} \end{aligned}$$

each of these parts is preserved. For example, the antisymmetric part of the new matrix is a linear combination of the components of only the antisymmetric part of the original matrix,

$$\begin{aligned} O\frac{1}{2}\left(M - M^t\right)O^{-1} &= \frac{1}{2}\left(OMO^{-1} - OM^tO^{-1}\right) \\ &= \frac{1}{2}\left(\tilde{M} - \tilde{M}^t\right) \end{aligned}$$

We say that the usual matrix representation M_{ij} is reducible, and from the fact that these three invariant subspace have one degree of freedom for the trace, three for the antisymmetric part, and five degrees of freedom for the traceless symmetric part, we might guess that we can write M as a combination of the three vector spaces,

$$|0,0\rangle, |1,m\rangle, |2,m\rangle$$

which are of dimensions $1, 3$ and 5, respectively. What we have accomplished is to find the *irreducible representations* of the rotation group.

There is notation for this equivalence. Letting the boldface number **3** stand for each index of M, we think of the nine components of M as the outer product of 3-dimensional things,

$$M \rightarrow \mathbf{3} \otimes \mathbf{3}$$

and we write this as the sum, in the new notation, of three irreducible vector spaces:

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5}$$

There are more general objects that rotations can act on. By taking outer products of vectors, we construct “tensors” with arbitrarily many indices,

$$T_{ij\ldots k} = u_i v_j \ldots w_k$$

Since we can rotate each vector, we know how $T_{ij\ldots k}$ changes under rotations. We may take abritrary linear combinations of objects of this form to construct n-index objects with 3^n degrees of freedom. For example, a general tensor with three indices, T_{ijk}, has $3^3 = 27$ independent components.

A systematic analysis along these same lines shows that a rank three tensor, that is, an object with three indices like the Levi-Civita tensor, T_{ijk}, may be decomposed into four irreducible parts,

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$$

The 1-dimensional subspace **1** is the totally antisymmetric part of T_{ijk}. The two **8**s are of definite mixed symmetry and the **10** is the totally symmetric part. Notice that the degrees of freedom always match, $3^3 = 27 = 1+8+8+10$, so we have accounted for all 27 independent components of T_{ijk}. There are general techniques for finding this decomposition for any tensor.

One familiar example of this sort of decomposition is given by the spherical harmonics. If we have any bounded, piecewise continuous function on a sphere, $f(\theta, \varphi)$, it may be expanded in spherical harmonics,

$$f(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} Y^l_m(\theta, \varphi)$$

But such functions form an infinite dimensional vector space, since sums of such functions give other functions on the sphere. The collection of spherical harmonics for any fixed l, $\left\{ Y^l_m(\theta, \varphi) \,|\, m = -l, -l+1, \ldots, l \right\}$ also form a vector space, since we may take linear combinations of any two linear combinations of these, to form another linear combinations of the same set. Moreover, these sets are rotationally invariant: any rotation of the sphere $(\theta, \varphi) \to (\theta + \alpha, \varphi + \beta)$ mixes m but leaves l fixed. Since the dimension of these invariant subspaces is $2l+1$, while the dimension of the function space is infinite, the sum above gives us an infinite decomposition,

$$\boldsymbol{\infty} = \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5} \oplus \cdots \oplus (\mathbf{2l} + \mathbf{1}) \oplus \cdots$$

We show in the next set of notes that these odd-dimensional vector spaces are, in fact, spanned by the spherical harmonics.

The importance of such decompositions becomes evident when we look at atoms, nuclei, mesons or baryons, all of which are composite. Atoms are described by electrons orbiting nuclei, while the others are comprised of quarks and gluons. In each of these multi-particle systems, the constituents may have both orbital angular momentum and spin, and we need to know how these various contributions to the total angular momentum combine to give a total number of states for the system. Therefore, we will develop rules for the addition of angular momentum states.

3.6 Measuring spin

We have found the Lie algebra for $Spin(p,q)$ in Eq.(2.14), which reduces in 4-dimensions to

$$\left[\sigma^{\alpha\beta}, \sigma^{\mu\nu}\right] = \eta^{\beta\mu}\sigma^{\alpha\nu} - \eta^{\alpha\mu}\sigma^{\beta\nu} - \eta^{\beta\nu}\sigma^{\alpha\mu} + \eta^{\alpha\nu}\sigma^{\beta\mu}$$

where

$$\sigma^{\alpha\beta} \equiv \frac{1}{4}\left[\gamma^\alpha, \gamma^\beta\right]$$

Define,

$$P^\mu \equiv \frac{1}{2}(1 - \gamma_5)\gamma^\mu$$

Then, since

$$[P^\mu, P^\nu] = \frac{1}{2}(1-\gamma_5)\gamma^\mu \frac{1}{2}(1-\gamma_5)\gamma^\nu = \frac{1}{4}(1-\gamma_5)(1+\gamma_5)\gamma^\mu\gamma^\nu = 0$$

we have

$$[P^\mu, P^\nu] = 0$$

Then, consider the commutator of P^μ with the $Spin\,(1,3)$ generators.

$$\begin{aligned}
\left[\sigma^{\alpha\beta}, P^\mu\right] &= \frac{1}{4}\left[\left[\gamma^\alpha, \gamma^\beta\right], \frac{1}{2}(1-\gamma_5)\gamma^\mu\right] \\
&= \frac{1}{4}\left[\gamma^\alpha\gamma^\beta - \gamma^\beta\gamma^\alpha, \frac{1}{2}(1-\gamma_5)\gamma^\mu\right] \\
&= \frac{1}{8}\left((1-\gamma_5)\left(\gamma^\alpha\gamma^\beta\gamma^\mu - \gamma^\mu\gamma^\alpha\gamma^\beta\right) - (1-\gamma_5)\left(\gamma^\beta\gamma^\alpha\gamma^\mu - \gamma^\mu\gamma^\beta\gamma^\alpha\right)\right) \\
&= \frac{1}{4}(1-\gamma_5)\left(\gamma^\alpha\gamma^\beta\gamma^\mu - \gamma^\mu\gamma^\alpha\gamma^\beta - \gamma^\beta\gamma^\alpha\gamma^\mu + \gamma^\mu\gamma^\beta\gamma^\alpha\right) \\
&= \frac{1}{8}(1-\gamma_5)\left(\gamma^\alpha\gamma^\beta\gamma^\mu - \gamma^\mu\gamma^\alpha\gamma^\beta - \left(-\gamma^\alpha\gamma^\beta + 2\eta^{\alpha\beta}\right)\gamma^\mu + \gamma^\mu\left(-\gamma^\alpha\gamma^\beta + 2\eta^{\alpha\beta}\right)\right) \\
&= \frac{1}{8}(1-\gamma_5)\left(2\gamma^\alpha\gamma^\beta\gamma^\mu - 2\gamma^\mu\gamma^\alpha\gamma^\beta\right)
\end{aligned}$$

and since

$$\begin{aligned}
\gamma^\mu\gamma^\alpha\gamma^\beta &= \left(-\gamma^\alpha\gamma^\mu + 2\eta^{\mu\alpha}\right)\gamma^\beta \\
&= -\gamma^\alpha\left(-\gamma^\beta\gamma^\mu + 2\eta^{\mu\beta}\right) + 2\eta^{\mu\alpha}\gamma^\beta \\
&= \gamma^\alpha\gamma^\beta\gamma^\mu - 2\eta^{\mu\beta}\gamma^\alpha + 2\eta^{\mu\alpha}\gamma^\beta
\end{aligned}$$

this becomes

$$\begin{aligned}
\left[\sigma^{\alpha\beta}, P^\mu\right] &= \frac{1}{8}(1-\gamma_5)\left(2\gamma^\alpha\gamma^\beta\gamma^\mu - 2\gamma^\mu\gamma^\alpha\gamma^\beta\right) \\
&= \frac{1}{8}(1-\gamma_5)\left(2\gamma^\alpha\gamma^\beta\gamma^\mu - 2\left(\gamma^\alpha\gamma^\beta\gamma^\mu - 2\eta^{\mu\beta}\gamma^\alpha + 2\eta^{\mu\alpha}\gamma^\beta\right)\right) \\
&= \frac{1}{8}(1-\gamma_5)\left(4\eta^{\mu\beta}\gamma^\alpha - 4\eta^{\mu\alpha}\gamma^\beta\right) \\
&= \eta^{\mu\beta}P^\alpha - \eta^{\mu\alpha}P^\beta
\end{aligned}$$

Comparing the commutators,

$$\begin{aligned}
\left[\sigma^{\alpha\beta}, \sigma^{\mu\nu}\right] &= \eta^{\beta\mu}\sigma^{\alpha\nu} - \eta^{\alpha\mu}\sigma^{\beta\nu} - \eta^{\beta\nu}\sigma^{\alpha\mu} + \eta^{\alpha\nu}\sigma^{\beta\mu} \\
\left[\sigma^{\alpha\beta}, P^\mu\right] &= \eta^{\mu\beta}P^\alpha - \eta^{\mu\alpha}P^\beta \\
\left[P^\mu, P^\nu\right] &= 0
\end{aligned}$$

to the Lie algebra of the Poincaré group, Eqs.(2.55), (2.56) and (2.57), we see that we have the complete Lie algebra of the Poincaré group in the spin representation.

Now, recall that the square of the spin operator, $W^\mu W_\mu$ is a Casimir operator of the Poincaré group, where, modifying Eq.(2.60) to use the $Spin\,(1,3)$ representation,

$$W^\mu \equiv \frac{1}{2}\varepsilon^\mu{}_{\nu\alpha\beta}P^\nu\sigma^{\alpha\beta}$$

is a Casimir operator of the Poincaré group. We now fix the normalization of W^μ to reproduce our usual spin operators, J^k.

Choosing the rest frame of the momentum, $P^\mu = mc\,(1,\mathbf{0})$, W^μ takes the form,

$$W^\mu \equiv \frac{mc}{2}\varepsilon^\mu{}_{0\alpha\beta}\sigma^{\alpha\beta}$$

In this frame, $W^0 = 0$ so we may write

$$W^i = -\frac{mc}{2}\varepsilon^i{}_{mn}\sigma^{mn}$$

where the Latin indices run over $1,2,3$. We also need

$$\begin{aligned} W_i\varepsilon^{ijk} &= -\frac{mc}{2}\varepsilon^{ijk}\varepsilon_{imn}\sigma^{mn} \\ &= -\frac{mc}{2}\left(\delta^j_m\delta^k_n-\delta^k_m\delta^j_n\right)\sigma^{mn} \\ &= -mc\sigma^{jk} \end{aligned}$$

so that $\sigma^{jk}=-\frac{1}{mc}W_i\varepsilon^{ijk}$. The commutator $\left[W^i,W^j\right]$ is

$$\begin{aligned} \left[W^i,W^j\right] &= \left[\frac{mc}{2}\varepsilon^i{}_{mn}\sigma^{mn},\frac{mc}{2}\varepsilon^j{}_{kl}\sigma^{kl}\right] \\ &= \frac{m^2c^2}{4}\varepsilon^i{}_{mn}\varepsilon^j{}_{kl}\left[\sigma^{mn},\sigma^{kl}\right] \end{aligned}$$

so restricting the indices to spacelike in Eq.(2.55), $\eta^{\alpha\beta}\to\delta^{mn}$ and

$$\left[\sigma^{mn},\sigma^{kl}\right] = \delta^{nk}\sigma^{ml}-\delta^{mk}\sigma^{nl}-\delta^{nl}\sigma^{mk}+\delta^{ml}\sigma^{nk}$$

the commutator becomes

$$\begin{aligned} \left[W^i,W^j\right] &= \frac{m^2c^2}{4}\varepsilon^i{}_{mn}\varepsilon^j{}_{kl}\left(\delta^{nk}\sigma^{ml}-\delta^{mk}\sigma^{nl}-\delta^{nl}\sigma^{mk}+\delta^{ml}\sigma^{nk}\right) \\ &= \frac{m^2c^2}{4}\left(\varepsilon^i{}_{mn}\varepsilon^{jn}{}_l\sigma^{ml}-\varepsilon^i{}_{mn}\varepsilon^{jm}{}_l\sigma^{nl}-\varepsilon^i{}_{mn}\varepsilon^j{}_k{}^n\sigma^{mk}+\varepsilon^i{}_{mn}\varepsilon^j{}_k{}^m\sigma^{nk}\right) \\ &= \frac{m^2c^2}{4}\left(-\left(\delta^{ij}\delta_{ml}-\delta^i_l\delta^j_m\right)\sigma^{ml}-\left(\delta^{ij}\delta_{nl}-\delta^i_l\delta^j_n\right)\sigma^{nl}-\left(\delta^{ij}\delta_{mk}-\delta^i_k\delta^j_m\right)\sigma^{mk}-\left(\delta^{ij}\delta_{nk}-\delta^i_k\delta^j_n\right)\sigma^{nk}\right) \\ &= \frac{m^2c^2}{4}\left(\sigma^{ji}+\sigma^{ji}+\sigma^{ji}+\sigma^{ji}\right) \\ &= -\frac{m^2c^2}{4}\sigma^{ij} \end{aligned}$$

and therefore

$$\left[W^i,W^j\right] = \frac{mc}{4}W_k\varepsilon^{kij}$$

Lowering indices and setting

$$J_i\equiv\frac{4i\hbar}{mc}W_i$$

we have the *Spin* (3) Lie algebra,

$$\begin{aligned} \left[\frac{4i\hbar}{mc}W_i,\frac{4i\hbar}{mc}W_j\right] &= \frac{4i\hbar}{mc}\frac{mc}{4}\varepsilon_{ijk}\frac{4i\hbar}{mc}W^k \\ \left[J_i,J_j\right] &= i\hbar\varepsilon_{ijk}J^k \end{aligned}$$

Normalized in this way, the spin operators of the Poincaré group become

$$W^\mu\equiv\frac{2i\hbar}{mc}\varepsilon^\mu{}_{\nu\alpha\beta}P^\nu\sigma^{\alpha\beta}$$

Chapter 4

Quantization of scalar fields

We have introduced several distinct types of fields, with actions that give their field equations. These include scalar fields,

$$S = \frac{1}{2} \int \left(\partial^\alpha \varphi \partial_\alpha \varphi - m^2 \varphi^2 \right) d^4x \tag{4.1}$$

and complex scalar fields,

$$S = \int \left(\partial^\alpha \varphi^* \partial_\alpha \varphi - m^2 \varphi^* \varphi \right) d^4x \tag{4.2}$$

These are often called *charged scalar fields* because they have a nontrivial global $U(1)$ symmetry that allows them to couple to electromagnetic fields. Scalar fields have spin 0 and mass m.

The next possible value of $W^2 \sim J^2$ is spin-$\frac{1}{2}$, which is possessed by spinors. Dirac spinors satisfy the Dirac equation, which follows from the action of Eq.(2.47),

$$S_D = \int d^4x \, \bar{\psi} \left(i\gamma^\mu \partial_\mu - m \right) \psi$$

Once again, the mass is m. For higher spin, we have the zero mass, spin-1 electromagnetic field, with action

$$S_{EM} = \int d^4x \, \left(\frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} + J^\alpha A_\alpha \right) \tag{4.3}$$

Electromagnetic theory has an important generalization in the Yang-Mills field, $F^A{}_{\alpha\beta}$ where the additional index corresponds to an $SU(n)$ symmetry, where A is an $SU(n)$ index. These fields have a similar action,

$$S_{YM} = \int d^4x \, \left(\frac{1}{4} F^{A\alpha\beta} F^A{}_{\alpha\beta} + J^\alpha A_\alpha \right)$$

and are also massless in order to preserve the gauge symmetry.

We could continue with the spin-$\frac{3}{2}$ Rarita-Schwinger field and the spin-2 metric field, $g_{\alpha\beta}$ of general relativity. The latter obeys the Einstein-Hilbert action,

$$S = \int d^4x \, \sqrt{-\det(g_{\alpha\beta})} g^{\alpha\beta} R^\mu{}_{\alpha\mu\beta} \tag{4.4}$$

where $R^\mu{}_{\alpha\mu\beta}$ is the Riemann curvature tensor computed from $g_{\alpha\beta}$ and its first and second derivatives.

In this, and the next two chapters, we will quantizing the scalar, Dirac, and electromagnetic field in turn.

We need the Hamiltonian formulation of field theory to do this properly, and that will require a bit of functional differentiation. It's actually kind of fun.

4.1 Functional differentiation

What distinguishes a functional such as the action $S[x(t)]$ from a function $f(x(t))$, is that $f(x(t)$ is a number for each value of t, whereas the value of $S[x(t)]$ cannot be computed without knowing the entire function $x(t)$. Thus, functionals are nonlocal. If we think of functions and functionals as maps, then a function maps the reals to the reals (or more generally, R^n to R^m)

$$f \quad : \quad R \to R$$

and $f(x)$ is a number for each real number x. A functional, by contrast, maps an entire function space into R,

$$\begin{aligned} S &\quad : \quad \mathcal{F} \to R \\ \mathcal{F} &\quad = \quad \{f | f : R \to R\} \end{aligned}$$

In this section we develop the *functional derivative*, that is, the generalization of differentiation to functionals.

We would like the functional derivative to formalize finding the extremum of an action integral, so it makes sense to review the variation of an action. The usual argument is that we replace $x(t)$ by $x(t) + h(t)$ in the functional $S[x(t)]$, then demand that to first order in $h(t)$,

$$\delta S \quad \equiv \quad S[x+h] - S[x] = 0$$

We want to replace this statement by the demand that at the extremum, the first functional derivative of $S[x]$ vanishes,

$$\frac{\delta S[x(t)]}{\delta x(t)} = 0$$

Now, suppose S is given by

$$S[x(t)] \quad = \quad \int L(x(t), \dot{x}(t)) dt$$

Then replacing x by $x + h$ and subtracting S gives

$$\begin{aligned} \delta S &\equiv \int L(x+h, \dot{x}+\dot{h}) dt - \int L(x, \dot{x}) dt \\ &= \int \left(L(x, \dot{x}) + \frac{\partial L(x, \dot{x})}{\partial x} h + \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \dot{h} \right) dt - \int L(x, \dot{x}) dt \\ &= \int \left(\frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) h(t)\ dt \end{aligned}$$

Setting $\delta x = h(t)$ we may write this as

$$\delta S = \int \left(\frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt'} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \delta x(t')\ dt'$$

Now write

$$\delta S = \frac{\delta S}{\delta x(t)} \delta x(t) = \left(\int \left(\frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt'} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \frac{\delta x(t')}{\delta x(t)}\ dt' \right) \delta x(t)$$

or simply

$$\frac{\delta S}{\delta x(t)} = \int \left(\frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt'} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \frac{\delta x(t')}{\delta x(t)}\ dt' \tag{4.5}$$

We might write this much by just using the chain rule. What we need is to evaluate the basic functional derivative,

$$\frac{\delta x(t')}{\delta x(t)}$$

To see what this might be, consider the analogous derivative for a countable number of degrees of freedom. Beginning with

$$\frac{\partial q^j}{\partial q^i} = \delta_i^j$$

we notice that when we sum over the i index holding j fixed, we have

$$\sum_i \frac{\partial q^j}{\partial q^i} = \sum_j \delta_i^j = 1$$

since $j = i$ for only one value of j. We demand the continuous version of this relationship. The sum over independent coordinates becomes an integral, $\sum_i \rightarrow \int dt'$, so we demand

$$\int \frac{\delta x(t')}{\delta x(t)} dt' = 1$$

This will be true provided we use a Dirac delta function for the derivative:

$$\frac{\delta x(t')}{\delta x(t)} = \delta(t' - t) \tag{4.6}$$

Substituting this expression into Eq.(4.5) gives the desired result for $\frac{\delta S}{\delta x(t)}$:

$$\begin{aligned}
\frac{\delta S}{\delta x(t)} &= \int \left(\frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt'} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \delta(t' - t)\ dt' \\
&= \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}}
\end{aligned}$$

The Dirac delta function extracts the equation of motion.

4.1.1 The details

Using a delta function works ideally to extract the Euler-Lagrange equation works, but it makes no obvious sense as a perturbation of the path. To connect these ideas we resort to the definition of the delta function as a limit of test functions,

$$\delta\left(t_0 - t'\right) = \lim_{n \rightarrow \infty} h_n\left(t_0, t'\right) \tag{4.7}$$

where for any smooth function $f\left(t'\right)$ of compact support,

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f\left(t'\right) h_n\left(t_0, t'\right) dt' = f\left(t_0\right)$$

We can choose the sequence of functions $h_n\left(t_0, t'\right)$ to charactrize deformations of the path, so that the change in the path is

$$x_n\left(\varepsilon, t'\right) = x\left(t'\right) + \varepsilon h_n\left(t_0, t'\right) \tag{4.8}$$

For this x_n, the change the action is

$$S\left[x_n\left(\varepsilon, t'\right)\right] = \int L\left(x\left(t'\right) + \varepsilon h_n\left(t_0, t'\right), \dot{x}\left(t'\right) + \varepsilon \dot{h}_n\left(t_0, t'\right)\right) dt'$$

Our previous procedure was to subtract the original action and keep terms to first order, but this may now be accomplished by differentiating with respect to ε:

$$\begin{aligned}
\frac{d}{d\varepsilon} S\left[x_n\left(\varepsilon, t'\right)\right] &= \int \frac{d}{d\varepsilon} L\left(x\left(t'\right) + \varepsilon h_n\left(t_0, t'\right), \dot{x}\left(t'\right) + \varepsilon \dot{h}_n\left(t_0, t'\right)\right) \\
&= \int \frac{\partial L}{\partial x} h_n\left(t_0, t'\right) + \frac{\partial L}{\partial \dot{x}} \dot{h}_n\left(t_0, t'\right)
\end{aligned}$$

where the partial $\frac{\partial L}{\partial x}$ is with respect to the first variable in $L(x, \dot{x})$ and the second partial with respect to the second. The chain rule then leaves h_n and $\dot{h}_n$, respectively. Integrating the second term by parts and discarding the surface term,

$$\frac{d}{d\varepsilon} S\left[x_n\left(\varepsilon, t'\right)\right] = \int \left(\frac{\partial L}{\partial x} - \frac{d}{dt'}\left(\frac{\partial L}{\partial \dot{x}}\right)\right) h_n\left(t_0, t'\right)$$

The form of the Lagrangian here is still

$$L\left(x\left(t'\right) + \varepsilon h_n\left(t_0, t'\right), \dot{x}\left(t'\right) + \varepsilon \dot{h}_n\left(t_0, t'\right)\right)$$

but after differentiating we may set $\varepsilon = 0$. This is equivalent to keeping only the first order term, and reduces the Lagrangian to $L\left(x\left(t'\right), \dot{x}\left(t'\right)\right)$.

Since this is valid for every $h_n\left(t_0, t'\right)$, we may take the limit and define

$$\begin{aligned}
\frac{\delta S}{\delta x} &\equiv \lim_{n\to\infty}\left[\frac{d}{d\varepsilon} S\left[x_n\left(\varepsilon, t'\right)\right]\right]_{\varepsilon=0} \\
&= \lim_{n\to\infty}\int \left(\frac{\partial L\left(x\left(t'\right), \dot{x}\left(t'\right)\right)}{\partial x} - \frac{d}{dt'}\left(\frac{\partial L\left(x\left(t'\right), \dot{x}\left(t'\right)\right)}{\partial \dot{x}}\right)\right) h_n\left(t_0, t'\right)\, dt' \\
&= \int \left(\frac{\partial L}{\partial x} - \frac{d}{dt'}\left(\frac{\partial L}{\partial \dot{x}}\right)\right) \delta\left(t' - t_0\right)\, dt' \\
&= \left[\frac{\partial}{\partial x} L\left(x\left(t'\right), \dot{x}\left(t'\right)\right) - \frac{d}{dt'}\left(\frac{\partial}{\partial \dot{x}} L\left(x\left(t'\right), \dot{x}\left(t'\right)\right)\right)\right]_{t'=t_0}
\end{aligned}$$

Since t_0 is arbitrary, the final expression holds for any t.

We therefore define the functional derivative of a functional $S[x]$ to be

$$\frac{\delta S}{\delta x} \equiv \lim_{n\to\infty}\left[\frac{d}{d\varepsilon} S\left[x_n\left(\varepsilon, t'\right)\right]\right]_{\varepsilon=0}$$

where the smooth functions $h_n\left(t, t'\right)$ satisfy Eq.(4.7) and $x_n\left(\varepsilon, t'\right)$ is given by Eq.(4.8).

The derivative with respect to ε accomplishes the usual variation of the action, and then setting $\varepsilon = 0$ selects the linear part of the variation. Then we take the limit of a carefully chosen sequence of variations h_n to extract the variational coefficient from the integral.

There are various advantages to this more formal approach. One is that we can equally well apply the technique to classical field theory, and another is that we may iterate the definition to take higher functional derivatives.

4.1.2 Example: Field equations as functional derivatives

We can vary field actions in the same way, and the results make sense directly. Consider varying the functional derivative of the Klein-Gordon scalar field action:

$$S = \frac{1}{2}\int \left(\partial^\alpha \varphi \partial_\alpha \varphi - m^2 \varphi^2\right) d^4 x$$

with respect to the field φ. The Lagrangian is now a functional,

$$\begin{aligned}
L &= \frac{1}{2}\int \left(\partial^\alpha \varphi \partial_\alpha \varphi - m^2 \varphi^2\right) d^3 x \qquad (4.9) \\
S &= \int L dt
\end{aligned}$$

The functional derivative of S is

$$\begin{aligned}
\frac{\delta S[\varphi]}{\delta\varphi(\mathbf{y})} &= \frac{1}{2}\frac{\delta}{\delta\varphi(\mathbf{y})}\int\left(\partial^{\alpha}\varphi\left(\mathbf{x}\right)\partial_{\alpha}\varphi\left(\mathbf{x}\right)-m^{2}\varphi^{2}\left(\mathbf{x}\right)\right)d^{4}x \\
&= \int\left(\partial^{\alpha}\varphi\frac{\partial}{\partial x^{\alpha}}\left(\frac{\delta\varphi\left(\mathbf{x}\right)}{\delta\varphi\left(\mathbf{y}\right)}\right)-m^{2}\varphi\frac{\delta\varphi\left(\mathbf{x}\right)}{\delta\varphi\left(\mathbf{y}\right)}\right)d^{4}x \\
&= \int\left(-\partial_{\alpha}\partial^{\alpha}\varphi-m^{2}\varphi\right)\frac{\delta\varphi\left(\mathbf{x}\right)}{\delta\varphi\left(\mathbf{y}\right)}d^{4}x \\
&= \int\left(-\partial_{\alpha}\partial^{\alpha}\varphi-m^{2}\varphi\right)\delta^{4}(\mathbf{x}-\mathbf{y})d^{4}x \\
&= -\Box\varphi\left(\mathbf{y}\right)-m^{2}\varphi\left(\mathbf{y}\right)
\end{aligned}$$

and the vanishing of the first functional derivative is the field equation, $\left(\Box+m^{2}\right)\varphi\left(\mathbf{y}\right)=0$.

Exercise: Find the field equation for the complex scalar field by taking the functional derivative of its action, eq.(4.2).

Exercise: Find the field equation for the Dirac field by taking the functional derivative of its action, eq.(2.47).

Exercise: Find the Maxwell equations by taking the functional derivative of its action, eq.(4.3).

With this new tool at our disposal, we turn to quantization.

4.2 Quantization of the Klein-Gordon (scalar) field

We develop the Hamiltonian formulation, then canonically quantize.

4.2.1 The conjugate momentum

To begin quantization, we require the Hamiltonian formulation of scalar field theory. Beginning with the Lagrangian,

$$L = \frac{1}{2}\int\left(\partial^{\alpha}\varphi\partial_{\alpha}\varphi-m^{2}\varphi^{2}\right)d^{3}x$$

the only modification to the definition of the conjugate momentum as

$$p=\frac{\partial L}{\partial\dot{x}}$$

is the recognition that (a) the independent variable is dependent on four parameters $\varphi=\varphi\left(x^{\alpha}\right)$ instead of just one, and (b) the Lagrangian is now a functional, Eq.(4.9). Just as the time derivative must be changed from a total to a partial derivative,

$$\dot{x}=\frac{dx}{dt}\Longrightarrow\dot{\varphi}=\frac{\partial\varphi}{\partial t}$$

the derivative of the Lagrangian must go to a functional derivative of the Lagrangian

$$\frac{\partial L}{\partial\dot{x}} \Longrightarrow \frac{\delta L}{\delta\dot{\varphi}}$$

Writing $\pi(y^\mu) = \pi(y)$, the conjugate momentum is therefore,

$$\begin{aligned}
\pi(y) &\equiv \frac{\delta L[x]}{\delta(\partial_0 \varphi(y))} \\
&= \frac{\delta}{\delta(\partial_0 \varphi(y))} \frac{1}{2} \int \left(\partial^\alpha \varphi(x) \partial_\alpha \varphi(x) - m^2 \varphi^2(x)\right) d^3x \\
&= \int \partial^0 \varphi(x) \delta^3(\mathbf{y} - \mathbf{x}) d^3x' \\
&= \partial^0 \varphi(t, \mathbf{y})
\end{aligned}$$

Notice that we treat $\varphi(\mathbf{x})$ and its derivatives $\partial_\alpha \varphi(\mathbf{x})$ as independent. In terms of the momentum density, the action and Lagrangian density are

$$S = \frac{1}{2} \int \left(\pi^2 - \nabla\varphi \cdot \nabla\varphi - m^2 \varphi^2\right) d^4x \tag{4.10}$$

$$\mathcal{L} = \frac{1}{2} \left(\pi^2 - \nabla\varphi \cdot \nabla\varphi - m^2 \varphi^2\right) \tag{4.11}$$

4.2.2 The Hamiltonian and Poisson brackets

We must also generalize the expression for the Hamiltonian. For the infinite number of field degrees of freedom (labeled by the spatial coordinates $\mathbf{x}$), the sum in the expression for the Hamiltonian becomes an integral, so that $H = \sum p_i \dot{q}^i - L$ generalizes to

$$H = \int \pi(\mathbf{x}, t) \dot{\varphi}(\mathbf{x}, t) d^3x - L \tag{4.12}$$

Therefore,

$$\begin{aligned}
H &= \int \pi(\mathbf{x}) \dot{\varphi}(\mathbf{x}) d^3x - \frac{1}{2} \int \left(\partial^\alpha \varphi \partial_\alpha \varphi - m^2 \varphi^2\right) d^3x \\
&= \frac{1}{2} \int \left(\pi^2 + \nabla\varphi \cdot \nabla\varphi + m^2 \varphi^2\right) d^3x
\end{aligned} \tag{4.13}$$

We can define the *Hamiltonian density,*

$$\mathcal{H} = \frac{1}{2} \left(\pi^2 + \nabla\varphi \cdot \nabla\varphi + m^2 \varphi^2\right) \tag{4.14}$$

Hamilton's equations can also be expressed in terms of densities. Starting from Hamilton's equations in the familiar form,

$$\begin{aligned}
\dot{q}^i &= \frac{\partial H}{\partial p_i} \\
\dot{p}_i &= -\frac{\partial H}{\partial q^i}
\end{aligned}$$

we replace $\left(q^i, p_j\right)$ with (φ, π) and since the Hamiltonian is a functional, replace the partial derivative with functional derivatives,

$$\dot{\varphi}(\mathbf{x}) = \frac{\delta H}{\delta \pi(\mathbf{x})} \tag{4.15}$$

$$\dot{\pi}(\mathbf{x}) = -\frac{\delta H}{\delta \varphi(\mathbf{x})} \tag{4.16}$$

and check that our procedure reproduces the correct field equation by taking the indicated derivatives. Carrying out the functional derivative for $\dot{\varphi}$,

$$\begin{aligned}
\dot{\varphi}(\mathbf{x}) &= \frac{\delta H}{\delta \pi_i(\mathbf{x})} \\
&= \frac{1}{2}\frac{\delta}{\delta \pi_i(\mathbf{x})}\int \left(\pi^2 + \nabla\varphi\cdot\nabla\varphi + m^2\varphi^2\right) d^3y \\
&= \frac{1}{2}\int \left(2\pi(\mathbf{y})\frac{\delta\pi(\mathbf{y})}{\delta\pi(\mathbf{x})}\right) d^3y \\
&= \int \pi(\mathbf{y})\,\delta^3(\mathbf{x}-\mathbf{y})\, d^3y \\
&= \pi(\mathbf{x})
\end{aligned}$$

This agrees with our definition of $\pi(\mathbf{x})$. For π we find

$$\begin{aligned}
\dot{\pi}(\mathbf{x}) &= -\frac{\delta H}{\delta \varphi(\mathbf{x})} \\
&= -\frac{1}{2}\frac{\delta}{\delta \varphi(\mathbf{x})}\int \left(\pi^2 + \nabla\varphi\cdot\nabla\varphi + m^2\varphi^2\right) d^3y \\
&= -\int \left(\nabla\varphi\cdot\nabla\frac{\delta\varphi(\mathbf{y})}{\delta\varphi(\mathbf{x})} + m^2\varphi\frac{\delta\varphi(\mathbf{y})}{\delta\varphi(\mathbf{x})}\right) d^3y \\
&= \int \left(\nabla^2\varphi\cdot\frac{\delta\varphi(\mathbf{y})}{\delta\varphi(\mathbf{x})} - m^2\varphi\frac{\delta\varphi(\mathbf{y})}{\delta\varphi(\mathbf{x})}\right) d^3y \\
&= \int \left(\nabla^2\varphi\delta^3(\mathbf{y}-\mathbf{x}) - m^2\varphi\delta^3(\mathbf{y}-\mathbf{x})\right) d^3y \\
&= \nabla^2\varphi(\mathbf{x}) - m^2\varphi(\mathbf{x})
\end{aligned}$$

But $\dot{\pi} = \partial_0\pi = \partial_0\partial^0\varphi$ so

$$\Box\varphi = -m^2\varphi$$

and we recover the Klein-Gordon field equation.

We move toward quantization by writing the field equations in terms of functional Poisson brackets. Let

$$\{f(\varphi,\pi), g(\varphi,\pi)\} \equiv \int \left(\frac{\delta f}{\delta\pi(\mathbf{x})}\frac{\delta g}{\delta\varphi(\mathbf{x})} - \frac{\delta f}{\delta\varphi(\mathbf{x})}\frac{\delta g}{\delta\pi(\mathbf{x})}\right) d^3x \tag{4.17}$$

where we replace the sum over all p_i and q^i by an integral over $\mathbf{x}$, and where $f(\varphi,\pi) = f(\varphi(\mathbf{y},t), \pi(\mathbf{y},t))$ and $g(\varphi,\pi) = g(\varphi(\mathbf{z},t), \pi(\mathbf{z},t))$. The bracket is evaluated at a constant time. Then we have

$$\begin{aligned}
\{\pi(\mathbf{y},t), \varphi(\mathbf{z},t)\} &= \int \left(\frac{\delta\pi(\mathbf{y},t)}{\delta\pi(\mathbf{x})}\frac{\delta\varphi(\mathbf{z},t)}{\delta\varphi(\mathbf{x})} - \frac{\delta\pi(\mathbf{y},t)}{\delta\varphi(\mathbf{x})}\frac{\delta\varphi(\mathbf{z},t)}{\delta\pi(\mathbf{x})}\right) d^3x \\
&= \int \delta^3(\mathbf{y}-\mathbf{x})\,\delta^3(\mathbf{z}-\mathbf{x})\, d^3x \\
&= \delta^3(\mathbf{z}-\mathbf{y})
\end{aligned}$$

while

$$\{\pi(\mathbf{y},t), \pi(\mathbf{z},t)\} = \{\varphi(\mathbf{y},t), \varphi(\mathbf{z},t)\} = 0$$

Hamilton's equations work out correctly:

$$\begin{aligned}
\dot{\varphi}(\mathbf{x}) &= \{H(\varphi,\pi),\varphi(\mathbf{x}')\} \\
&= \int\left(\frac{\delta H(\varphi,\pi)}{\delta\pi(\mathbf{x})}\frac{\delta\varphi(\mathbf{x}')}{\delta\varphi(\mathbf{x})}-\frac{\delta H}{\delta\varphi(\mathbf{x})}\frac{\delta\varphi(\mathbf{x}')}{\delta\pi(\mathbf{x})}\right)d^3x \\
&= \int\frac{\delta H(\varphi,\pi)}{\delta\pi(\mathbf{x})}\delta^3(\mathbf{x}-\mathbf{x}')d^3x \\
&= \frac{\delta H(\varphi(\mathbf{x}),\pi(\mathbf{x}))}{\delta\pi(\mathbf{x})}
\end{aligned}$$

and

$$\begin{aligned}
\dot{\pi}(\mathbf{x}) &= \{H(\varphi,\pi),\pi(\mathbf{x}')\} \\
&= \int\left(\frac{\delta H(\varphi,\pi)}{\delta\pi(\mathbf{x})}\frac{\delta\pi(\mathbf{x}')}{\delta\varphi(\mathbf{x})}-\frac{\delta H}{\delta\varphi(\mathbf{x})}\frac{\delta\pi(\mathbf{x}')}{\delta\pi(\mathbf{x})}\right)d^3x \\
&= -\int\frac{\delta H(\varphi,\pi)}{\delta\varphi(\mathbf{x})}\delta^3(\mathbf{x}-\mathbf{x}')d^3x \\
&= -\frac{\delta H(\varphi(\mathbf{x}),\pi(\mathbf{x}))}{\delta\varphi(\mathbf{x})}
\end{aligned}$$

Now we quantize, canonically. The field and its conjugate momentum become operators and the fundamental Poisson brackets become commutators:

$$\{\pi(\mathbf{x}'),\varphi(\mathbf{x}'')\}=\delta^3(\mathbf{x}''-\mathbf{x}')\Rightarrow[\hat{\pi}(\mathbf{x}'),\hat{\varphi}(\mathbf{x}'')]=i\delta^3(\mathbf{x}''-\mathbf{x}')$$

(where $\hbar=1$) while

$$[\hat{\varphi}(\mathbf{x}'),\hat{\varphi}(\mathbf{x}'')]=[\hat{\pi}(\mathbf{x}'),\hat{\pi}(\mathbf{x}'')]=0$$

These are the fundamental commutation relations of the quantum field theory. Because the commutator of the field operators $\hat{\pi}(\mathbf{x})$ and $\hat{\varphi}(\mathbf{x})$ are evaluated at the same value of t, these are called *equal time commutation relations.* More explicitly,

$$\begin{aligned}
[\hat{\pi}(\mathbf{x}',t),\hat{\varphi}(\mathbf{x}'',t)] &= i\delta^3(\mathbf{x}''-\mathbf{x}') \\
[\hat{\varphi}(\mathbf{x}',t),\hat{\varphi}(\mathbf{x}'',t)] &= [\hat{\pi}(\mathbf{x}',t),\hat{\pi}(\mathbf{x}'',t)]=0
\end{aligned}\tag{4.18}$$

This completes the canonical quantization. The trick, of course, is to characterize the states these operators act on.

4.2.3 Solution for the free classical Klein-Gordon field

Having written commutation relations for the field, we still have the problem of finding solutions and interpreting them. To begin, we look at solutions the classical theory. The field equation

$$\Box\varphi = -\frac{m^2}{\hbar^2}\varphi$$

(where we keep $\hbar$, but set $c=1$) is not hard to solve. Consider plane waves,

$$\begin{aligned}
\varphi(\mathbf{x},t) &= Ae^{\frac{i}{\hbar}(p_\alpha x^\alpha)}+A^*e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \\
&= Ae^{\frac{i}{\hbar}(Et-\mathbf{p}\cdot\mathbf{x})}+A^*e^{-\frac{i}{\hbar}(Et-\mathbf{p}\cdot\mathbf{x})}
\end{aligned}$$

where we add the complex conjugate because φ is real. Substituting into the field equation we have

$$A\left(\frac{i}{\hbar}\right)^2 p_\alpha p^\alpha\exp\frac{i}{\hbar}(p_\alpha x^\alpha)=-\frac{m^2}{\hbar^2}A\exp\frac{i}{\hbar}(p_\alpha x^\alpha)$$

so we need the usual mass-energy-momentum relation:

$$p_\alpha p^\alpha = m^2$$

We can solve this for the energy,

$$\begin{aligned} E_+ &= \sqrt{\mathbf{p}^2 + m^2} \\ E_- &= -\sqrt{\mathbf{p}^2 + m^2} \end{aligned}$$

then construct the general solution by Fourier superposition. To keep the result manifestly relativistic, we use a Dirac delta function to impose the energy condition, $p_\alpha p^\alpha = m^2$. We also insert a unit step function, $\Theta(E)$, to insure positivity of the energy. This insertion may seem a bit *ad hoc*, and it is – we will save discussion of the negative energy solutions and antiparticles for the last section of this chapter. Then,

$$\begin{aligned} \varphi(\mathbf{x}, t) &= \frac{1}{(2\pi)^{3/2}} \int \sqrt{2E} \left(a(E, \mathbf{p}) e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^*(E, \mathbf{p}) e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \right) \\ & \quad \times \delta\left(p_\alpha p^\alpha - m^2\right) \Theta(E) \hbar^{-4} d^4 p \end{aligned} \tag{4.19}$$

where $A = \sqrt{2E}a(E, \mathbf{p})$ is the arbitrary complex amplitude of each wave mode and $\frac{1}{(2\pi)^{3/2}}$ is the conventional normalization for Fourier integrals.

Recall that for a function $f(x)$ with zeros at x_i, $i = 1, 2, \ldots, n$, $\delta(f)$ gives a contribution at each zero:

$$\delta(f) = \sum_{i=1}^{n} \frac{1}{|f'(x_i)|} \delta(x - x_i) \tag{4.20}$$

so the quadratic delta function can be written as

$$\begin{aligned} \delta\left(p_\alpha p^\alpha - m^2\right) &= \delta\left(E^2 - \mathbf{p}^2 - m^2\right) \\ &= \frac{1}{2|E|} \delta\left(E - \sqrt{\mathbf{p}^2 + m^2}\right) + \frac{1}{2|E|} \delta\left(E + \sqrt{\mathbf{p}^2 + m^2}\right) \end{aligned} \tag{4.21}$$

Exercise: Prove eq.(4.20).

Exercise: Argue that $\Theta(E)$ is Lorentz invariant.

The integral for the solution $\varphi(\mathbf{x}, t)$ becomes

$$\begin{aligned} \varphi(\mathbf{x}, t) &= \frac{1}{(2\pi)^{3/2}} \int \sqrt{2\hbar E} \left\{ \left(a e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^* e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \right) \frac{1}{2|E|} \delta\left(E - \sqrt{\mathbf{p}^2 + m^2}\right) \right. \\ & \quad \left. + \left(a e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^\dagger e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \right) \frac{1}{2|E|} \delta\left(E + \sqrt{\mathbf{p}^2 + m^2}\right) \right\} \Theta(E) \hbar^{-3} d^4 p \\ &= \frac{1}{(2\pi)^{3/2}} \int \left(a e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^* e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \right) \frac{1}{\sqrt{\frac{2}{\hbar}|E|}} \delta\left(E - \sqrt{\mathbf{p}^2 + m^2}\right) \hbar^{-3} d^4 p \\ &= \frac{1}{(2\pi)^{3/2}} \int \left(a e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^* e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \right) \frac{1}{\sqrt{2\omega}} d^3 k \\ &= \frac{1}{(2\pi)^{3/2}} \int \left(a e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^* e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \right) \frac{1}{\sqrt{2\omega}} d^3 k \end{aligned}$$

Define the wave vector k^μ,

$$\begin{aligned} k^\mu &= (\omega, \mathbf{k}) \\ \mathbf{k} &= \frac{\mathbf{p}}{\hbar} \\ \omega &= +\frac{1}{\hbar} \sqrt{\mathbf{p}^2 + m^2} = +\sqrt{\mathbf{k}^2 + \left(\frac{m}{\hbar}\right)^2} \end{aligned}$$

Then integrating over the energy delta function,

$$\varphi(\mathbf{x},t) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega}} \left(a(\mathbf{k})\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} + a^*(\mathbf{k})\, e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{4.22}$$

This is the general classical solution for the Klein-Gordon field. Notice that since $\omega = \omega(\mathbf{k})$, the amplitudes a and a^* depend only on $\mathbf{k}$. We also need the conjugate momentum,

$$\begin{aligned} \pi(\mathbf{x},t) &= \partial_0 \varphi(\mathbf{x},t) \\ &= \frac{i}{(2\pi)^{3/2}} \int d^3k \sqrt{\frac{\omega}{2}} \left(a(\mathbf{k})\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} - a^*(\mathbf{k})\, e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \end{aligned} \tag{4.23}$$

To check that our solution satisfies the Klein-Gordon equation, we need only apply the wave operator to the right side. This pulls down an overall factor of $(ik_\mu)(ik^\mu) = -\frac{1}{\hbar^2}\left(E^2 - \mathbf{p}^2\right) = -\frac{m^2}{\hbar^2}$. Since this is constant, it comes out of the integral, giving $-\frac{m^2}{\hbar^2}\varphi$ as required.

4.2.4 Quantization of the mode amplitudes

Now we need to quantize the classical *solution*. We know the fundamental commutation relations that $\hat{\varphi}$ and $\hat{\pi}$ satisfy, Eqs.(4.18) as operators, but we need to see the effect this has on the right hand side of the solution, Eq.(4.22). To do this, we first invert the classical Fourier integrals to solve for the coefficients in terms of the fields. To this end, multiply $\varphi(\mathbf{x},t)$ by $\frac{1}{(2\pi)^{3/2}} d^3x e^{i\mathbf{k}'\cdot\mathbf{x}}$ and integrate. It proves sufficient to evaluate the expression at $t = 0$.

$$\begin{aligned} \frac{1}{(2\pi)^{3/2}} \int \varphi(\mathbf{x},0)\, e^{i\mathbf{k}'\cdot\mathbf{x}} d^3x &= \frac{1}{(2\pi)^3} \int\int \frac{d^3x d^3k}{\sqrt{2\omega}} \left(a(\mathbf{k})\, e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}} + a^*(\mathbf{k})\, e^{i(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}} \right) \\ &= \frac{1}{(2\pi)^3} \int \frac{d^3k}{\sqrt{2\omega}} \left(a(\mathbf{k})\,(2\pi)^3 \delta^3(\mathbf{k}'-\mathbf{k}) + a^*(\mathbf{k})\,(2\pi)^3 \delta^3(\mathbf{k}'+\mathbf{k}) \right) \\ &= \frac{1}{\sqrt{2\omega'}} \left(a(\mathbf{k}') + a^*(-\mathbf{k}') \right) \end{aligned} \tag{4.24}$$

where we have used the Fourier representation of the Dirac delta function

$$\delta^3(\mathbf{k}) = \frac{1}{(2\pi)^3} \int d^3x\; e^{i\mathbf{k}\cdot\mathbf{x}}$$

Once again taking the Fourier transform, $\frac{1}{(2\pi)^{3/2}} \int \pi(\mathbf{x},0)\, e^{i\mathbf{k}'\cdot\mathbf{x}} d^3x$, of the momentum density, we find it equal to

$$\begin{aligned} \frac{1}{(2\pi)^{3/2}} \int \pi(\mathbf{x},0)\, e^{i\mathbf{k}'\cdot\mathbf{x}} d^3x &= \frac{i}{(2\pi)^3} \int d^3x \int d^3k \sqrt{\frac{\omega}{2}} \left(a(\mathbf{k})\, e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}} - a^\dagger(\mathbf{k})\, e^{i(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}} \right) \\ &= i \int d^3k \sqrt{\frac{\omega}{2}} \left(a(\mathbf{k})\, \delta^3(\mathbf{k}'-\mathbf{k}) - a^\dagger(\mathbf{k})\, \delta^3(\mathbf{k}'+\mathbf{k}) \right) \\ &= i \sqrt{\frac{\omega'}{2}} \left(a(\mathbf{k}') - a^\dagger(-\mathbf{k}') \right) \end{aligned} \tag{4.25}$$

These results combine to solve for the amplitudes. Adding $\sqrt{2\omega'}$ times Eq.(4.24) to $-i\sqrt{\frac{2}{\omega'}}$ times (4.23) gives $a(\mathbf{k}')$:

$$\left(a(\mathbf{k}') + a^\dagger(-\mathbf{k}') \right) + \left(a(\mathbf{k}') - a^\dagger(-\mathbf{k}') \right) = \frac{\sqrt{2\omega'}}{(2\pi)^{3/2}} \int \varphi(\mathbf{x},0)\, e^{i\mathbf{k}'\cdot\mathbf{x}} d^3x - \frac{i}{(2\pi)^{3/2}} \sqrt{\frac{2}{\omega'}} \int \pi(\mathbf{x},0)\, e^{i\mathbf{k}'\cdot\mathbf{x}} d^3x$$

Simplifying, we have solved for the mode amplitudes,

$$a(\mathbf{k}') = \frac{\sqrt{2\omega'}}{2(2\pi)^{3/2}} \int \left(\varphi(\mathbf{x},0) - \frac{i}{\omega'}\pi(\mathbf{x},0) \right) e^{i\mathbf{k}'\cdot\mathbf{x}} d^3x \tag{4.26}$$

Subtracting instead of adding, and replacing $\mathbf{k}' \to -\mathbf{k}'$ gives the conjugate mode amplitudes,

$$a^*(\mathbf{k}') = \frac{\sqrt{2\omega'}}{2(2\pi)^{3/2}} \int \left(\varphi(\mathbf{x},0) + \frac{i}{\omega'}\pi(\mathbf{x},0) \right) e^{-i\mathbf{k}'\cdot\mathbf{x}} d^3x \tag{4.27}$$

This gives the amplitudes in terms of the field and its conjugate momentum. So far, this result is classical.

Now we quantize the amplitudes by replacing φ and π by the operators, $\hat{\varphi}$ and $\hat{\pi}$. Clearly, once φ and π become operators, the amplitudes must too; there is no other field present that could become an operator instead. Dropping the primes and noting that the complex conjugate becomes the adjoint operator,

$$\hat{a}(\mathbf{k}) = \frac{\sqrt{2\omega}}{2(2\pi)^{3/2}} \int \left(\hat{\varphi}(\mathbf{x},0) - \frac{i}{\omega'}\hat{\pi}(\mathbf{x},0) \right) e^{i\mathbf{k}\cdot\mathbf{x}} d^3x \tag{4.28}$$

$$\hat{a}^\dagger(\mathbf{k}) = \frac{\sqrt{2\omega}}{2(2\pi)^{3/2}} \int \left(\hat{\varphi}(\mathbf{x},0) + \frac{i}{\omega'}\hat{\pi}(\mathbf{x},0) \right) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3x \tag{4.29}$$

From the commutation relations for φ and π we can compute those for a and $a^\dagger$.

$$\begin{aligned}
\left[\hat{a}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')\right] &= \frac{\sqrt{\omega\omega'}}{2(2\pi)^3} \int\int e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} d^3x\, d^3x' \left[\hat{\varphi}(\mathbf{x},0) - \frac{i}{\omega}\hat{\pi}(\mathbf{x},0), \hat{\varphi}(\mathbf{x}',0) + \frac{i}{\omega'}\hat{\pi}(\mathbf{x}',0) \right] \\
&= \frac{\sqrt{\omega\omega'}}{2(2\pi)^3} \int\int e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} d^3x\, d^3x' \left(\frac{i}{\omega'} \left[\hat{\varphi}(\mathbf{x},0), \hat{\pi}(\mathbf{x}',0)\right] - \frac{i}{\omega}\left[\hat{\pi}(\mathbf{x},0), \hat{\varphi}(\mathbf{x}',0)\right] \right) \\
&= \frac{\sqrt{\omega\omega'}}{2(2\pi)^3} \int\int e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} d^3x\, d^3x' \left(\frac{2}{\omega'}\delta^3(\mathbf{x}-\mathbf{x}') \right)
\end{aligned}$$

The Dirac delta function allows us to evaluate the integrals,

$$\begin{aligned}
\left[\hat{a}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')\right] &= \sqrt{\frac{\omega}{\omega'}} \frac{1}{(2\pi)^3} \int\int e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} d^3x\, d^3x' \delta^3(\mathbf{x}-\mathbf{x}') \\
&= \sqrt{\frac{\omega}{\omega'}} \frac{1}{(2\pi)^3} \int e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} d^3x \\
&= \delta^3(\mathbf{k}-\mathbf{k}')
\end{aligned}$$

Notice that the delta function makes $\omega = \omega'$.

Exercise: Show that $[\hat{a}(\mathbf{k}), \hat{a}(\mathbf{k}')] = 0$.

Exercise: Show that $\left[\hat{a}^\dagger(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')\right] = 0$.

Finally, we summarize by the field and momentum density operators in terms of the mode amplitude operators:

$$\hat{\varphi}(\mathbf{x},t) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega}} \left(\hat{a}(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} + \hat{a}^\dagger(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{4.30}$$

$$\hat{\pi}(\mathbf{x},t) = \frac{i}{(2\pi)^{3/2}} \int d^3k \sqrt{\frac{\omega}{2}} \left(\hat{a}(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} - \hat{a}^\dagger(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{4.31}$$

Next, we turn to a study of states. To begin, we require the Hamiltonian operator, which requires a bit of calculation.

4.2.5 Calculation of the Hamiltonian operator

This is our first typical quantum field theory calculation. They're a bit tricky to keep track of, but not really that hard. Our goal is to compute the expression for the Hamiltonian operator

$$\hat{H} \equiv \frac{\hbar}{2}\int\left(\hat{\pi}^2+\boldsymbol{\nabla}\hat{\varphi}\cdot\boldsymbol{\nabla}\hat{\varphi}+m^2\hat{\varphi}^2\right)d^3x \tag{4.32}$$

in terms of the mode operators. Because the techniques involved are used frequently in field theory calculations, we include all the gory details.

Let's consider one term at a time. For the first,

$$\begin{aligned}
I_\pi &= \frac{\hbar}{2}\int\hat{\pi}^2 d^3x \\
&= \frac{\hbar}{2}\int d^3x\left[\frac{i}{(2\pi)^{3/2}}\int d^3k\sqrt{\frac{\omega}{2}}\left(\hat{a}(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}-\hat{a}^\dagger(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)\right. \\
&\qquad\left.\times\frac{i}{(2\pi)^{3/2}}\int d^3k'\sqrt{\frac{\omega'}{2}}\left(\hat{a}(\mathbf{k}')e^{i\left(\omega' t-\mathbf{k}'\cdot\mathbf{x}\right)}-\hat{a}^\dagger(\mathbf{k}')e^{-i\left(\omega' t-\mathbf{k}'\cdot\mathbf{x}\right)}\right)\right] \\
&= -\frac{1}{4}\frac{\hbar}{(2\pi)^3}\int d^3x\int d^3k\int d^3k'\sqrt{\omega\omega'}\left(\hat{a}(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}-\hat{a}^\dagger(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)\left(\hat{a}(\mathbf{k}')e^{i\left(\omega' t-\mathbf{k}'\cdot\mathbf{x}\right)}-\hat{a}^\dagger(\mathbf{k}')e^{-i\left(\omega' t-\mathbf{k}'\cdot\mathbf{x}\right)}\right) \\
&= -\frac{1}{4}\frac{\hbar}{(2\pi)^3}\int d^3x\int d^3k\int d^3k'\sqrt{\omega\omega'}\left[\hat{a}(\mathbf{k})\hat{a}(\mathbf{k}')e^{i\left(\left(\omega+\omega'\right)t-\left(\mathbf{k}+\mathbf{k}'\right)\cdot\mathbf{x}\right)}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')e^{i\left(\left(\omega-\omega'\right)t-\left(\mathbf{k}-\mathbf{k}'\right)\cdot\mathbf{x}\right)}\right. \\
&\qquad\left.-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}')e^{-i\left(\left(\omega-\omega'\right)t-\left(\mathbf{k}-\mathbf{k}'\right)\cdot\mathbf{x}\right)}+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')e^{-i\left(\left(\omega+\omega'\right)t-\left(\mathbf{k}+\mathbf{k}'\right)\cdot\mathbf{x}\right)}\right]
\end{aligned}$$

The integral over d^3x, produces Dirac delta functions, which we integrate immediately:

$$\begin{aligned}
I_\pi &= -\frac{\hbar}{4}\int d^3k\int d^3k'\sqrt{\omega\omega'}\left[\hat{a}(\mathbf{k})\hat{a}(\mathbf{k}')\delta^3(\mathbf{k}+\mathbf{k}')e^{i\left(\omega+\omega'\right)t}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')\delta^3(\mathbf{k}-\mathbf{k}')e^{i\left(\omega-\omega'\right)t}\right. \\
&\qquad\left.-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}')\delta^3(\mathbf{k}-\mathbf{k}')e^{-i\left(\omega-\omega'\right)t}+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')\delta^3(\mathbf{k}+\mathbf{k}')e^{-i\left(\omega+\omega'\right)t}\right] \\
&= -\frac{\hbar}{4}\int d^3k\,\omega\left[\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}\right]
\end{aligned}$$

We follow the same steps for the remaining two terms in the Hamiltonian. Inserting the gradient of Eq.(4.30), the second term becomes

$$\begin{aligned}
I_{\nabla\varphi} &= \frac{\hbar}{2}\int\boldsymbol{\nabla}\hat{\varphi}\cdot\boldsymbol{\nabla}\hat{\varphi}d^3x \\
&= \frac{\hbar}{2}\frac{1}{(2\pi)^3}\int d^3x\int\frac{d^3k}{\sqrt{2\omega}}\int\frac{d^3k'}{\sqrt{2\omega'}}(-i\mathbf{k})\cdot(-i\mathbf{k}')\left(\hat{a}(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}-\hat{a}^\dagger(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)\left(\hat{a}(\mathbf{k}')e^{i\left(\omega' t-\mathbf{k}'\cdot\mathbf{x}\right)}-\hat{a}^\dagger(\mathbf{k}')e^{-i}\right. \\
&= -\frac{\hbar}{2}\frac{1}{(2\pi)^3}\int d^3x\int\frac{d^3k}{\sqrt{2\omega}}\int\frac{d^3k'}{\sqrt{2\omega'}}\mathbf{k}\cdot\mathbf{k}'\left[\left(\hat{a}(\mathbf{k})\hat{a}(\mathbf{k}')e^{i\left(\left(\omega+\omega'\right)t-\left(\mathbf{k}+\mathbf{k}'\right)\cdot\mathbf{x}\right)}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')e^{i\left(\left(\omega-\omega'\right)t-\left(\mathbf{k}-\mathbf{k}'\right)\cdot\mathbf{x}\right)}\right)\right. \\
&\qquad\left.-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}')e^{-i\left(\left(\omega-\omega'\right)t-\left(\mathbf{k}-\mathbf{k}'\right)\cdot\mathbf{x}\right)}+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')e^{-i\left(\left(\omega+\omega'\right)t-\left(\mathbf{k}+\mathbf{k}'\right)\cdot\mathbf{x}\right)}\right] \\
&= -\frac{\hbar}{2}\int\frac{d^3k}{\sqrt{2\omega}}\int\frac{d^3k'}{\sqrt{2\omega'}}\mathbf{k}\cdot\mathbf{k}'\left[\left(\hat{a}(\mathbf{k})\hat{a}(\mathbf{k}')\delta^3(\mathbf{k}+\mathbf{k}')e^{i\left(\omega+\omega'\right)t}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')\delta^3(\mathbf{k}-\mathbf{k}')e^{i\left(\omega-\omega'\right)t}\right)\right. \\
&\qquad\left.-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}')\delta^3(\mathbf{k}-\mathbf{k}')e^{-i\left(\omega-\omega'\right)t}+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')\delta^3(\mathbf{k}+\mathbf{k}')e^{-i\left(\omega+\omega'\right)t}\right] \\
&= -\frac{\hbar}{4}\int d^3k\,\frac{\mathbf{k}^2}{\omega}\left[-\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})-\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}\right]
\end{aligned}$$

As before, the d^3x integrals of the four terms give four Dirac delta functions and the d^3k' integrals become trivial. It is not hard to see the pattern that is emerging. The $\frac{\mathbf{k}\cdot\mathbf{k}}{\omega}$ term will combine nicely with the ω from the $\hat{\pi}^2$ integral and a corresponding m^2 term from the final integral to give a cancellation. The crucial thing is to keep track of the signs.

The third and final term is

$$\begin{aligned}
I_m &= \frac{\hbar}{2}\int m^2\hat{\varphi}^2 d^3x \\
&= \frac{\hbar}{2}\frac{1}{(2\pi)^3}\int d^3x\int\frac{d^3k}{\sqrt{2\omega}}\int\frac{d^3k'}{\sqrt{2\omega'}}m^2\left(\hat{a}(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}+\hat{a}^\dagger(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)\left(\hat{a}(\mathbf{k}')e^{i(\omega' t-\mathbf{k}'\cdot\mathbf{x})}+\hat{a}^\dagger(\mathbf{k}')e^{-i(\omega' t-\mathbf{k}'\cdot\mathbf{x})}\right) \\
&= \frac{\hbar}{2}\frac{1}{(2\pi)^3}\int d^3x\int\frac{d^3k}{\sqrt{2\omega}}\int\frac{d^3k'}{\sqrt{2\omega'}}m^2\left[\left(\hat{a}(\mathbf{k})\hat{a}(\mathbf{k}')e^{i((\omega+\omega')t-(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x})}+\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')e^{i((\omega-\omega')t-(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x})}\right)\right. \\
&\qquad\left.+\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}')e^{-i((\omega-\omega')t-(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x})}+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')e^{-i((\omega+\omega')t-(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x})}\right] \\
&= \frac{\hbar}{2}\int\frac{d^3k}{\sqrt{2\omega}}\int\frac{d^3k'}{\sqrt{2\omega'}}m^2\left[\left(\hat{a}(\mathbf{k})\hat{a}(\mathbf{k}')\delta^3(\mathbf{k}+\mathbf{k}')e^{i(\omega+\omega')t}+\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')\delta^3(\mathbf{k}-\mathbf{k}')e^{i(\omega-\omega')t}\right)\right. \\
&\qquad\left.+\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}')\delta^3(\mathbf{k}-\mathbf{k}')e^{-i(\omega-\omega')t}+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')\delta^3(\mathbf{k}+\mathbf{k}')e^{-i(\omega+\omega')t}\right] \\
&= \frac{\hbar}{4}\int d^3k\,\frac{m^2}{\omega}\left[\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t}+\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}\right]
\end{aligned}$$

Now we can combine all three terms:

$$\begin{aligned}
\hat{H} &\equiv \frac{\hbar}{2}\int\left(\hat{\pi}^2+\nabla\hat{\varphi}\cdot\nabla\hat{\varphi}+m^2\hat{\varphi}^2\right)d^3x \\
&= I_\pi+I_{\nabla\varphi}+I_m \\
&= -\frac{\hbar}{4}\int d^3k\,\omega\left[\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}\right] \\
&\quad -\frac{\hbar}{4}\int d^3k\,\frac{\mathbf{k}^2}{\omega}\left[-\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t}-\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})-\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})-\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}\right] \\
&\quad +\frac{\hbar}{4}\int d^3k\,\frac{m^2}{\omega}\left[\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t}+\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}\right] \\
&= -\frac{\hbar}{4}\int d^3k\left(\omega-\frac{\mathbf{k}^2}{\omega}-\frac{m^2}{\omega}\right)\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t}-\frac{\hbar}{4}\int d^3k\left(-\omega-\frac{\mathbf{k}^2}{\omega}-\frac{m^2}{\omega}\right)\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}) \\
&\quad -\frac{\hbar}{4}\int d^3k\left(-\omega-\frac{\mathbf{k}^2}{\omega}-\frac{m^2}{\omega}\right)\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})-\frac{\hbar}{4}\int d^3k\left(\omega-\frac{\mathbf{k}^2}{\omega}-\frac{m^2}{\omega}\right)\hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}
\end{aligned}$$

Since

$$\omega^2-\mathbf{k}^2=m^2$$

the Hamiltonian becomes

$$\begin{aligned}
\hat{H} &= -\frac{\hbar}{4}\int d^3k\,\left(-2\omega\right)\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})-\frac{\hbar}{4}\int d^3k\,\left(-2\omega\right)\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}) \\
&= \frac{1}{2}\int d^3k\,\hbar\omega\left(\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})\right)
\end{aligned}$$

If we commute $\hat{a}(\mathbf{k})$ and $\hat{a}^\dagger(\mathbf{k})$ in the first term on the right, we encounter a problem:

$$\begin{aligned}
\hat{H} &= \frac{1}{2}\int d^3k\,\hbar\omega\left(\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})+\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})+\delta^3(\mathbf{k}-\mathbf{k})\right) \\
&= \int d^3k\,\hbar\omega\left(\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})+\frac{1}{2}\delta^3(\mathbf{k}-\mathbf{k})\right)
\end{aligned}$$

This is very close to a sensible result, but the constant term is problematic.

4.2.6 Our first infinity

The form of the Hamiltonian found above displays an obvious problem – the final term,

$$\frac{1}{2}\int d^3k\,\hbar\omega\delta^3(0)$$

diverges in several ways. Most obviously, the triple Dirac delta function is evaluated at 0 and therefore diverges. Even if it were not present, the remaining integral, $\int d^3k\,\omega$, itself diverges.

While the constant "ground state energy" of the harmonic oscillator, $\frac{1}{2}\hbar\omega$, causes no probem in quantum mechanics, the presence of such an energy term for each mode of quantum field theory leads to an infinite energy for the vacuum state. Fortunately, a simple trick allows us to eliminate this divergence throughout our calculations. To see how it works, notice that anytime we have a product of two or more fields at the same point, we develop some terms of the general form

$$\hat{\varphi}(\mathbf{x})\,\hat{\varphi}(\mathbf{x}) \sim \hat{a}(\omega,\mathbf{k})\,\hat{a}^\dagger(\omega,\mathbf{k}) + \ldots$$

which have $\hat{a}^\dagger(\omega,\mathbf{k})$ to the right of $\hat{a}(\omega,\mathbf{k})$. When such products act on the vacuum state, the $\hat{a}^\dagger(\omega,\mathbf{k})$ gives a nonvanishing contribution, and if we sum over all wave vectors we get a divergence. The solution is simply to impose a rule that changes the order of the creation and annihilation operators. This is called *normal ordering*, and is denoted by enclosing the product in colons. Thus, we define

$$:\hat{a}(\omega,\mathbf{k})\,\hat{a}^\dagger(\omega,\mathbf{k}): \;\equiv\; \hat{a}^\dagger(\omega,\mathbf{k})\,\hat{a}(\omega,\mathbf{k})$$

and more generally, normal ordering requires us to place all creation operators to the left of anihilation operators. There is always an ordering ambituity when building functions of $\hat{\varphi}$ and $\hat{\pi}$, since these do not commute. We resolve the ordering ambiguity by writing the function in terms of $\hat{a}(\omega,\mathbf{k})$ and $\hat{a}^\dagger(\omega,\mathbf{k})$ and normal ordering,

$$f(\varphi,\pi) \;\Rightarrow\; :f(\hat{\varphi},\hat{\pi}):$$

Applied to the Hamiltonian, we define

$$\begin{aligned}\hat{H} &= \frac{\hbar}{2}\int :\left(\hat{\pi}^2 + \nabla\hat{\varphi}\cdot\nabla\hat{\varphi} + m^2\hat{\varphi}^2\right): d^3x \\ &= \frac{1}{2}\int d^3k\,\hbar\omega :\left(\hat{a}(\mathbf{k})\,\hat{a}^\dagger(\mathbf{k}) + \hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k})\right):\end{aligned}$$

and this results in

$$\hat{H} = \int d^3k\,\hbar\omega\hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k}) \tag{4.33}$$

This expression gives zero for the vacuum state, and is finite for all states with a finite number of particles. While this procedure may seem a bit ad hoc, recall that the ordering of operators in any quantum expression is one thing that cannot be determined from the classical framework using canonical quantization. It is therefore reasonable to use whatever ordering convention gives the most sensible results.

4.2.7 States of the Klein-Gordon field

The similarity between the field Hamiltonian and the harmonic oscillator makes it easy to interpret this result. We begin the observation that the expectation values of $\hat{H}$ are bounded below. This follows because for *any* normalized state $|\alpha\rangle$ we have

$$\begin{aligned}\langle\alpha|\,\hat{H}\,|\alpha\rangle &= \langle\alpha|\int d^3k\,\hbar\omega\hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k})\,|\alpha\rangle \\ &= \int d^3k\;\hbar\omega\left(\langle\alpha|\,\hat{a}^\dagger(\mathbf{k})\right)\left(\hat{a}(\mathbf{k})\,|\alpha\rangle\right)\end{aligned}$$

This is positive definite, since if we let $|\beta\rangle = \hat{a}(\mathbf{k})\,|\alpha\rangle$, then $\langle\beta| = \langle\alpha|\,\hat{a}^\dagger(\mathbf{k})$, so

$$\begin{aligned}
\langle\alpha|\,\hat{H}\,|\alpha\rangle &= \int d^3k\ \hbar\omega\,\langle\alpha|\,\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})\,|\alpha\rangle \\
&= \int d^3k\ \hbar\omega\,\langle\beta|\beta\rangle \\
&> 0
\end{aligned}$$

since the integrand is positive definite. However, we can show that the action of $\hat{a}\,(\mathbf{k})$ lowers the eigenvalues of $\hat{H}$. Consider the commutator of $\hat{a}\,(\mathbf{k})$ with the Hamiltonian,

$$\begin{aligned}
\left[\hat{H},\hat{a}\,(\mathbf{k})\right] &= \left[\int d^3k'\hbar\omega'\hat{a}^\dagger\,(\mathbf{k}')\,\hat{a}\,(\mathbf{k}')\,,\hat{a}\,(\mathbf{k})\right] \\
&= \int d^3k'\hbar\omega'\left[\hat{a}^\dagger\,(\mathbf{k}')\,,\hat{a}\,(\mathbf{k})\right]\hat{a}\,(\mathbf{k}') \\
&= -\int d^3k'\hbar\omega'\delta^3\,(\mathbf{k}-\mathbf{k}')\,\hat{a}\,(\mathbf{k}') \\
&= -\hbar\omega\,\hat{a}\,(\mathbf{k})
\end{aligned}$$

Therefore, if $|\alpha\rangle$ is an eigenstate of $\hat{H}$ with $\hat{H}\,|\alpha\rangle = \alpha\,|\alpha\rangle$ then so is $\hat{a}(\mathbf{k})\,|\alpha\rangle$ because

$$\begin{aligned}
\hat{H}\,(\hat{a}\,(\mathbf{k})\,|\alpha\rangle) &= \left[\hat{H},\hat{a}\,(\mathbf{k})\right]|\alpha\rangle + \hat{a}\,(\mathbf{k})\,\hat{H}\,|\alpha\rangle \\
&= -\hbar\omega\hat{a}\,(\mathbf{k})\,|\alpha\rangle + \hat{a}\,(\mathbf{k})\,\alpha\,|\alpha\rangle \\
&= (\alpha-\hbar\omega)\,(\hat{a}\,(\mathbf{k})\,|\alpha\rangle)
\end{aligned}$$

Moreover, the eigenvalue of the new eigenstate is *lower* than α. Since the eigenvalues are bounded below, there must exist a state such that

$$\hat{a}\,(\mathbf{k})\,|0\rangle = 0 \tag{4.34}$$

for all values of $\mathbf{k}$. The state $|0\rangle$ is called the *vacuum state* and the operators $\hat{a}\,(\mathbf{k})$ are called annihilation operators. From the vacuum state, we can construct the entire spectrum of eigenstates of the Hamiltonian. First, notice that the vacuum state is a minimal eigenstate of $\hat{H}$:

$$\begin{aligned}
\hat{H}\,|0\rangle &= \int d^3k'\hbar\omega'\hat{a}^\dagger\,(\mathbf{k}')\,\hat{a}\,(\mathbf{k}')\,|0\rangle \\
&= 0
\end{aligned}$$

Now, we act on the vacuum state with $\hat{a}^\dagger(\mathbf{k})$ to produce new eigenstates.

Exercise: Prove that $|\mathbf{k}\rangle = \hat{a}^\dagger(\mathbf{k})\,|0\rangle$ is an eigenstate of $\hat{H}$ with energy eigenvalue $\hbar\omega$.

We can build infinitely many states in two ways. First, just like the harmonic oscillator states, we can apply the creation operator $\hat{a}^\dagger(\mathbf{k})$ as many times as we like. Such a state contains multiple particles with energy $\hbar\omega$. Second, we can apply creation operators of different $\mathbf{k}$:

$$|\mathbf{k}',\mathbf{k}\rangle = \hat{a}^\dagger(\mathbf{k}')\hat{a}^\dagger(\mathbf{k})\,|0\rangle = \hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(\mathbf{k}')\,|0\rangle$$

This state contains two particles, one with energy $\hbar\omega$ and the other with energy $\hbar\omega'$.

As with the harmonic oscillator, we can introduce a number operator to measure the number of quanta in a given state. The number operator is just the sum over all modes of the number operator for a given mode:

$$\begin{aligned}
\hat{N} &= \int :\left(\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})\right): d^3k \\
&= \int \hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})d^3k
\end{aligned}$$

Exercise: By applying $\hat{N}$, compute the number of particles in the state

$$|\mathbf{k}', \mathbf{k}\rangle = \hat{a}^{\dagger}(\mathbf{k}')\hat{a}^{\dagger}(\mathbf{k})|0\rangle$$

Notice that creation and annihilation operators for different modes all commute with one another, e.g.,

$$\left[\hat{a}^{\dagger}(\mathbf{k}'), \hat{a}(\mathbf{k})\right] = 0$$

when $\mathbf{k}' \neq \mathbf{k}$.

4.2.8 Poincaré transformations of Klein-Gordon fields

Now let's examine the Lorentz transformation and translation properties of scalar fields. For this we need to construct quantum operators which generate the required transformations. Since the translations are the simplest, we begin with them.

We have observed that the spacetime translation generators forming a basis for the Lie algebra of translations (and part of the basis of the Poincaré Lie algebra) resemble the energy and momentum operators of quantum mechanics. Moreover, Noether's theorem tells us that energy and momentum are conserved as a result of translation symmetry of the action. We now need to bring these insights into the realm of quantum fields.

From our discussion in Chapter 1, using the Klein-Gordon Lagrangian density from eq.(4.11), we have the conserved stress-energy tensor,

$$\begin{aligned} T^{\mu\nu} &= \frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu}\phi\right)}\partial^{\nu}\phi - \mathcal{L}\eta^{\mu\nu} \\ &= \partial^{\mu}\varphi\partial^{\nu}\varphi - \frac{1}{2}\eta^{\mu\nu}\left(\pi^{2} - \nabla\varphi\cdot\nabla\varphi - m^{2}\varphi^{2}\right) \end{aligned}$$

which leads to the conserved charges,

$$P^{\mu} = \int T^{\mu 0} d^{3}x$$

and the natural extension of this observation is to simply replace the products of fields in $T^{\mu 0}$ with normal-ordered field operators. We therefore write

$$\hat{P}^{\mu} \equiv \int : \hat{T}^{\mu 0} : d^{3}x$$

First, for the time component,

$$\begin{aligned} \hat{P}^{0} &= \int : \hat{T}^{00} : d^{3}x \\ &= \int : \partial^{0}\hat{\varphi}\partial^{0}\hat{\varphi} - \frac{1}{2}\eta^{00}\left(\hat{\pi}^{2} - \nabla\hat{\varphi}\cdot\hat{\varphi} - m^{2}\hat{\varphi}^{2}\right) : d^{3}x \\ &= \frac{1}{2}\int : \hat{\pi}^{2} + \nabla\hat{\varphi}\cdot\nabla\hat{\varphi} + m^{2}\hat{\varphi}^{2} : d^{3}x \\ &= \hat{H} \end{aligned}$$

This is promising!

Now consider the momentum operators:

$$\begin{aligned}\hat{P}^i &\equiv \int :\hat{T}^{i0}: d^3x \\ &= \int :\partial^i\hat{\varphi}\partial^0\hat{\varphi} - \frac{1}{2}\eta^{i0}\left(\hat{\pi}^2 - \nabla\hat{\varphi}\cdot\hat{\varphi} - m^2\hat{\varphi}^2\right): d^3x \\ &= \int :\partial^i\hat{\varphi}\;\hat{\pi}: d^3x \\ \hat{\mathbf{P}} &= \int :\boldsymbol{\nabla}\hat{\varphi}\;\hat{\pi}: d^3x\end{aligned}$$

Exercise: By substituting the field operators, eq.(4.30) and eq.(4.31), into the integral for $\hat{P}^i$, show that

$$\hat{\mathbf{P}} = \frac{1}{2}\int \hbar\mathbf{k} : \left[-\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t} + \hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}) + \hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}) - \hat{a}^\dagger(\mathbf{k})\hat{a}^\dagger(-\mathbf{k})e^{-2i\omega t}\right] : d^3k$$

The calculation is similar to the computation of the Hamiltonian operator above, except there is only one term to consider.

We can simplify this result for $\hat{\mathbf{P}}$ using a parity argument. Consider the effect of parity on the first integral. Since the volume form together with the limits is invariant under $\mathbf{k}\to-\mathbf{k}$,

$$\iiint_{-\infty}^{\infty} d^3k \quad\to\quad \iiint_{\infty}^{-\infty} (-1)^3\, d^3k = \iiint_{-\infty}^{\infty} d^3k$$

and $\omega(-\mathbf{k}) = \omega(\mathbf{k})$, the first integral satisfies

$$\begin{aligned}\hat{\mathbf{I}}_1 &= \frac{1}{2}\int d^3k\;\hbar\mathbf{k}\;\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t} \\ &= \frac{1}{2}\int d^3k\;\left(-\hbar\mathbf{k}\right)\hat{a}(-\mathbf{k})\hat{a}(\mathbf{k})e^{2i\omega t} \\ &= -\frac{1}{2}\int d^3k\;\hbar\mathbf{k}\;\hat{a}(\mathbf{k})\hat{a}(-\mathbf{k})e^{2i\omega t} \\ &= -\hat{\mathbf{I}}_1\end{aligned}$$

and therefore $\hat{\mathbf{I}}_1 = 0$. The final term vanishes in the same way, so the momentum operator reduces to

$$\begin{aligned}\hat{\mathbf{P}} &= \int :\partial^i\hat{\varphi}\;\hat{\pi}: d^3x \\ &= \frac{1}{2}\int \hbar\mathbf{k} : \left(\hat{a}(\mathbf{k})\hat{a}^\dagger(\mathbf{k}) + \hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})\right) : d^3k \\ &= \int \hbar\mathbf{k}\;\hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})d^3k\end{aligned}$$

Once again, this makes sense; moreover, they are suitable for translation generators since they all commute.

In a similar way, we can compute the operators $\hat{M}^{\alpha\beta}$, and show that the commutation relations of the full set reproduce the Poincaré Lie algebra,

$$\begin{aligned}\left[\hat{M}^{\alpha\beta}, \hat{M}^{\mu\nu}\right] &= \eta^{\beta\mu}\hat{M}^{\alpha\nu} - \eta^{\beta\nu}\hat{M}^{\alpha\mu} - \eta^{\alpha\mu}\hat{M}^{\beta\nu} - \eta^{\alpha\nu}\hat{M}^{\beta\mu} \\ \left[\hat{M}^{\alpha\beta}, \hat{P}^{\mu}\right] &= \eta^{\mu\alpha}\hat{P}^{\beta} - \eta^{\mu\beta}\hat{P}^{\alpha} \\ \left[\hat{P}^{\alpha}, \hat{P}^{\beta}\right] &= 0\end{aligned}$$

The notable accomplishment here is that we have shown that even after quantization, the symmetry algebra not only survives, but can be built from the quantum field operators. This is far from obvious, because the commutation relations for the field operators are simply imposed by the rules of canonical quantization and have nothing to do, a priori, with the commutators of the symmetry algebra. One consequence, as noted above, is that the Casimir operators of the Poincaré algebra may be used to label *quantum* states.

4.3 Quantization of the complex scalar field

4.3.1 Classical Hamiltonian formulation

The complex scalar field provides a slight generalization of the real scalar field. As before we begin with the Lagrangian, Eq.(4.2)

$$L = \int \left(\partial^\alpha \varphi^* \partial_\alpha \varphi - m^2 \varphi^* \varphi\right) d^3x \tag{4.35}$$

This has twice the degrees of freedom as the real Klein-Gordon field, and introduces an extra symmetry. While we could realize the two degrees of freedom by expanding $\varphi = \varphi_R + i\varphi_I$, treating φ and φ^* as the independent variables yields the same results.

We define the conjugate momentum densities to each of φ and φ^* as the functional derivatives L with respect to φ and φ^* :

$$\pi \equiv \frac{\delta L}{\delta\left(\partial_0 \varphi\right)} = \partial^0 \varphi^* (\mathbf{x}) \tag{4.36}$$

and similarly

$$\pi^* \equiv \frac{\delta L}{\delta\left(\partial_0 \varphi^*\right)} = \partial^0 \varphi (\mathbf{x}) \tag{4.37}$$

The action and Lagrangian density, written in terms of these momenta, are therefore

$$\begin{aligned} S &= \int \left(\pi\pi^* - \nabla\varphi^* \cdot \nabla\varphi - m^2 \varphi^* \varphi\right) d^4x \\ \mathcal{L} &= \pi\pi^* - \nabla\varphi^* \cdot \nabla\varphi - m^2 \varphi^* \varphi \end{aligned}$$

The Hamiltonian is defined as

$$\begin{aligned} H &\equiv \int \left(\pi \partial_0 \varphi + \pi^* \partial_0 \varphi^*\right) d^3x - L \\ &= \int \left(\pi\pi^* + \pi^*\pi\right) - \left(\pi\pi^* - \nabla\varphi^* \cdot \nabla\varphi - m^2 \varphi^* \varphi\right) d^3x \end{aligned}$$

and therefore

$$H = \int \left(\pi^*\pi + \nabla\varphi^* \cdot \nabla\varphi + m^2 \varphi^* \varphi\right) d^3x \tag{4.38}$$

Hamilton's equations are:

$$\begin{aligned} \dot{\varphi}(\mathbf{x}) &= \frac{\delta H}{\delta \pi (\mathbf{x})} \\ \dot{\pi}(\mathbf{x}) &= -\frac{\delta H}{\delta \varphi (\mathbf{x})} \\ \dot{\varphi}^*(\mathbf{x}) &= \frac{\delta H}{\delta \pi^* (\mathbf{x})} \\ \dot{\pi}^*(\mathbf{x}) &= -\frac{\delta H}{\delta \varphi^* (\mathbf{x})} \end{aligned}$$

Exercise: Prove that Hamilton's equations reproduce the field equations for φ and φ^*.

Now write the field equations in terms of functional Poisson brackets which – remembering to sum the derivatives over all independent fields – are given for functionals $f = f[\varphi, \pi, \varphi^*, \pi^*]$ and $g = g[\varphi, \pi, \varphi^*, \pi^*]$ by

$$\{f, g\} \equiv \int d^3x \left(\frac{\delta f}{\delta \pi(\mathbf{x})} \frac{\delta g}{\delta \varphi(\mathbf{x})} + \frac{\delta f}{\delta \pi^*(\mathbf{x})} \frac{\delta g}{\delta \varphi^*(\mathbf{x})} - \frac{\delta f}{\delta \varphi(\mathbf{x})} \frac{\delta g}{\delta \pi(\mathbf{x})} - \frac{\delta f}{\delta \varphi^*(\mathbf{x})} \frac{\delta g}{\delta \pi^*(\mathbf{x})} \right) \tag{4.39}$$

The result is the just what we would guess from the real case,

$$\begin{aligned}
\{\pi(\mathbf{x}), \varphi(\mathbf{y})\} &= \int \left(\frac{\delta \pi(\mathbf{x})}{\delta \pi(\mathbf{x}')} \frac{\delta \varphi(\mathbf{y})}{\delta \varphi(\mathbf{x}')} + 0 - \frac{\delta \pi(\mathbf{x}))}{\delta \varphi(\mathbf{x}')} \frac{\delta \varphi(\mathbf{y})}{\delta \pi(\mathbf{x}')} - 0 \right) d^3x \\
&= \int \delta^3(\mathbf{x} - \mathbf{x}') \, \delta^3(\mathbf{y} - \mathbf{x}) \, d^3x \\
&= \delta^3(\mathbf{x} - \mathbf{y}) \\
\{\pi^*(\mathbf{x}), \varphi^*(\mathbf{y})\} &= \delta^3(\mathbf{x} - \mathbf{y})
\end{aligned}$$

with all other brackets vanishing.

Exercise: Check that Hamilton's equations

$$\begin{aligned}
\dot{\varphi}(\mathbf{x}) &= \{H(\varphi, \pi, \varphi^*, \pi^*), \varphi(\mathbf{x}')\} \\
\dot{\varphi}^*(\mathbf{x}) &= \{H(\varphi, \pi, \varphi^*, \pi^*), \varphi^*(\mathbf{x}')\} \\
\dot{\pi}(\mathbf{x}) &= \{H(\varphi, \pi, \varphi^*, \pi^*), \pi(\mathbf{x}')\} \\
\dot{\pi}^*(\mathbf{x}) &= \{H(\varphi, \pi, \varphi^*, \pi^*), \pi^*(\mathbf{x}')\}
\end{aligned}$$

reproduce Hamilton's equations.

Now we quantize, replacing fields by operators and Poisson brackets by equal-time commutators:

$$[\hat{\pi}(\mathbf{x}, t), \hat{\varphi}(\mathbf{y}, t)] = i\hbar \delta^3(\mathbf{x} - \mathbf{y}) \tag{4.40}$$

$$\left[\hat{\pi}^\dagger(\mathbf{x}, t), \hat{\varphi}^\dagger(\mathbf{y}, t)\right] = i\hbar \delta^3(\mathbf{x} - \mathbf{y}) \tag{4.41}$$

with all other pairs commuting. Now we seek free field solutions satisfying these quantization relations.

4.3.2 Mode amplitudes of the complex scalar field

The solution proceeds as before, by starting with solutions for the classical theory. The field equations

$$\begin{aligned}
\Box \varphi &= -\frac{m^2}{\hbar^2} \varphi \\
\Box \varphi^* &= -\frac{m^2}{\hbar^2} \varphi^*
\end{aligned}$$

are complex conjugates of each other. The only difference from the real case is that we no longer restrict to real plane waves. This leaves the amplitudes independent:

$$\varphi(\mathbf{x}, t) = a e^{\frac{i}{\hbar}(Et - \mathbf{p}\cdot\mathbf{x})} + b^* e^{-\frac{i}{\hbar}(Et - \mathbf{p}\cdot\mathbf{x})} \tag{4.42}$$

Substituting into the field equation we have

$$\begin{aligned}
\Box \varphi &= \left(\frac{i}{\hbar}\right)^2 p_\alpha p^\alpha a e^{\frac{i}{\hbar}(Et - \mathbf{p}\cdot\mathbf{x})} + \left(-\frac{i}{\hbar}\right)^2 p_\alpha p^\alpha b^* e^{-\frac{i}{\hbar}(Et - \mathbf{p}\cdot\mathbf{x})} \\
&= -\frac{1}{\hbar^2} p_\alpha p^\alpha \varphi(\mathbf{x}, t)
\end{aligned}$$

so again we require the energy condition

$$p_\alpha p^\alpha = m^2$$

We can solve this for the energy,

$$\begin{aligned} E_+ &= \sqrt{\mathbf{p}^2 + m^2} \\ E_- &= -\sqrt{\mathbf{p}^2 + m^2} \end{aligned}$$

The general Fourier superposition is

$$\begin{aligned} \varphi(\mathbf{x}, t) &= \frac{1}{(2\pi)^{3/2}} \int \sqrt{2E} \left(a(E, \mathbf{p}) e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + b^*(E, \mathbf{p}) e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)} \right) \delta\left(p_\alpha p^\alpha - m^2\right) \Theta(E) \hbar^{-4} d^4 p \\ &= \frac{1}{(2\pi)^{3/2}} \int \frac{d^3 k}{\sqrt{2\omega}} \left(a(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} + b^*(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \end{aligned}$$

Collecting this together with the the conjugate field and the momenta,

$$\varphi(\mathbf{x}, t) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3 k}{\sqrt{2\omega}} \left(a(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} + b^\dagger(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{4.43}$$

$$\varphi^*(\mathbf{x}, t) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3 k}{\sqrt{2\omega}} \left(b(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} + a^\dagger(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{4.44}$$

$$\pi(\mathbf{x}, t) = \frac{i}{(2\pi)^{3/2}} \int \sqrt{\frac{\omega}{2}} d^3 k \left(b(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} - a^\dagger(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{4.45}$$

$$\pi^*(\mathbf{x}, t) = \frac{i}{(2\pi)^{3/2}} \int \sqrt{\frac{\omega}{2}} d^3 k \left(a(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} - b^\dagger(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{4.46}$$

Notice that we may obtain the conjugate expressions, Eqs.(4.44) and (4.46) simply by interchanging $a$with b, and interchanging $a^\dagger$ with $b^\dagger$.

We need to invert these Fourier integrals to solve for $a(\mathbf{k}), b(\mathbf{k}), a^\dagger(\mathbf{k})$ and $b^\dagger(\mathbf{k})$.

Exercise: By taking inverse Fourier integrals, show that

$$a(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{\omega}{2}} \int d^3 x \left(\varphi(\mathbf{x}, 0) - \frac{i}{\omega} \pi^*(\mathbf{x}, 0) \right) e^{i\mathbf{k}\cdot\mathbf{x}} \tag{4.47}$$

$$b(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{\omega}{2}} \int d^3 x \left(\varphi^*(\mathbf{x}, 0) - \frac{i}{\omega} \pi(\mathbf{x}, 0) \right) e^{i\mathbf{k}\cdot\mathbf{x}} \tag{4.48}$$

It follows immediately from this exercise that the conjugate mode amplitudes are given by

$$a^*(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{\omega}{2}} \int d^3 x \left(\varphi^*(\mathbf{x}, 0) + \frac{i}{\omega} \pi(\mathbf{x}, 0) \right) e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{4.49}$$

$$b^*(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{\omega}{2}} \int d^3 x \left(\varphi(\mathbf{x}, 0) + \frac{i}{\omega} \pi^*(\mathbf{x}, 0) \right) e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{4.50}$$

4.3.3 Quantization

We can now move to study the quantum operators. When the fields become operators the complex conjugates above become adjoints (for example, $a^*(\mathbf{k}) \to a^\dagger(\mathbf{k})$). We next find the commutation relations that hold among the four operators $\hat{a}(\mathbf{k}), \hat{b}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k})$ and $\hat{b}^\dagger(\mathbf{k})$.

Exercise: From the commutation relations for the fields and conjugate momenta, Eqs.(4.40) and (4.41), show that

$$\begin{aligned}\left[\hat{a}(\mathbf{k}), \hat{a}^{\dagger}(\mathbf{k}')\right] &= \delta^3(\mathbf{k}-\mathbf{k}') \\ \left[\hat{b}(\mathbf{k}), \hat{b}^{\dagger}(\mathbf{k}')\right] &= \delta^3(\mathbf{k}-\mathbf{k}')\end{aligned}$$

Exercise: From the commutation relations for the fields and conjugate momenta, eqs.(4.40) and (4.41), show that

$$\begin{aligned}\left[\hat{a}(\mathbf{k}), \hat{b}(\mathbf{k}')\right] &= 0 \\ \left[\hat{a}(\mathbf{k}), \hat{b}^{\dagger}(\mathbf{k}')\right] &= 0\end{aligned}$$

As we did for for the Klein-Gordon field, we could go on to construct the Poincaré currents, writing the energy, momentum and angular momentum in terms of the creation and anihilation operators. These emerge much as before. However, for the charged scalar field, there is an additional symmetry.

Exercise: Find the Hamiltonian operator

$$\hat{H} = \int :\left(\pi^*\pi + \boldsymbol{\nabla}\varphi^* \cdot \boldsymbol{\nabla}\hat{\varphi} + m^2\hat{\varphi}^*\hat{\varphi}\right): d^3x$$

in terms of the creation and annihilation operators.

4.3.4 Noether current and current operator

The transformation

$$\begin{aligned}\varphi(\mathbf{x},t) &\rightarrow e^{i\alpha}\varphi(\mathbf{x},t) \\ \varphi^*(\mathbf{x},t) &\rightarrow e^{-i\alpha}\varphi^*(\mathbf{x},t)\end{aligned} \tag{4.51}$$

leaves the action, Eq.(4.2), invariant, so the complex scalar field has a global $U(1)$ symmetry. Therefore, there is an additional Noether current. In this case, the variation of the Lagrangian, Eq.(4.35), under the $U(1)$ symmetry is also zero, so from eq.(1.35) the Noether current is simply

$$J^{\mu} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\phi^A)}\Delta^A$$

where

$$\phi^A \rightarrow \phi^A + \Delta^A\left(\phi^B, x\right)$$

defines the infinitesimal transformation Δ^A. For an infinitesimal phase change, $e^{i\alpha} \approx 1 + i\alpha$ so the fields change by

$$\begin{aligned}\varphi &\rightarrow \varphi + i\alpha\varphi \\ \varphi^* &\rightarrow \varphi^* - i\alpha\varphi^*\end{aligned}$$

so the current is

$$\begin{aligned}J^{\alpha} &\equiv \frac{\partial \mathcal{L}}{\partial(\partial_{\alpha}\phi)}\Delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_{\alpha}\phi^*)}\Delta\varphi^* \\ &= (\partial^{\alpha}\varphi^*)\, i\alpha\varphi - (\partial^{\alpha}\varphi)\, i\alpha\varphi^* \\ &= i\alpha\left((\partial^{\alpha}\varphi^*)\varphi - (\partial^{\alpha}\varphi)\varphi^*\right)\end{aligned} \tag{4.52}$$

We are guaranteed that the divergence of J^α must vanish and can easily check using the field equations:

$$\begin{aligned}
\partial_\alpha J^\alpha &= i\alpha\partial_\alpha\left(\left(\partial^\alpha\varphi^*\right)\varphi - \left(\partial^\alpha\varphi\right)\varphi^*\right) \\
&= i\alpha\left(\left(\partial_\alpha\partial^\alpha\varphi^*\right)\varphi - \left(\partial^\alpha\varphi\right)\left(\partial_\alpha\varphi^*\right) + \left(\partial^\alpha\varphi^*\right)\left(\partial_\alpha\varphi\right) - \left(\partial_\alpha\partial^\alpha\varphi\right)\varphi^*\right) \\
&= -i\alpha\left(\frac{m^2}{\hbar^2}\varphi^*\varphi - \frac{m^2}{\hbar^2}\varphi\varphi^*\right) \\
&= 0
\end{aligned}$$

In general, when new fields are introduced to make a global symmetry into a local symmetry, the new fields produce interactions between the original, symmetric fields. The strength of this interaction is governed by the Noether currents of the symmetry. In the present case, when this $U(1)$ (phase) invariance is gauged to produce an interaction, the new field that is introduced is the photon field, and it is this current J^α that carries the electric charge. Therefore, writing e for α, and writing the 4-current as $J^\alpha = (\rho, \mathbf{J})$, we see that

$$\rho = ie\left(\partial_0\varphi^*\varphi - \partial_0\varphi\varphi^*\right) \tag{4.53}$$

$$\mathbf{J} = ie\left(\varphi\boldsymbol{\nabla}\varphi^* - \varphi^*\boldsymbol{\nabla}\varphi\right) \tag{4.54}$$

4.3.5 Conserved charge operator

Classically, the spatial integral of the charge density ρ gives us conserved charge,

$$\begin{aligned}
Q &= \int J^0 d^3x \\
&= \int \rho d^3x
\end{aligned}$$

While all of the current may be expressed in terms of operators on quantum states, we will be particuarly interested in the total charge. Substituting the operator expressions for the fields, we find that the conserved charge is given by

$$\begin{aligned}
\hat{Q} &= \int :\hat{\rho}: d^3x \\
&= ie\int :\left(\partial_0\hat{\varphi}^*\hat{\varphi} - \partial_0\hat{\varphi}\hat{\varphi}^*\right): d^3x \\
&= ie\int\left(\hat{\pi}\hat{\varphi} - \hat{\pi}^*\hat{\varphi}^*\right)d^3x
\end{aligned}$$

Substituting the fields from Eqs.(4.43) - (4.46), this becomes

$$\begin{aligned}
\hat{Q} &= -\frac{e}{(2\pi)^3}\int d^3x : \left(\int\sqrt{\frac{\omega}{2}}d^3k\left(\hat{b}(\mathbf{k})e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} - \hat{a}^\dagger(\mathbf{k})e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})}\right)\right)\left(\int\frac{d^3k'}{\sqrt{2\omega'}}\left(\hat{a}(\mathbf{k}')e^{i(\omega' t - \mathbf{k}'\cdot\mathbf{x})} + \hat{b}^\dagger(\mathbf{k}')e^{-i(\omega' t - \mathbf{k}'\cdot\mathbf{x})}\right.\right. \\
&\quad - \left(\hat{a} \leftrightarrow \hat{b}\ and\ \hat{a}^\dagger \leftrightarrow \hat{b}^\dagger\right) \\
&= -\frac{e}{2(2\pi)^3}\int d^3x\int d^3k\int d^3k'\sqrt{\frac{\omega}{\omega'}} : \left(\hat{b}(\mathbf{k})e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} - \hat{a}^\dagger(\mathbf{k})e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})}\right)\left(\hat{a}(\mathbf{k}')e^{i(\omega' t - \mathbf{k}'\cdot\mathbf{x})} + \hat{b}^\dagger(\mathbf{k}')e^{-i(\omega' t - \mathbf{k}'\cdot\mathbf{x})}\right) : \\
&\quad - \left(\hat{a} \leftrightarrow \hat{b}\ and\ \hat{a}^\dagger \leftrightarrow \hat{b}^\dagger\right) \\
&= -\frac{e}{2(2\pi)^3}\int d^3x\int d^3k\int d^3k'\sqrt{\frac{\omega}{\omega'}} : \left[\left(\hat{b}(\mathbf{k})\hat{a}(\mathbf{k}')e^{i((\omega+\omega')t - (\mathbf{k}+\mathbf{k}')\cdot\mathbf{x})} - \hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}')e^{-i((\omega-\omega')t - (\mathbf{k}-\mathbf{k}')\cdot\mathbf{x})}\right)\right. \\
&\qquad \left. + \hat{b}(\mathbf{k})\hat{b}^\dagger(\mathbf{k}')e^{i((\omega-\omega')t - (\mathbf{k}-\mathbf{k}')\cdot\mathbf{x})} - \hat{a}^\dagger(\mathbf{k})\hat{b}^\dagger(\mathbf{k}')e^{-i((\omega+\omega')t - (\mathbf{k}+\mathbf{k}')\cdot\mathbf{x})}\right] : \\
&\quad - \left(\hat{a} \leftrightarrow \hat{b}\ and\ \hat{a}^\dagger \leftrightarrow \hat{b}^\dagger\right)
\end{aligned}$$

Integrate over d^3x, giving Dirac delta functions, $\delta(\mathbf{k}+\mathbf{k}')$ or $\delta(\mathbf{k}-\mathbf{k}')$, then integrate over d^3k':

$$\begin{aligned}
\hat{Q} &= -\frac{e}{2}\int d^3k \int d^3k' \sqrt{\frac{\omega}{\omega'}} : \Big[\Big(\hat{b}(\mathbf{k})\,\hat{a}(\mathbf{k}')\,\delta(\mathbf{k}+\mathbf{k}')\,e^{i(\omega+\omega')t} - \hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k}')\,\delta(\mathbf{k}-\mathbf{k}')\,e^{-i(\omega-\omega')t}\Big) \\
&\qquad + \hat{b}(\mathbf{k})\,\hat{b}^\dagger(\mathbf{k}')\,\delta(\mathbf{k}-\mathbf{k}')\,e^{i(\omega-\omega')t} - \hat{a}^\dagger(\mathbf{k})\,\hat{b}^\dagger(\mathbf{k}')\,\delta(\mathbf{k}+\mathbf{k}')\,e^{-i(\omega+\omega')t}\Big] : \\
&\quad - \Big(\hat{a}\leftrightarrow\hat{b} \; and \; \hat{a}^\dagger\leftrightarrow\hat{b}^\dagger\Big) \\
&= -\frac{e}{2}\int d^3k : \Big[\hat{b}(\mathbf{k})\,\hat{a}(-\mathbf{k})\,e^{2i\omega t} - \hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k}) + \hat{b}(\mathbf{k})\,\hat{b}^\dagger(\mathbf{k}) - \hat{a}^\dagger(\mathbf{k})\,\hat{b}^\dagger(-\mathbf{k})\,e^{-2i\omega t}\Big] : \\
&\quad +\frac{e}{2}\int d^3k : \Big[\hat{a}(\mathbf{k})\,\hat{b}(-\mathbf{k})\,e^{2i\omega t} - \hat{b}^\dagger(\mathbf{k})\,\hat{b}(\mathbf{k}) + \hat{a}(\mathbf{k})\,\hat{a}^\dagger(\mathbf{k}) - \hat{b}^\dagger(\mathbf{k})\,\hat{a}^\dagger(-\mathbf{k})\,e^{-2i\omega t}\Big] :
\end{aligned}$$

Noticing that changing variable $\mathbf{k}\to-\mathbf{k}$ produces

$$\iiint_{-\infty}^{\infty} d^3k\,\hat{b}(\mathbf{k})\,\hat{a}(-\mathbf{k})\,e^{2i\omega t} = \iiint_{-\infty}^{\infty} d^3k\,\hat{b}(-\mathbf{k})\,\hat{a}(\mathbf{k})\,e^{2i\omega t}$$

shows that the two $e^{2i\omega t}$ terms cancel, as do the final two $e^{-2i\omega t}$ terms. This leaves

$$\hat{Q} = -\frac{e}{2}\int d^3k : \Big[-\hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k}) + \hat{b}(\mathbf{k})\,\hat{b}^\dagger(\mathbf{k}) + \hat{b}^\dagger(\mathbf{k})\,\hat{b}(\mathbf{k}) - \hat{a}(\mathbf{k})\,\hat{a}^\dagger(\mathbf{k})\Big] :$$

Normal ordering, we have

$$\hat{Q} = e\int d^3k\,\Big[\hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k}) - \hat{b}^\dagger(\mathbf{k})\,\hat{b}(\mathbf{k})\Big]$$

Writing this in terms of number operators gives a new insight. Define

$$\begin{aligned}
\hat{N}_a(\mathbf{k}) &\equiv \hat{a}^\dagger(\mathbf{k})\,\hat{a}(\mathbf{k}) \\
\hat{N}_b(\mathbf{k}) &\equiv \hat{b}^\dagger(\mathbf{k})\,\hat{b}(\mathbf{k})
\end{aligned}$$

and acting on various states we find that these count the number of a-type and b-type particles at any given $\mathbf{k}$, respectively. If we integrate over $\mathbf{k}$ we find the total number of a-type and b-type particles in a state.

In terms of number operators, the charge operator is

$$\hat{Q} = \int d^3k\,\Big[e\hat{N}_a(\mathbf{k}) - e\hat{N}_b(\mathbf{k})\Big] \tag{4.55}$$

so the a and b-type particles have opposite charge.

It proves to be of some importance that the charge e appears as the phase of the $U(1)$ symmetry transformation. This means that complex conjugation has the effect of changing the signs of all charges. This *charge conjugation symmetry* is one of the central discrete symmetries associated with the Lorentz group, and it plays a role when we consider the meaning of antiparticles later in this chapter. Notice, in particular, in the solution for the complex scalar field, eq.(4.43), that the phase of the antiparticle is just reversed from the phase for the particle.

4.4 Scalar multiplets

Suppose we have n scalar fields, φ^i, $i=1,\ldots,n$ governed by the action

$$S = \frac{1}{2}\int \sum \left(\partial^\alpha\varphi^i\partial_\alpha\varphi^i - m^2\varphi^i\varphi^i\right) d^4x$$

The quantization is similar to the previous cases. We find the conjugate momenta, $\pi^i = \frac{\delta L}{\delta \dot{\varphi}^i} = \dot{\varphi}^i$ and the Hamiltonian is

$$H = \frac{1}{2}\int \left(\pi^i\pi^i + \nabla\varphi^i\cdot\nabla\varphi^i + m^2\varphi^i\varphi^i\right) d^3x$$

The fundamental commutation relations are

$$\left[\hat{\pi}^i(\mathbf{x},t), \hat{\varphi}^j(\mathbf{x}',t)\right] = i\delta^{ij}\delta^3\left(\mathbf{x}-\mathbf{x}'\right)$$

with all others vanishing. These lead to creation and annihilation operators as before,

$$\left[\hat{a}^i(\mathbf{k}), \hat{a}^{j\dagger}(\mathbf{k}')\right] = \delta^{ij}\delta^3\left(\mathbf{k}-\mathbf{k}'\right)$$

and a number operator for each field,

$$\hat{N}(\mathbf{k}) = \hat{a}^{i\dagger}(\mathbf{k})\hat{a}^i(\mathbf{k})$$

The interesting feature of this case is the presence of a more general symmetry. The action S is left invariant by orthogonal rotations of the fields into one another. Thus, if $O^i{}_j$ is an orthogonal transformation, we can define new fields

$$\varphi^{i\prime} = O^i{}_j\varphi^i$$

It is easy to see that the action is unchanged by such a transformation. For each infinitesimal generator of a rotation, $\left[\varepsilon_{(rs)}\right]^{ij} = \frac{1}{2}\left(\delta^i_r\delta^j_s - \delta^i_s\delta^j_r\right)$, there is a conserved Noether current found from the infinitesimal transformation,

$$\varphi^i \rightarrow \varphi^i + \left[\varepsilon_{(rs)}\right]^{ij}\varphi^j$$

Since the Lagrangian is invariant, the current is

$$\begin{aligned} J^\alpha_{(rs)} &\equiv \frac{\partial\mathcal{L}}{\partial\left(\partial_\alpha\phi^i\right)}\Delta_{(rs)}\varphi^i \\ &= \partial^\alpha\varphi^i\left[\varepsilon_{(rs)}\right]^{ij}\varphi^j \\ &= \varphi^r\partial^\alpha\varphi^s - \varphi^s\partial^\alpha\varphi^r \end{aligned}$$

We are guaranteed that the divergence of J^μ vanishes when the field equations are satisfied.

Chapter 5

Antiparticles

Until this section we have dodged the issue of the negative energy solutions to scalar field theories by inserting a step function, $\Theta(E)$, in the Fourier series for the solution. Now let's consider these in more detail. We will see that the negative energy states may be interpreted as antiparticles. While the discussion applies to all fields we consider, it is simplest to look at the real scalar field. The same considerations apply to the complex and multiplet fields.

5.1 Green functions for scalar fields

To begin, let's look at sources for an interacting scalar field. For example, consider a term in the particle action that couples a scalar field to a spinor field. One possible action is

$$S = \frac{1}{2}\int d^4x \left(\partial^\alpha \phi \partial_\alpha \phi - m^2\phi^2 - 2\phi\bar{\psi}\psi\right) \tag{5.1}$$

In this simple case, the spinor field provides a source for the scalar field. We neeed not consider the dynamics of the spinor fields. The field equation for ϕ is then

$$\Box\phi + m^2\phi = -J \tag{5.2}$$

where $J = \bar{\psi}\psi$. For our purposes it is sufficient to consider solutions to equations of the general form given in eq.(5.2).

To solve eq.(5.2), we use Green's theorem. For a complete treatment of the method, see e.g., Jackson or Arfken. Simply put, if we can first solve

$$\left(\Box + m^2\right) G\left(x, x'\right) = -\delta^4\left(x - x'\right) \tag{5.3}$$

for a function $G\left(x, x'\right)$ satisfying the relevant boundary conditions, then linearity of Eq.(5.2) shows that the complete solution is

$$\phi\left(x\right) = \int d^4x' G\left(x, x'\right) J\left(x'\right)$$

for the same boundary conditions when the source is $J(x)$. To check, apply the Klein-Gordon operator to $\phi(x)$

$$\left(\Box + m^2\right)\phi\left(x\right) = \left(\Box + m^2\right)\int d^4x' G\left(x, x'\right) J\left(x'\right)$$

Since $\left(\Box + m^2\right)$ depends on x and the integral is over x', we may bring the operator inside the integral:

$$\begin{aligned}
\left(\Box + m^2\right)\phi(x) &= \int d^4x' \left(\Box + m^2\right) G\left(x, x'\right) J\left(x'\right) \\
&= -\int d^4x' \delta^4\left(x - x'\right) J\left(x'\right) \\
&= -J\left(x\right)
\end{aligned}$$

The technique is successful because a solution to eq.(5.3) may be built from solutions to the homogeneous equation.

To find the Green function $G\left(x, x'\right)$ explicitly when the boundary conditions are at infinity, we again use a Fourier series. Write

$$\begin{aligned}
G\left(x, x'\right) &= \frac{1}{(2\pi)^4} \int d^4k\, \tilde{G}\left(k\right) e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)} \\
&= \frac{1}{(2\pi)^4} \int d^4k\, \tilde{G}\left(k\right) e^{i\left(k_0\left(t-t'\right) - \mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)\right)} \\
\delta^4\left(x - x'\right) &= \frac{1}{(2\pi)^4} \int d^4k e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)}
\end{aligned}$$

Then, substituting into Eq.(5.3) and cancelling the overall factor of $(2\pi)^4$,

$$\begin{aligned}
\left(\Box + m^2\right) \int d^4k\, \tilde{G}\left(k\right) e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)} &= -\int d^4k e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)} \\
\int d^4k\, \tilde{G}\left(k\right)\left(k\right)\left(\Box + m^2\right) e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)} &= -\int d^4k e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)} \\
\int d^4k \left(\tilde{G}\left(k\right)\left(-k_\alpha k^\alpha + m^2\right) + 1\right) e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)} &= 0 \qquad (5.4)
\end{aligned}$$

Eq.(5.4) is the Fourier transform of the factor in parentheses. Since the Fourier integral is invertible, we must have

$$\tilde{G}\left(k\right) = \frac{1}{k_\alpha k^\alpha - m^2}$$

Now invert the Fourier transform to find the Green function:

$$G\left(x, x'\right) = \frac{1}{(2\pi)^4} \int d^4k \frac{e^{ik_\alpha\left(x^\alpha - x'^\alpha\right)}}{\left(k^0\right)^2 - \mathbf{k}^2 - m^2} \qquad (5.5)$$

The interesting feature here is the divergence when the denominator vanishes. To compute it we resort to a contour integral and the residue theorem.

The poles are given by factoring the divergent factor as

$$\frac{1}{\left(k^0\right)^2 - \mathbf{k}^2 - m^2} = \frac{1}{2k_0}\left(\frac{1}{k^0 - \sqrt{\mathbf{k}^2 + m^2}} + \frac{1}{k^0 + \sqrt{\mathbf{k}^2 + m^2}}\right) \qquad (5.6)$$

The poles lie on the real axis in the complex k^0 plane, but we can displace the poles slightly by replacing $k_0 \to k_0 + i\varepsilon$ or $k_0 \to k_0 - i\varepsilon$. The direction we push the pole depends on the boundary conditions we want to impose.

5.1.1 First pole

Consider the various alternatives. For each of the two simple poles we have two choices, so there are four possible contributions to the Green function. We compute them in turn. The first pole occurs when

$$k^0 = +\sqrt{\mathbf{k}^2+m^2}$$

Displacing this point leads to two cases:

$$\begin{aligned} k^0 &= +\sqrt{\mathbf{k}^2+m^2}+i\varepsilon \\ k^0 &= +\sqrt{\mathbf{k}^2+m^2}-i\varepsilon \end{aligned}$$

The first choice gives the Green function

$$\begin{aligned} G_{+E,+t}(x,x') &\equiv \frac{1}{(2\pi)^4}\int \frac{d^4k}{2k^0}\frac{e^{i\left(k_0\left(t-t'\right)-\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)\right)}}{k^0-\sqrt{\mathbf{k}^2+m^2}-i\varepsilon} \\ &= \frac{1}{(2\pi)^4}\int d^3k\, e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)}\int \frac{dk^0}{2k^0}\frac{e^{ik^0\left(t-t'\right)}}{k^0-\sqrt{\mathbf{k}^2+m^2}-i\varepsilon} \end{aligned}$$

To close the k^0 contour, we add a half-circle and let its radius tend to infinity. When $t>t'$, we must add this half circle in the upper half plane, $k^0=k_x^0+ik_y^0$, so that on the circle

$$e^{ik^0\left(t-t'\right)}=e^{i\left(k_x^0+ik_y^0\right)\left(t-t'\right)}=e^{ik_x^0\left(t-t'\right)}e^{-k_y^0\left(t-t'\right)}$$

converges to zero as $k_y^0\to\infty$. For $t<t'$, we must close the contour in the lower half plane. Since the pole is in the upper half plane, the integral for $t<t'$ gives zero, while for $t>t'$ the residue at the pole is

$$\lim_{\varepsilon\to 0} Res\left(\frac{1}{2k^0}\frac{e^{ik^0\left(t-t'\right)}}{k^0-\sqrt{\mathbf{k}^2+m^2}-i\varepsilon}\right)=\frac{dk^0}{2\sqrt{\mathbf{k}^2+m^2}}e^{i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)}$$

so including unit a step function,

$$\Theta(t-t') \equiv \begin{cases} 1 & t>t' \\ 0 & t<t' \end{cases}$$

the Green function becomes

$$G_{+E,+t}(x,x') = \frac{i}{(2\pi)^3}\Theta(t-t')\int\frac{d^3k}{2\sqrt{\mathbf{k}^2+m^2}}e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)+i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)}$$

We will not need the explicit form of the remaining integral, so define

$$H_I(x,x') \equiv \frac{i}{(2\pi)^3}\int\frac{d^3k}{2\sqrt{\mathbf{k}^2+m^2}}e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)+i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)}$$

and write the Green function as $G_{+E,+t}(x,x')=H_I(x,x')\,\Theta(t-t')$.

For the second displacement, the upper contour (for $t>t'$) gives zero contribution while for $t<t'$ we compute

$$\begin{aligned} G_{+E,-t}(x,x') &= \frac{1}{(2\pi)^4}\Theta(t'-t)\int\frac{d^4k}{2k^0}\frac{1}{k^0-\sqrt{\mathbf{k}^2+m^2}}e^{i\left(k_0\left(t-t'\right)-\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)\right)} \\ &= \Theta(t'-t)\lim_{\varepsilon\to 0}\frac{1}{(2\pi)^4}\int d^3k e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)}\oint_{upper\ \frac{1}{2}-plane}\frac{dk^0}{2k^0}\frac{e^{-ik^0\left(t'-t\right)}}{k^0-\sqrt{\mathbf{k}^2+m^2}+i\varepsilon} \\ &= \Theta(t'-t)\frac{2\pi i}{(2\pi)^4}\int\frac{d^3k}{2\sqrt{\mathbf{k}^2+m^2}}e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)}e^{i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)} \\ &= H_I(x,x') \end{aligned}$$

5.1.2 Second pole

For the second pole, at $k^0 = -\sqrt{\mathbf{k}^2 + m^2}$, we again have two possible displacements,

$$
\begin{aligned}
k^0 &= -\sqrt{\mathbf{k}^2 + m^2} + i\varepsilon \\
k^0 &= -\sqrt{\mathbf{k}^2 + m^2} - i\varepsilon
\end{aligned}
$$

Choosing the first,

$$G_{-E,+t}(x, x') = \lim_{\varepsilon \to 0} \frac{1}{(2\pi)^4} \int \frac{d^4k}{2k^0} \frac{e^{i\left(k^0(t-t') - \mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')\right)}}{k^0 + \sqrt{\mathbf{k}^2 + m^2} - i\varepsilon}$$

has the pole in the upper half plane. When $t > t'$, we must close the contour in the upper half plane and we get a residue,

$$
\begin{aligned}
G_{-E,+t}(x, x') &= \Theta(t - t') \frac{-2\pi i}{(2\pi)^4} \int \frac{d^3k}{2\sqrt{\mathbf{k}^2 + m^2}} e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} e^{-i\sqrt{\mathbf{k}^2+m^2}(t-t')} \\
&\equiv \Theta(t - t') H_{II}(x, x')
\end{aligned}
$$

Finally, pushing the pole to the lower half plane gives

$$G_{-E,-t}(x, x') = \Theta(t' - t) H_{II}(x, x')$$

Collecting all four possible Green functions,

$$
\begin{aligned}
G_{+E,+t}(x, x') &= \Theta(t - t') H_I(x, x') && (5.7) \\
G_{+E,-t}(x, x') &= \Theta(t' - t) H_I(x, x') && (5.8) \\
G_{-E,+t}(x, x') &= \Theta(t - t') H_{II}(x, x') && (5.9) \\
G_{-E,-t}(x, x') &= \Theta(t' - t) H_{II}(x, x') && (5.10)
\end{aligned}
$$

The $+E$ and $-E$ subscripts indicate whether the solution describes a positive or negative energy solution. Restoring $\hbar$ and c,

$$
\begin{aligned}
E &= \hbar k^0 = +\sqrt{\hbar^2 \mathbf{k}^2 + m^2 c^4} \\
E &= \hbar k^0 = -\sqrt{\hbar^2 \mathbf{k}^2 + m^2 c^4}
\end{aligned}
$$

We now show that the $+t$ and $-t$ subscripts indicate whether solutions progress causally toward the future or toward the past.

From Eqs.(5.5) and (5.6), we see that the full Green function is a sum of one of the H_I terms with one of the H_{II} terms, that is, one positive energy Green function and one negative energy Green function. Classically, we would choose the Green function to be

$$G(x, x') = G_{+E,+t}(x, x') + G_{-E,+t}(x, x') \tag{5.11}$$

because then the solution is for ϕ is given by

$$
\begin{aligned}
\phi(t, \mathbf{x}) &= \int d^4x' G(x, x') J(\mathbf{x}') \\
&= \int_{-\infty}^{\infty} dt' \Theta(t - t') \int d^3x' \left(H_I(x, x') + H_{II}(x, x')\right) J(t', \mathbf{x}') \\
&= \int_{-\infty}^{t} dt' \int d^3x' \left(H_I(x, x') + H_{II}(x, x')\right) J(t', \mathbf{x}')
\end{aligned}
$$

The limits on the final time integral show that the field at time t is determined only by sources $J(t', \mathbf{x}')$ evaluated for times t' earlier than t. This is our usual minimal expectation for causality. However, Feynman has shown that using any of the Green functions is consistent with causality, and proposes pairing

$G_{+e+t}(x,x')$ with $G_{-E-t}(x,x')$. Indeed, the Feynman choice is actually *more* consistent with causality as we now understand it. Causality, in essence, is the preservation of the spacetime light cone. No physical propagation that begins in the future-pointing light cone may exceed the speed of light – it must remain in the future-pointing light cone. We call such motion *futurelike.* Correspondingly, we are justified in asserting that any propagation beginning in a direction inside the past-pointing light cone must remain within this past-pointing light cone. Motion into the past light cone is called *pastlike.* A similar prohibition applies for causal tachyons – particles whose motion remains in spacelike directions. The symmetry of the situation suggests that it is reasonable to consider both directions of time propagation equally. Doing so leads us to a clearer understanding of antiparticles.

Choosing the Green function in the form which associates positive energy solutions with futurelike motion in time and negative energy solutions to pastlike motion,

$$G(x,x') = G_{+E,+t}(x,x') + G_{-E,-t}(x,x')$$

leads to fields of the *past-future symmetric Green function*

$$\phi(t,\mathbf{x}) = \int_{-\infty}^{t} dt' \int d^3x' H_I(x,x')\, J(t',\mathbf{x}') + \int_{t}^{\infty} dt' \int d^3x' H_{II}(x,x')\, J(t',\mathbf{x}') \tag{5.12}$$

As a result of this choice, $\phi(t,\mathbf{x})$ can depend on events in both its forward and backward light cones. The benefit of this choice is that it gives a clear physical meaning to the negative energy solutions, for the following reason. Suppose a particle travels backward in time, from point $A(t_2,x_2)$ to point $B(t_1,x_1)$ with $t_1 < t_2$. Then an observer moving forward in time will experience the particle first at t_1 and later at t_2 and the particle will appear to move in the opposite direction, from x_1 to x_2. Moreover, if the particle carries negative energy from A to B, the observer sees the negative energy arrive at B, then depart later from A. This means that the energy at B decreases and the energy at A increases, so to the futurelike observer a positive amount of energy has moved from B to A, *forward* in time. The same argument applies to electric or other charges. If a negative charge moves backward in time from A to B, the forward moving observer sees a positive charge leave B then arrive at A.

The pastlike propagation is also consistent with the energy-momentum 4-vectors,

$$\begin{aligned} p_+^\alpha &= \left(+\sqrt{\hbar^2\mathbf{k}^2 + m^2c^4}, \mathbf{k}\right) \\ p_-^\alpha &= \left(-\sqrt{\hbar^2\mathbf{k}^2 + m^2c^4}, \mathbf{k}\right) \end{aligned}$$

These show the direction of the motion of the particle in spacetime, pointing into the future and past light cones, respectively.

To summarize: fix a set of Cartesian coordinates on spacetime, $(t,\mathbf{x})$ where the sign of t distinguishes the two halves of the light cone, "future" and "past". Now consider a futurelike observer, that is, moving in such a way that the time coordinate t associated with their position increases. To this observer, particles moving into the future light cone will have positive energy E, momentum $\mathbf{p}$, and may have a charge q. When this same futurelike observer observes a negative energy state of the *same* type of particle travelling into the past light cone, (with decreasing time t, energy $-E$, momentum $\mathbf{p}$, and charge q), the particle appears to the futurelike observer to move in the direction of *increasing* t, have energy $+E$, momentum $-\mathbf{p}$, and charge $-q$.

5.1.3 Charged scalars

We can also see how the creation and anihilation operators depend on currents,

$$\begin{aligned} a(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega}{2}} \int d^3x \left(\varphi(\mathbf{x},0) - \frac{i}{\omega}\pi^*(\mathbf{x},0)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \\ b(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega}{2}} \int d^3x \left(\varphi^*(\mathbf{x},0) - \frac{i}{\omega}\pi(\mathbf{x},0)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \end{aligned}$$

and the past-future symmetric Green function,

$$\begin{aligned}
\phi(t,\mathbf{x}) &= \int_{-\infty}^{t} dt' \int d^3x' H_I\left(x,x'\right) J(t',\mathbf{x}') + \int_t^{\infty} \int d^3x' dt' H_{II}\left(x,x'\right) J\left(t',\mathbf{x}'\right) \\
&= \frac{i}{2\left(2\pi\right)^3} \int_{-\infty}^{t} dt' \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)+i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)} J\left(t',\mathbf{x}'\right) \\
&\quad -\frac{i}{2\left(2\pi\right)^3} \int_t^{\infty} dt' \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} e^{-i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)} J\left(t',\mathbf{x}'\right)
\end{aligned}$$

so the momentum is

$$\begin{aligned}
\pi^*\left(\mathbf{x},t\right) &= \partial_0 \varphi\left(\mathbf{x},t\right) \\
&= \frac{i}{2\left(2\pi\right)^3} \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} J(t',\mathbf{x}') - \frac{1}{2\left(2\pi\right)^3} \int_{-\infty}^{t} dt' \int d^3x' \int d^3k e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)+i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)} J\left(t',\right. \\
&\quad +\frac{i}{2\left(2\pi\right)^3} \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} J(t',\mathbf{x}') - \frac{1}{2\left(2\pi\right)^3} \int_t^{\infty} dt' \int d^3x' \int d^3k e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} e^{-i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)} J\left(\right. \\
&= \frac{i}{\left(2\pi\right)^3} \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} J(t',\mathbf{x}') - \frac{1}{2\left(2\pi\right)^3} \int_{-\infty}^{t} dt' \int d^3x' \int d^3k e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)+i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)} J\left(t',\mathbf{x}\right. \\
&\quad -\frac{1}{2\left(2\pi\right)^3} \int_t^{\infty} dt' \int d^3x' \int d^3k e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} e^{-i\sqrt{\mathbf{k}^2+m^2}\left(t-t'\right)} J\left(t',\mathbf{x}'\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}
a\left(\mathbf{k}\right) &= \frac{1}{\left(2\pi\right)^{3/2}} \sqrt{\frac{\omega}{2}} \int d^3x \left(\varphi\left(\mathbf{x},0\right) - \frac{i}{\omega}\pi^*\left(\mathbf{x},0\right)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \\
&= \frac{1}{\left(2\pi\right)^{3/2}} \sqrt{\frac{\omega}{2}} \int d^3x \left(\frac{i}{2\left(2\pi\right)^3} \int_{-\infty}^{0} dt' \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)-i\sqrt{\mathbf{k}^2+m^2}t'} J\left(t',\mathbf{x}'\right)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \\
&\quad +\frac{1}{\left(2\pi\right)^{3/2}} \sqrt{\frac{\omega}{2}} \int d^3x \left(-\frac{i}{2\left(2\pi\right)^3} \int_0^{\infty} dt' \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} e^{i\sqrt{\mathbf{k}^2+m^2}t'} J\left(t',\mathbf{x}'\right)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \\
&\quad -\frac{i}{\left(2\pi\right)^{3/2}} \frac{1}{\sqrt{2\omega}} \int d^3x \left(\frac{i}{\left(2\pi\right)^3} \int d^3x' \int \frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}} e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} J\left(t',\mathbf{x}'\right)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \\
&\quad -\frac{i}{\left(2\pi\right)^{3/2}} \frac{1}{\sqrt{2\omega}} \int d^3x \left(-\frac{1}{2\left(2\pi\right)^3} \int_{-\infty}^{0} dt' \int d^3x' \int d^3k e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)-i\sqrt{\mathbf{k}^2+m^2}t'} J\left(t',\mathbf{x}'\right)\right) e^{i\mathbf{k}\cdot\mathbf{x}} \\
&\quad -\frac{i}{\left(2\pi\right)^{3/2}} \frac{1}{\sqrt{2\omega}} \int d^3x \left(-\frac{1}{2\left(2\pi\right)^3} \int_0^{\infty} dt' \int d^3x' \int d^3k e^{-i\mathbf{k}\cdot\left(\mathbf{x}-\mathbf{x}'\right)} e^{i\sqrt{\mathbf{k}^2+m^2}t'} J\left(t',\mathbf{x}'\right)\right) e^{i\mathbf{k}\cdot\mathbf{x}}
\end{aligned}$$

so

$$\begin{aligned}
a(\mathbf{k}) &= \frac{i}{2(2\pi)^{9/2}}\sqrt{\frac{\omega}{2}}\int d^3x\int_{-\infty}^{0}dt'\int d^3x'\int\frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}}e^{i\mathbf{k}\cdot\mathbf{x}'}e^{-i\sqrt{\mathbf{k}^2+m^2}t'}J(t',\mathbf{x}') \\
&\quad -\frac{i}{2(2\pi)^{9/2}}\sqrt{\frac{\omega}{2}}\int d^3x\int_{0}^{\infty}dt'\int d^3x'\int\frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}}e^{i\mathbf{k}\cdot\mathbf{x}'}e^{i\sqrt{\mathbf{k}^2+m^2}t'}J(t',\mathbf{x}') \\
&\quad +\frac{1}{(2\pi)^3}\frac{1}{(2\pi)^{3/2}}\frac{1}{\sqrt{2\omega}}\int d^3x\int d^3x'\int\frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}}e^{i\mathbf{k}\cdot\mathbf{x}'}J(t',\mathbf{x}') \\
&\quad +\frac{i}{2(2\pi)^{9/2}}\frac{1}{\sqrt{2\omega}}\int d^3x\int_{-\infty}^{0}dt'\int d^3x'\int d^3k e^{i\mathbf{k}\cdot\mathbf{x}'}e^{-i\sqrt{\mathbf{k}^2+m^2}t'}J(t',\mathbf{x}') \\
&\quad +\frac{i}{2(2\pi)^{9/2}}\frac{1}{\sqrt{2\omega}}\int d^3x\int_{0}^{\infty}dt'\int d^3x'\int d^3k e^{i\mathbf{k}\cdot\mathbf{x}'}e^{i\sqrt{\mathbf{k}^2+m^2}t'}J(t',\mathbf{x}')
\end{aligned}$$

Collecting terms

$$\begin{aligned}
a(\mathbf{k}) &= \frac{1}{(2\pi)^{9/2}}\frac{1}{\sqrt{2\omega}}\int d^3x\int d^3x'\int\frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}}e^{i\mathbf{k}\cdot\mathbf{x}'}J(t',\mathbf{x}') \\
&\quad +\frac{i}{2(2\pi)^{9/2}}\sqrt{\frac{\omega}{2}}\int d^3x\int d^3x'\int\frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}}e^{i\mathbf{k}\cdot\mathbf{x}'}\left(\int_{-\infty}^{0}dt'e^{-i\sqrt{\mathbf{k}^2+m^2}t'}J(t,\mathbf{x}')-\int_{0}^{\infty}dt'e^{i\sqrt{\mathbf{k}^2+m^2}t'}J(t',\mathbf{x}')\right) \\
&\quad +\frac{i}{2(2\pi)^{9/2}}\frac{1}{\sqrt{2\omega}}\int d^3x\int d^3x'\int d^3k e^{i\mathbf{k}\cdot\mathbf{x}'}\left(\int_{-\infty}^{0}dt'e^{-i\sqrt{\mathbf{k}^2+m^2}t'}J(t,\mathbf{x}')+\int_{0}^{\infty}dt'e^{i\sqrt{\mathbf{k}^2+m^2}t'}J(t',\mathbf{x}')\right)
\end{aligned}$$

Let the Fourier transformation of J be written as

$$\tilde{J}(t',\mathbf{k})=\frac{1}{(2\pi)^{3/2}}\int d^3x'e^{i\mathbf{k}\cdot\mathbf{x}'}J(tt,\mathbf{x}')$$

and look at the anihilation per unit volume,

$$\begin{aligned}
\frac{a(\mathbf{k})}{\int d^3x} &= \frac{1}{(2\pi)^3}\frac{1}{\sqrt{2\omega}}\int\frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}}\tilde{J}(t',\mathbf{k}) \\
&\quad +\frac{i}{2(2\pi)^3}\sqrt{\frac{\omega}{2}}\int\frac{d^3k}{\sqrt{\mathbf{k}^2+m^2}}\left(\int_{-\infty}^{0}dt'e^{-i\sqrt{\mathbf{k}^2+m^2}t'}\tilde{J}(t',\mathbf{k})-\int_{0}^{\infty}dt'e^{i\sqrt{\mathbf{k}^2+m^2}t'}\tilde{J}(t',\mathbf{k})\right) \\
&\quad +\frac{i}{2(2\pi)^3}\frac{1}{\sqrt{2\omega}}\int d^3k\left(\int_{-\infty}^{0}dt'e^{-i\sqrt{\mathbf{k}^2+m^2}t'}\tilde{J}(t',\mathbf{k})+\int_{0}^{\infty}dt'e^{i\sqrt{\mathbf{k}^2+m^2}t'}\tilde{J}(t',\mathbf{k})\right)
\end{aligned}$$

Like the fields, the creation and anihilation operators may depend on either past or future sources.

5.1.4 Chronicity (skip)

Let's make this precise using the discrete Lorentz transformations.

5.1.4.1 Past-future interchange

Choose a foliation of spacetime by timelike curves. (While choosing any inertial frame of reference will to in flat spacetim, this is possible even in curved spacetimes as long as they are globally hyperbolic, that is, they admit a Cauchy surface. This makes the manifold diffeomorphic to a product of the Cauchy surface with $\mathbb{R}$. See Bernal and Sánchez [1]). At each point $x=(t,\mathbf{x})$ along each such curve, let the 4-velocity (the unit tangent to the curve) be given by

$$u^{\alpha}(x)=\frac{dx^{\alpha}}{d\tau}$$

From these we construct the projection operator,

$$P^{\alpha}{}_{\beta} \equiv \frac{1}{c^2} u^{\alpha} u_{\beta}$$

This is a projection since it is idempotent,

$$\begin{aligned} P^{\alpha}{}_{\mu} P^{\mu}{}_{\beta} &\equiv \frac{1}{c^2} u^{\alpha} \left(\frac{1}{c^2} u_{\mu} u^{\mu} \right) u_{\beta} \\ &= \frac{1}{c^2} u^{\alpha} u_{\beta} \end{aligned}$$

This projects any 4-vector, $w^{\alpha} = \left(w^0, \mathbf{w}\right)$ into the timelike direction of this reference frame. View the action of $P^{\alpha}{}_{\beta}$ in an arbitrary inertial frame, where $u^{\alpha} = \gamma\left(c, \mathbf{v}\right)$. The action on w^{α} is

$$\begin{aligned} P^{\alpha}{}_{\beta} w^{\beta} &= \frac{1}{c^2} u^{\alpha} \left(u_{\beta} w^{\beta} \right) \\ &= \frac{1}{c^2} \gamma \left(c w^0 - \mathbf{v} \cdot \mathbf{w} \right) u^{\alpha} \end{aligned}$$

Now consider the effect of the operator

$$T^{\alpha}{}_{\beta} \equiv \delta^{\alpha}{}_{\beta} - 2P^{\alpha}{}_{\beta}$$

In the rest frame of u^{α}, we find

$$\begin{aligned} T^{\alpha}{}_{\beta} w^{\beta} &= w^{\alpha} - \frac{2}{c^2} u^{\alpha} \left(u_{\beta} w^{\beta} \right) \\ &= w^{\alpha} - \frac{2}{c^2} \left(c w^0 \right) u^{\alpha} \\ &= w^{\alpha} - 2 w^0 \left(1, \mathbf{0} \right) \\ &= \left(-w^0, \mathbf{w} \right) \end{aligned}$$

and this result is Lorentz covariant, so though the components change in a different frame, they describe the same 4-vector. Thus, the effect of this transformation is to interchange the future and past light cones. However, the path of particles changes. Suppose a particle is traveling, in the frame of u^{α} at constant velocity to the right along the curve $x^{\alpha}(t) = (ct, vt, 0, 0)$. Then the 4-velocity is

$$\begin{aligned} v^{\alpha} &= \frac{dx^{\alpha}}{d\tau} \\ &= \gamma_v \left(c, v, 0.0 \right) \end{aligned}$$

Applying $T^{\alpha}{}_{\beta}$ this becomes

$$T^{\alpha}{}_{\beta} v^{\beta} = \gamma_v \left(-c, v, 0.0 \right)$$

and a spacetime curve with this tangent vector passes through the origin and moves down and to the *right*. The future and past light cones have been interchanged, and curves are mirror reflected across a spatial surface. The transformation $T^{\alpha}{}_{\beta}$ accomplishes *past-future* interchange. This is a symmetry of the second order scalar field equation, but it will not be a symmetry of the Dirac equation. Instead,

$$i\gamma^0 \partial_0 \psi + i\boldsymbol{\gamma} \cdot \boldsymbol{\nabla} \psi - m\psi = 0$$

becomes

$$-i\gamma^0 \partial_0 \psi + i\boldsymbol{\gamma} \cdot \boldsymbol{\nabla} \psi - m\psi = 0$$

If we look at the condition the Dirac equation places on a plane wave, $\psi = A \exp i \left(Et - \mathbf{p} \cdot \mathbf{x} \right)$,

$$-\gamma^0 E \psi + \boldsymbol{\gamma} \cdot \mathbf{p} \psi - m\psi = 0$$

while the transformed equation becomes

$$\gamma^0 E\psi + \boldsymbol{\gamma}\cdot\mathbf{p}\psi - m\psi = 0$$

This is just the negative energy solution. When we study solutions of the Dirac equation in detail, we will see that it has both positive and negative energy solutions, so the transformed equation remains a solution.

In general, it is clear that $T^{\alpha}{}_{\beta}$ is a discrete Lorentz transformation because the length of any 4-vector,

$$\tau = \sqrt{c^2t^2 - \mathbf{x}^2}$$

is unchanged if t is replaced by $-t$.

5.1.4.2 Reversal of motion

We contrast T with *reversal of motion*, R. Along every timelike curve in spacetime, replace the proper time parameter τ with $-\tau$. Then the positions are unchanged, but tangent vectors to curves point in the opposite direction. This is reverses the direction of motion:

$$\begin{aligned} R\tau &= -\tau \\ Rx^{\alpha} &= x^{\alpha} \end{aligned}$$

so that

$$\begin{aligned} u^{\alpha} &= \frac{dx^{\alpha}}{d\tau} \\ R(u^{\alpha}) &= \frac{dx^{\alpha}}{d(-\tau)} \\ &= -u^{\alpha} \end{aligned}$$

Each particle follows the same timelike curve before and after the transformation, but the direction of motion is now from future to past. Reversal of motion is also a discrete Lorentz transformation, since $\tau^2 = \eta_{\mu\nu}x^{\mu}x^{\nu}$ is preserved.

In our discussion of discrete Lorentz transformations, we defined *chronicity*, Θ, as follows:

$$\begin{aligned} \Theta &: t \to -t \\ \Theta &: \mathbf{x} \to \mathbf{x} \\ \Theta &: E \to -E \\ \Theta &: \mathbf{p} \to \mathbf{p} \\ \Theta &: q \to q \end{aligned}$$

We also need the actions of *charge conjugation*,

$$\begin{aligned} \mathcal{C} &: t \to t \\ \mathcal{C} &: \mathbf{x} \to \mathbf{x} \\ \mathcal{C} &: E \to E \\ \mathcal{C} &: \mathbf{p} \to \mathbf{p} \\ \mathcal{C} &: q \to -q \end{aligned}$$

which we implement by complex conjugation, and *parity*

$$\begin{aligned} \mathcal{P} &: t \to t \\ \mathcal{P} &: \mathbf{x} \to -\mathbf{x} \\ \mathcal{P} &: E \to E \\ \mathcal{P} &: \mathbf{p} \to -\mathbf{p} \\ \mathcal{P} &: q \to q \end{aligned}$$

The effect of combining all three operations at once is then

$$\begin{aligned}
\mathcal{CP}\Theta &: t \to -t \\
\mathcal{CP}\Theta &: \mathbf{x} \to -\mathbf{x} \\
\mathcal{CP}\Theta &: E \to -E \\
\mathcal{CP}\Theta &: \mathbf{p} \to -\mathbf{p} \\
\mathcal{CP}\Theta &: q \to -q
\end{aligned}$$

The action on the phase of a field is

$$\mathcal{CP}\Theta : \frac{i}{\hbar}(Et - \mathbf{p}\cdot\mathbf{x}) \to -\frac{i}{\hbar}((-E)(-t) - (-\mathbf{p})\cdot(-\mathbf{x})) = -\frac{i}{\hbar}(Et - \mathbf{p}\cdot\mathbf{x}) \tag{5.13}$$

Therefore, if we always choose our field expansions to include $\varphi^\dagger$ symmetrically with φ, the field theory will be $\mathcal{CP}\Theta$-invariant. This combined action of discrete transformations gives us the picture we want. By simply changing the sign , we turn the phase a particle of 4-momentum $p^\mu = (E, \mathbf{p})$ and charge q into the phase of a particle 4-momentum $p^\mu = (-E, -\mathbf{p})$ and charge $-q$ travelling backward in time in a parity flipped space.

Now our interpretation of the negative energy states is clear. By choosing the Green function to be

$$G(x, x') = G_{+E,+t}(x, x') + G_{-E,-t}(x, x') \tag{5.14}$$

and the field expansion to be

$$\varphi(\mathbf{x}, t) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega}} \left(a(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} + a^\dagger(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{5.15}$$

we always associate the negative energy solutions with pastlike motion. The appearance of such a pastlike particle (with energy $-E$, momentum $\mathbf{p}$ and charge q) to a futurelike observer is the $\mathcal{CP}\Theta$ transform of the field, i.e., a futurelike particle of energy $+E$, momentum $-\mathbf{p}$ and charge $-q$.

Define: Suppose a given variety of particle exists in futurelike states described by physical field $\phi_+ (E, \mathbf{p}, q \ldots)$, and also in pastlike states described by $\phi_- (-E, \mathbf{p}, q \ldots)$ having negative energy. Then $\phi_+ (E, \mathbf{p}, q \ldots)$ has an *antiparticle* state defined as $\mathcal{CP}\Theta\phi_- (-E, \mathbf{p}, q \ldots)$.

Since

$$\mathcal{CP}\Theta\phi_- (-E, \mathbf{p}, q \ldots) = \mathcal{CP}\Theta\phi_- (E, -\mathbf{p}, -q \ldots)$$

antiparticle states are positive energy and futurelike. It is easy to see that all other quantum numbers are reversed, because a pastlike particle carrying any quantum charge g into the past will be experienced by a futurelike observer as carrying a charge $-g$ into the future.

We require field theories to be symmetric with respect to particles and antiparticles, so that for field operators

$$\mathcal{CP}\Theta\hat{\varphi}(t, \mathbf{x}) (\mathcal{CP}\Theta)^{-1} = \hat{\varphi}(t, \mathbf{x})$$

Since the conjugate momentum

$$\hat{\pi}(\mathbf{x}, t) = \frac{\partial}{\partial t}\hat{\varphi}(\mathbf{x}, t)$$

satisfies

$$\begin{aligned}
\mathcal{CP}\Theta\hat{\pi}(\mathbf{x}, t) (\mathcal{CP}\Theta)^{-1} &= \mathcal{CP}\Theta \left(\frac{\partial}{\partial t}\hat{\varphi}(t, \mathbf{x}) \right) (\mathcal{CP}\Theta)^{-1} \\
&= -\frac{\partial}{\partial t}\hat{\varphi}(t, \mathbf{x}) \\
&= -\hat{\pi}(\mathbf{x}, t)
\end{aligned}$$

we find that the effect of $\mathcal{CP}\Theta$ on $a(\mathbf{k}) = \hat{\varphi}(\mathbf{x}) - \frac{i}{\omega}\hat{\pi}(\mathbf{x})$ is

$$\begin{aligned}
\mathcal{CP}\Theta a(\mathbf{k})(\mathcal{CP}\Theta)^{-1} &= \mathcal{CP}\Theta\left(\hat{\varphi}(\mathbf{x}) - \frac{i}{\omega}\hat{\pi}(\mathbf{x})\right)(\mathcal{CP}\Theta)^{-1} \\
&= \hat{\varphi}(\mathbf{x}) - \frac{(-i)}{(-\omega)}(-\hat{\pi}(\mathbf{x})) \\
&= \hat{\varphi}(\mathbf{x}) + \frac{i}{\omega}\hat{\pi}(\mathbf{x}) \\
&= a^{\dagger}(\mathbf{k})
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
\mathcal{CP}\Theta i(\omega t - \mathbf{k}\cdot\mathbf{x})(\mathcal{CP}\Theta)^{-1} &= -i((-\omega)(-t) - (-\mathbf{k})\cdot(-\mathbf{x})) \\
&= -i(\omega t - \mathbf{k}\cdot\mathbf{x})
\end{aligned}$$

so that the action of $\mathcal{CP}\Theta$ on a plane wave is

$$\mathcal{CP}\Theta\left(a(E,\mathbf{p})e^{\frac{i}{\hbar}(p_\alpha x^\alpha)}\right)(\mathcal{CP}\Theta)^{-1} = a^{\dagger}(E,\mathbf{p})e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)}$$

Therefore, the full expansion for $\hat{\varphi}(t,\mathbf{x})$ will be symmetric under $\mathcal{CP}\Theta$ if it is an equal linear combination of terms

$$\hat{\varphi}(t,\mathbf{x}) \sim a(E,\mathbf{p})e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^{\dagger}(E,\mathbf{p})e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)}$$

This form agrees with eq.(5.15) for $\hat{\varphi}(t,\mathbf{x})$. But eq.(5.15) was found by setting

$$\varphi(\mathbf{x},t) = \frac{1}{(2\pi)^{3/2}}\int\sqrt{2E}\left(a(E,\mathbf{p})e^{\frac{i}{\hbar}(p_\alpha x^\alpha)} + a^{\dagger}(E,\mathbf{p})e^{-\frac{i}{\hbar}(p_\alpha x^\alpha)}\right)\delta\left(p_\alpha p^\alpha - m^2\right)\Theta(E)\hbar^{-4}d^4p$$

where we included the positive energy step function. The present calculation justifies our earlier step.

Our choice of boundary conditions leads us to ask what boundary conditions the other possible Green functions represent. A moment's reflection on the expression

$$G(x,x') = G_{+E,-t}(x,x') + G_{-E,+t}(x,x')$$

suggest that this is the proper Green function for an *observer* travelling backward in time. For such an observer, an antiparticle (also moving backward in time) would be assigned positive energy, hence the $G_{+E,-t}(x,x')$ term. To the same observer a matter particle would be a negative energy state travelling in the positive time direction.

5.1.5 Chronicity, time reversal and the Schrödinger equation

The relationship of chronicity and time reversal to quantum mechanics is also interesting. Consistent with energy in Newtonian mechanics, the action of time reversal is always taken to leave the Hamiltonian invariant. By contrast, the chronicity reverses the sign of the energy. We now consider the effect of these transformations on solutions of the Schrödinger equation.

Suppose a state ψ solves the Schrödinger equation,

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi$$

We want to know when a transformed state $T\psi$ is also a solution, where T is either time reversal or chronicity. In either case we have

$$i\hbar\frac{\partial(T\psi)}{\partial t} = \hat{H}(T\psi)$$

A sufficient condition for this to be the case is found by acting on the equation with T^{-1} and inserting appropriate identities:

$$\begin{aligned} T^{-1} i\hbar \frac{\partial (T\psi)}{\partial t} &= T^{-1}\hat{H}(T\psi) \\ \left(T^{-1} i\hbar T\right)\left(T^{-1}\frac{\partial}{\partial t}T\right)\psi &= \left(T^{-1}\hat{H}T\right)\psi \end{aligned}$$

Now, for both time reversal and chronicity, $T^{-1}\frac{\partial}{\partial t}T = -\frac{\partial}{\partial t}$. Therefore the transformed state $T\psi$ is a solution if

$$-\left(T^{-1} i T\right)\hbar\frac{\partial}{\partial t} = T^{-1}\hat{H}T$$

Time reversal and chronicity take advantage of the two simple ways to solve this equation. For time reversal, is accomplished by making the operator *anti*-unitary,

$$\mathcal{T} i \mathcal{T}^{-1} = -i$$

while chronicity is unitary but changes the sign of the Hamiltonian,

$$\Theta\hat{H}\Theta^{-1} = -\hat{H}$$

This is the reason that chronicity is not suitable for quantum *mechanics*: since quantum mechanics includes neither antiparticles nor pastlike particles, negative energy states cannot be reinterpreted as futurelike, positive energy states. Then, the presence of both positive and negative energy states of the same quantum system leads to runaway production of ever more negative energy states. As noted previously, the failure of energy and momentum to form a 4-vector under time reversal is not a problem in a non-relativistic theory.

With both pastlike and futurelike particles present symmetrically, we may consistently regard all negative energy states with futurelike positive energy states, so there will be no runaway solutions. Another way to think about this is to consider interactions. The interaction of a futurelike particle with a pastlike particle always occurs as if the futurelike particle were encountering a positive energy antiparticle. Nor can futurelike particles gain arbitrary energy by creating negative energy states, because the only negative energy states are pastlike. Futurelike particles can only produce pastlike particles under special conditions such as particle-antiparticle annihilation.

One further consequence of using chronicity is that, being hermitian, it is a quantum observable. Since $\mathcal{T}^2 = 1$, there will be two eigenvalues. We conjecture that these will correspond to *antiparticle number*, with particles assigned the eigenvalue $+1$ and their antiparticles the eigenvalue -1. Of course, it is arbitrary which is called the particle, but the two states are distinguishable. This assignment is equivalent to assigning plus one to futurelike observers and minus one to pastlike observers, which accounts for our observations revealing only the $+1$ eigenvalue and only the one pair of Green functions.

Chapter 6

Quantization of the Dirac field

6.1 Solution of the free classical Dirac equation

As with the scalar field, we can solve using a Fourier integral. First consider a single value of the momentum. Then we can write two plane wave solutions with fixed energy 4-momentum p^α in the form

$$\psi(\mathbf{x},t) = u(p^\alpha)e^{-ip_\alpha x^\alpha} + v(p^\alpha)e^{ip_\alpha x^\alpha} \tag{6.1}$$

where $u(p^\alpha)$ and $v(p^\alpha)$ are spinors, $p^\alpha = \left(E, p^i\right)$ and $p_\alpha = (E, p_i) = (E, -p^i)$. Substituting,

$$\begin{aligned}
0 &= (i\gamma^\alpha\partial_\alpha - m)\,\psi(\mathbf{x},t) \\
&= (i\gamma^\alpha\partial_\alpha - m)\left(u(p^\alpha)e^{-ip_\alpha x^\alpha} + v(p^\alpha)e^{ip_\alpha x^\alpha}\right) \\
&= (\gamma^\alpha p_\alpha - m)\,u(p^\alpha)e^{-ip_\alpha x^\alpha} - (\gamma^\alpha p_\alpha + m)\,v(p^\alpha)e^{ip_\alpha x^\alpha}
\end{aligned}$$

we find the pair of equations

$$(\gamma^\alpha p_\alpha - m)\,u(p^\alpha) = 0 \tag{6.2}$$

$$(\gamma^\alpha p_\alpha + m)\,v(p^\alpha) = 0 \tag{6.3}$$

for the $u(p^\alpha)$ and $v(p^\alpha)$ modes, respectively.

To begin, write out the equation using the Dirac matrices as given in eqs.(2.35),

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & \\ & -\mathbf{1} \end{pmatrix}, \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$$

and solve first for $u(p^\alpha)$. If we set

$$[u(p_\alpha)]^A = \begin{pmatrix} \alpha(p_\alpha) \\ \beta(p_\alpha) \end{pmatrix}$$

where $A = 1,2,3,4$, then we get the matrix equation

$$\begin{aligned}
0 &= (\gamma^\alpha p_\alpha - m)\,w(p^\alpha) \\
&= \begin{pmatrix} E - m & \sigma^i p_i \\ -\sigma^i p_i & -E - m \end{pmatrix}\begin{pmatrix} \alpha(p_\alpha) \\ \beta(p_\alpha) \end{pmatrix}
\end{aligned}$$

which gives the set of 2×2 equations

$$(E - m)\,\alpha(p_\alpha) + \sigma^i p_i \beta(p_\alpha) = 0 \tag{6.4}$$

$$-\sigma^i p_i \alpha(p_\alpha) - (E + m)\,\beta(p_\alpha) = 0 \tag{6.5}$$

Since $E > 0$, the quantity $E + m$ is nonzero so Eq.(6.5) may be solved for $\beta(p_\alpha)$,

$$\beta(p_\alpha) = -\left(\frac{\sigma^i p_i}{E+m}\right)\alpha(p_\alpha) \tag{6.6}$$

Substituting into the first:

$$\begin{aligned}(E-m)\,\alpha(p_\alpha) &= \sigma^i p_i\left(\frac{\sigma^i p_i}{E+m}\right)\alpha(p_\alpha)\\ \left(E^2 - \mathbf{p}^2 - m^2\right)\alpha(p_\alpha) &= 0\end{aligned}$$

where we use $\left(\sigma^i p_i\right)^2 = (-p^i)(-p^i) = \mathbf{p}^2$ in the last line. This just gives the usual relativistic expression relating mass, energy and momentum, with positive energy solution

$$E = \pm\sqrt{\mathbf{p}^2 + m^2} \tag{6.7}$$

This determines the possible energies. Notice that despite the first order character of the Dirac equation, E may still take positive and negative values.

6.1.1 Positive energy eigenstates

Now we need the eigenstates. These must satisfy Eqs.(6.6) and (6.7) with no further constraint on $\alpha(p_\alpha)$. We are free to choose any convenient pair of 2-spinors for $\alpha(p_\alpha)$. Therefore, let

$$\alpha_1(p_\alpha) \equiv \begin{pmatrix}1\\0\end{pmatrix} \tag{6.8}$$

$$\alpha_2(p_\alpha) \equiv \begin{pmatrix}0\\1\end{pmatrix} \tag{6.9}$$

For $\alpha_1(p_\alpha)$, (remembering that $p_i = -p^i$) we must have

$$\begin{aligned}\beta_1(p_\alpha) &= -\left(\frac{\sigma^i p_i}{E+m}\right)\alpha_1(p_\alpha)\\ &= \frac{1}{E+m}\begin{pmatrix}p^z & p^x - ip^y\\ p^x + ip^y & -p^z\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\\ &= \frac{1}{E+m}\begin{pmatrix}p^z\\ p^x + ip^y\end{pmatrix}\end{aligned}$$

while for $\alpha_2(p_\alpha)$ we find

$$\begin{aligned}\beta_2(p_\alpha) &= -\left(\frac{\sigma^i p_i}{E+m}\right)\alpha_2(p_\alpha)\\ &= \frac{1}{E+m}\begin{pmatrix}p^z & p^x - ip^y\\ p^x + ip^y & -p^z\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\\ &= \frac{1}{E+m}\begin{pmatrix}p^x - ip^y\\ -p^z\end{pmatrix}\end{aligned}$$

These relations define two independent, positive energy solutions, which we normalize and denote by $u_a(p^\alpha)$:

$$[u_1(p^\alpha)]^A = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \frac{p^z}{E+m} \\ \frac{p^x+ip^y}{E+m} \end{pmatrix} \tag{6.10}$$

$$[u_2(p^\alpha)]^A = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ 1 \\ \frac{p^x-ip^y}{E+m} \\ \frac{-p^z}{E+m} \end{pmatrix} \tag{6.11}$$

Exercise: Show that $u_1(p^\alpha)$ and $u_2(p^\alpha)$ are orthonormal,

$$\langle u_a, u_b \rangle = \delta_{ab} \tag{6.12}$$

where the inner product of two spinors is given by

$$\langle \chi, \psi \rangle \equiv \chi^\dagger h \psi = \bar{\chi}\psi \tag{6.13}$$

with h given by eq.(2.45). Notice that this inner product is Lorentz invariant, so our spinor basis remains orthonormal in every frame of reference.

6.1.2 Negative energy eigenstates

For the second set of mode amplitudes, we solve Eq.(6.3), which becomes

$$\begin{aligned} 0 &= (\gamma^\alpha p_\alpha + m)\, v(p^\alpha) \\ &= \begin{pmatrix} E+m & \sigma^i p_i \\ -\sigma^i p_i & -E+m \end{pmatrix} \begin{pmatrix} \alpha(p_\alpha) \\ \beta(p_\alpha) \end{pmatrix} \end{aligned}$$

Solving for $\alpha(p_\alpha)$ first:

$$\alpha(p_\alpha) = -\frac{\sigma^i p_i}{E+m}\beta(p_\alpha) \tag{6.14}$$

Once again, substituging into the second equation yields $E^2 - \mathbf{p}^2 - m^2 = 0$, so that $E = \pm\sqrt{\mathbf{p}^2 + m^2}$. There are again two solutions. Since $\beta(p_\alpha)$ is arbitrary and $\alpha(p_\alpha)$ is given by eq.(6.14), we choose

$$\beta_1(p_\alpha) \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{6.15}$$

$$\beta_2(p_\alpha) \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.16}$$

leading to two more independent, normalized solutions, $v_a(p^\alpha)$,

$$[v_1(p^\alpha)]^A = \sqrt{\frac{m+E}{2m}} \begin{pmatrix} \frac{p^z}{E+m} \\ \frac{p^x+ip^y}{E+m} \\ 1 \\ 0 \end{pmatrix} \tag{6.17}$$

$$[v_2(p^\alpha)]^A = \sqrt{\frac{m+E}{2m}} \begin{pmatrix} \frac{p^x-ip^y}{E+m} \\ \frac{-p^z}{E+m} \\ 0 \\ 1 \end{pmatrix} \tag{6.18}$$

The entire set of four spinors, $u_a(p^\alpha), v_a(p^\alpha),\ a = 1,2$ is a complete, pseudo-orthonormal basis.

Exercise: Check that $v_1(p^\alpha)$ and $v_2(p^\alpha)$ satisfy

$$\langle v_a(p^\alpha), v_b(p^\alpha)\rangle = -\delta_{ab} \tag{6.19}$$
$$\langle u_a(p^\alpha), v_b(p^\alpha)\rangle = 0 \tag{6.20}$$

Exercise: Prove the completeness relation,

$$\sum_{a=1}^{2}\left([u_a(p^\alpha)]^A [\bar{u}_a(p^\alpha)]_B - [v_a(p^\alpha)]^A [\bar{v}_a(p^\alpha)]_B\right) = \delta^A_B \tag{6.21}$$

where $A, B = 1, \ldots, 4$ index the components of the basis spinors.

The completeness relation, Eq.(6.21), guarantees that any spinor may be written as a linear combination of $(u_a(p^\alpha), v_a(p^\alpha),\ a = 1, 2)$, for given any spinor ψ^A we may write

$$\begin{aligned}
\psi^A &= \delta^A_B \psi^B \\
&= \sum_{a=1}^{2}\left([u_a(p^\alpha)]^A [\bar{u}_a(p^\alpha)]_B - [v_a(p^\alpha)]^A [\bar{v}_a(p^\alpha)]_B\right)\psi^B \\
&= \sum_{a=1}^{2}\left([u_a(p^\alpha)]^A (\bar{u}_a\psi) - [v_a(p^\alpha)]^A (\bar{v}_a\psi)\right) \\
&= \sum_{a=1}^{2}(\bar{u}_a\psi)\, u_a^A - \sum_{a=1}^{2}(\bar{v}_a\psi)\, v_a^A
\end{aligned}$$

where the four complex numbers $(\bar{u}_a\psi), (\bar{v}_a\psi)$ give the components in our basis.

6.1.3 Fourier superposition

Using this basis, we now have a complete solution to the free Dirac equation. Using $\Theta(E)$ to enforce positive energy condition, we have

$$\begin{aligned}
\psi(\mathbf{x}, t) &= \frac{1}{(2\pi)^{3/2}}\sum_{a=1}^{2}\int d^4k\; 2\sqrt{m\omega}\,\delta\left(E^2 - \mathbf{p}^2 - m^2\right)\Theta(E)\Big(b_a(p^\alpha)\, u_a(p^\alpha)\, e^{-\frac{i}{\hbar}p_\alpha x^\alpha} \\
&\quad + d_a^\dagger(p^\alpha)\, v_a^\dagger(p^\alpha)\, e^{\frac{i}{\hbar}p_\alpha x^\alpha}\Big)
\end{aligned}$$

Here the Dirac delta function imposes the energy condition while the spinor plane waves $u_a(p^\alpha)\, e^{-\frac{i}{\hbar}p_\alpha x^\alpha}$ and $v_a^\dagger(p^\alpha)\, e^{\frac{i}{\hbar}p_\alpha x^\alpha}$ satisfy the Dirac equation for any fixed p^α. The arbitrary mode amplitudes $b_a(p)$ and $d_a^\dagger(p)$ then form an arbitrary superposition. The factor $2\sqrt{m\omega}$ is a convenient normaliztion where we define $\omega \equiv +\sqrt{\mathbf{k}^2 + m^2}$. Replacing the delta function using Eq.(4.21) $\Theta(E)\,\delta\left(E^2 - \mathbf{p}^2 - m^2\right) = \frac{1}{2\omega}\delta\left(E - \sqrt{\mathbf{p}^2 + m^2}\right)$ and integrating over k^0,

$$\psi(\mathbf{x}, t) = \frac{1}{(2\pi)^{\frac{3}{2}}}\sum_{a=1}^{2}\int d^3k\sqrt{\frac{m}{\omega}}\left(b_a(\mathbf{k})\, u_a(\mathbf{k})\, e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} + d_a^\dagger(\mathbf{k}))v_a(\mathbf{k})\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})}\right) \tag{6.22}$$

Before turning to quantization, we consider the spin of spinors.

Notice that our writing $d_i^\dagger(\mathbf{k})$ instead of $d_i(\mathbf{k})$ in the expansion of ψ, while perfectly allowable, is consistent with what we found for scalar fields. It is purely a matter of definition. However, when we look at the commutation relations of the corresponding operators, this part of the field operator $\hat{\psi}$ should create an antiparticle, and therefore is most appropriately called $d_i^\dagger(\mathbf{k})$. This is consistent with $CP\Theta$ symmetry of the field.

6.2 The energy and spin of spinors

6.2.1 Energy projections

The basis spinors $(u_a(p^\alpha), v_a(p^\alpha))$ may be thought of as eigenvectors of the operator $p_\alpha\gamma^\alpha$. Rewriting Eqs.(6.2) and (6.3),

$$\begin{aligned} \gamma^\alpha p_\alpha u_a(p^\alpha) &= m u_a(p^\alpha) \\ \gamma^\alpha p_\alpha v_a(p^\alpha) &= -m v_a(p^\alpha) \end{aligned}$$

This allows us to construct projection operators that single out the $u_a(p^\alpha)$- and $v_a(p^\alpha)$-type spinors. We add the normalized operators $\pm\frac{1}{m}\gamma^\alpha p_\alpha$, normalized by the eigenvalus to the identity,

$$P_\pm = \frac{1}{2}\left(\mathbf{1} \pm \frac{1}{m}\gamma^\alpha p_\alpha\right) \tag{6.23}$$

then the resulting $P_\pm$ are idempotent,

$$\begin{aligned} P_\pm^2 &= \frac{1}{4}\left(\mathbf{1} \pm \frac{1}{m}\gamma^\alpha p_\alpha\right)\left(\mathbf{1} \pm \frac{1}{m}\gamma^\beta p_\beta\right) \\ &= \frac{1}{4}\left(\mathbf{1} \pm \frac{2}{m}\gamma^\alpha p_\alpha + \frac{1}{m^2}\gamma^\alpha p_\alpha \gamma^\beta p_\beta\right) \\ &= \frac{1}{4}\left(\mathbf{1} \pm \frac{2}{m}\gamma^\alpha p_\alpha + \frac{1}{m^2}p^2\right) \\ &= P_\pm \end{aligned}$$

where $\gamma^\alpha p_\alpha \gamma^\beta p_\beta = p^2 = m^2$. They are also orthogonal to one another,

$$\begin{aligned} P_+ P_- &= \frac{1}{4}\left(\mathbf{1} + \frac{1}{m}\gamma^\alpha p_\alpha\right)\left(\mathbf{1} - \frac{1}{m}\gamma^\beta p_\beta\right) \\ &= \frac{1}{4}\left(\mathbf{1} - \frac{1}{m^2}\gamma^\alpha p_\alpha \gamma^\beta p_\beta\right) \\ &= \frac{1}{4}\left(\mathbf{1} - \frac{1}{m^2}m^2\right) \\ &= 0 \end{aligned}$$

and together span the full space,

$$\begin{aligned} P_+ + P_- &= \frac{1}{2}\left(\mathbf{1} + \frac{1}{m}\gamma^\alpha p_\alpha\right) + \frac{1}{2}\left(\mathbf{1} - \frac{1}{m}\gamma^\alpha p_\alpha\right) \\ &= 1 \end{aligned}$$

Returning to the action of $P_\pm$ on states, we clearly have

$$\begin{aligned} P_+ u_a(p^\alpha) &= u_a(p^\alpha) \\ P_+ v_a(p^\alpha) &= 0 \end{aligned}$$

and we may build P_+ out of its nonzero eigenvectors,

$$P_+ = \sum_{a=1}^{2} u_a(p^\alpha)\bar{u}_a(p^\alpha) \tag{6.24}$$

Similarly, we have

$$\begin{aligned} P_- u_a(p^\alpha) &= 0 \\ P_- v_a(p^\alpha) &= v_a(p^\alpha) \\ P_- &= \sum_{a=1}^{2} v_a(p^\alpha)\bar{v}_a(p^\alpha) \end{aligned} \tag{6.25}$$

The operators $P_\pm$ distinguish between the positive and negative energy states.

6.2.2 Spin projections

Next, we seek a pair of operators which distinguishes between u_1 and u_2 and between v_1 and v_2. Since u_a and v_a are pseudo-orthonormal, Eqs.(6.12), (6.19) and (6.20), we can simply write

$$\begin{aligned} [\Pi_+]^A{}_B &= u_1 \otimes \bar{u}_1 - v_2 \otimes \bar{v}_2 \\ &= [u_1]^A \left[\gamma^0\right]_{BC} \left[u_1^\dagger\right]^C - [v_2]^A \left[\gamma^0\right]_{BC} \left[v_2^\dagger\right]^C \end{aligned}$$

to project into the u_1 and v_2 directions. This is immediately seen to satisfy $\Pi_+ u_2 = \Pi_+ v_1 = 0$ and $\Pi_+ u_1 = u_1$, $\Pi_+ v_2 = v_2$.

To find an explicit form for Π_+, consider first the rest frame of the particle, where the 4-momentum is given by $p^\alpha = (mc, 0)$. There we have

$$\begin{aligned} u_1(p^\alpha) &= (1,0,0,0) \\ u_2(p^\alpha) &= (0,1,0,0) \\ v_1(p^\alpha) &= (0,0,1,0) \\ v_2(p^\alpha) &= (0,0,0,1) \end{aligned} \tag{6.26}$$

so that

$$[\Pi_+]^A{}_B = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix}$$

This combination is easy to construct from the gamma matrices. With the forms given in Eq.(2.35) and with $\gamma_5 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, we note that

$$\gamma^3\gamma_5 = \begin{pmatrix} \sigma^3 & 0 \\ 0 & -\sigma^3 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix}$$

From this we see that *all* of the basis vectors, eq.(6.26), are eigenvectors of $\gamma^3\gamma_5$:

$$\begin{aligned} \gamma^3\gamma_5 u_1 &= u_1 \\ \gamma^3\gamma_5 u_2 &= -u_2 \\ \gamma^3\gamma_5 v_1 &= -v_1 \\ \gamma^3\gamma_5 v_2 &= v_2 \end{aligned}$$

and we construct two projection operators,

$$\begin{aligned}\Pi_+ &= \frac{1}{2}\left(1+\gamma^3\gamma_5\right) = \frac{1}{2}\left(1+n_\alpha\gamma^\alpha\gamma_5\right)\\ \Pi_- &= \frac{1}{2}\left(1-\gamma^3\gamma_5\right) = \frac{1}{2}\left(1-n_\alpha\gamma^\alpha\gamma_5\right)\end{aligned}$$

where $n_\alpha = (0,0,0,1)$. Notice that n_α is spacelike, with $n^2=-1$, and that $p^\alpha n_\alpha = 0$.

Now, we generalize these new projections by writing

$$\Pi_\pm \equiv \frac{1}{2}\left(1 \pm s_\mu\gamma^\mu\gamma_5\right) \tag{6.27}$$

where s_μ is any 4-vector. These are still projection operators provided s^α is spacelike, $s_\mu s_\nu \eta^{\mu\nu} = s^2 = -1$, since then we have

$$\begin{aligned}\Pi_\pm^2 &= \frac{1}{4}\left(1 \pm s_\mu\gamma^\mu\gamma_5\right)\left(1 \pm s_\nu\gamma^\nu\gamma_5\right)\\ &= \frac{1}{4}\left(1 \pm 2s_\mu\gamma^\mu\gamma_5 + s_\mu\gamma^\mu\gamma_5 s_\nu\gamma^\nu\gamma_5\right)\\ &= \frac{1}{4}\left(1 \pm 2s_\mu\gamma^\mu\gamma_5 - s_\mu s_\nu\gamma^\mu\gamma^\nu\gamma_5\gamma_5\right)\\ &= \frac{1}{4}\left(1 \pm 2s_\mu\gamma^\mu\gamma_5 - s_\mu s_\nu\eta^{\mu\nu}\right)\\ &= \Pi_\pm\end{aligned}$$

In addition, we can require these $\Pi_\pm$ to commute with P_+ and P_- . Consider

$$\begin{aligned}\left[\Pi_\pm, P_+\right] &= \left[\frac{1}{2}\left(1 \pm s_\mu\gamma^\mu\gamma_5\right), \frac{1}{2}\left(1+\frac{1}{m}\gamma^\alpha p_\alpha\right)\right]\\ &= \frac{1}{4}\left(1 \pm s_\mu\gamma^\mu\gamma_5 + \frac{1}{m}\gamma^\alpha p_\alpha \pm \frac{1}{m}s_\mu p_\alpha\gamma^\mu\gamma_5\gamma^\alpha\right)\\ &\quad -\frac{1}{4}\left(1+\frac{1}{m}\gamma^\alpha p_\alpha \pm s_\mu\gamma^\mu\gamma_5 \pm \frac{1}{m}p_\alpha s_\mu\gamma^\alpha\gamma^\mu\gamma_5\right)\\ &= \pm\frac{1}{4m}\left(-s_\mu p_\alpha\gamma^\mu\gamma^\alpha\gamma_5 + p_\alpha s_\mu\gamma^\alpha\gamma^\mu\gamma_5\right)\\ &= -\frac{1}{4m}s_\mu p_\alpha\left(\gamma^\mu\gamma^\alpha + \gamma^\alpha\gamma^\mu\right)\gamma_5\\ &= -\frac{1}{2m}s_\mu p_\alpha\eta^{\mu\alpha}\gamma_5\end{aligned}$$

This will vanish if s^α and p_α are orthogonal, $s^\alpha p_\alpha = 0$.

The vector s^α is the 4-dimensional generalization of the spin direction, s^i, reducing in the rest frame to $s^\alpha = \left(0, s^i\right)$ but transforming as a 4-vector.

Spin operators.

In quantum mechanics, the Pauli matrices σ_i give us three spin operators,

$$S_i = \frac{\hbar}{2}\sigma_i$$

In quantum field theory, we start instead with the spin representation of the Poincaré group, where we found the Lorentz generators,

$$\sigma^{\mu\nu} = \left[\gamma^\mu, \gamma^\nu\right]$$

These generate both rotations and boosts, and we are now interested in the rotational part. The rotations are given by the part of $\sigma^{\mu\nu}$ orthogonal to the momentum 4-vector,

$$J^{\alpha} \equiv \frac{1}{2}\varepsilon^{\alpha}{}_{\beta\mu\nu}P^{\beta}M^{\mu\nu}$$

u_a and v_a are eigenvectors of the z-component of spin.

Since, $P_+P_- = 0$ and $\Pi_+\Pi_- = 0$, the set of projection operators,

$$\{P_+, P_-, \Pi_+, \Pi_-\}$$

is fully commuting and therefore simultaneously diagonalizable. Moreover, they are independent. To see this, consider the four products

$$\{P_+\Pi_+, P_+\Pi_-, P_-\Pi_+, P_-\Pi_-\}$$

These are mutually orthogonal, i.e., $(P_+\Pi_+)(P_+\Pi_-) = P_+P_+\Pi_+\Pi_- = 0$ and so on. Each combination projects into a 1-dimensional subspace of the spinor space since,

$$\begin{aligned} tr\left(P_+\Pi_+\right) &= \frac{1}{4}tr\left(1 + s_\mu\gamma^\mu\gamma_5 + \frac{1}{m}\gamma^\alpha p_\alpha + \frac{1}{m}s_\mu p_\alpha\gamma^\mu\gamma_5\gamma^\alpha\right) \\ &= \frac{1}{4}\left(4+0+0+0\right) \\ &= 1 \end{aligned}$$

and similarly $tr\left(P_+\Pi_-\right) = tr\left(P_-\Pi_+\right) = tr\left(P_-\Pi_-\right) = 1$. Finally, they span the 4-dimensional space as we see from the completeness relation:

$$\begin{aligned} P_+\Pi_+ + P_+\Pi_- + P_-\Pi_+ + P_-\Pi_- &= P_+\left(\Pi_+ + \Pi_-\right) + P_-\left(\Pi_+ + \Pi_-\right) \\ &= P_+ + P_- \\ &= \mathbf{1} \end{aligned}$$

We are free to choose u_a and v_a to be eigenvectors of any 3-vector s^i, and therefore eigenspinors of the corresponding $\Pi_+(s^\alpha), \Pi_-(s^\alpha)$. As a result, we can label the spinors by their 4-momentum and their spin vectors,

$$\begin{aligned} &u_a\left(p^\alpha, s^\beta\right) \\ &v_a\left(p^\alpha, s^\beta\right) \end{aligned}$$

and since we have expressed the parameterization in terms of 4-vectors we have

$$\begin{aligned} \Pi_+ &= u_1\left(p^\alpha, s^\beta\right)\bar{u}_1\left(p^\alpha, s^\beta\right) - v_2\left(p^\alpha, s^\beta\right)\bar{v}_2\left(p^\alpha, s^\beta\right) = \frac{1}{2}\left(1 + s_\mu\gamma^\mu\gamma_5\right) \\ \Pi_- &= u_2\left(p^\alpha, s^\beta\right)\bar{u}_2\left(p^\alpha, s^\beta\right) - v_1\left(p^\alpha, s^\beta\right)\bar{v}_1\left(p^\alpha, s^\beta\right) = \frac{1}{2}\left(1 - s_\mu\gamma^\mu\gamma_5\right) \end{aligned}$$

in any frame of reference and for any choice of spin direction satisfying $s_\alpha s^\alpha = -1$ and $s^\alpha p_\alpha = 0$.

Using these expressions for the spin projection operators together with the corresponding expressions, eqs.(6.24) and (6.25), for the energy, we can rewrite the outer products of the completeness relation, eq.(6.21), as

$$\begin{aligned}
u_1\left(p^\alpha, s^\beta\right)\bar{u}_1\left(p^\alpha, s^\beta\right) &= P_+\Pi_+ \\
&= \frac{1}{2}\left(\mathbf{1}+\frac{1}{m}\gamma^\alpha p_\alpha\right)\frac{1}{2}\left(1+s_\mu\gamma^\mu\gamma_5\right) \\
u_2\left(p^\alpha, s^\beta\right)\bar{u}_2\left(p^\alpha, s^\beta\right) &= P_+\Pi_- \\
&= \frac{1}{2}\left(\mathbf{1}+\frac{1}{m}\gamma^\alpha p_\alpha\right)\frac{1}{2}\left(1-s_\mu\gamma^\mu\gamma_5\right) \\
&= u_1\left(p^\alpha, -s^\beta\right)\bar{u}_1\left(p^\alpha, -s^\beta\right) \\
v_1\left(p^\alpha, s^\beta\right)\bar{v}_1\left(p^\alpha, s^\beta\right) &= -P_-\Pi_- \\
&= -\frac{1}{2}\left(\mathbf{1}-\frac{1}{m}\gamma^\alpha p_\alpha\right)\frac{1}{2}\left(1-s_\mu\gamma^\mu\gamma_5\right) \\
v_1\left(p^\alpha, s^\beta\right)\bar{v}_1\left(p^\alpha, s^\beta\right) &= -P_-\Pi_+ \\
&= -\frac{1}{2}\left(\mathbf{1}-\frac{1}{m}\gamma^\alpha p_\alpha\right)\frac{1}{2}\left(1+s_\mu\gamma^\mu\gamma_5\right) \\
&= v_1\left(p^\alpha, -s^\beta\right)\bar{v}_1\left(p^\alpha, -s^\beta\right)
\end{aligned}$$

These identities will be useful for calculating scattering amplitudes, allowing us to express various spinor products in terms of gamma matrices.

6.3 Hamiltonian formulation

Now we turn to the quantization of the Dirac field. We rewrite the action, Eq.(2.47),

$$S = \int d^4x\ \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi$$

The conjugate momentum to ψ is the spinor field

$$\begin{aligned}
\pi_A &\equiv \frac{\delta L}{\delta\left(\partial_0\psi^A\right)} = i\bar{\psi}\gamma^0 \\
&= i\left[\psi^\dagger\right]^B h_{BC}\left[\gamma^0\right]^C{}_A
\end{aligned} \tag{6.28}$$

We can also write this as

$$\pi\gamma^0 = i\bar{\psi}$$

and in a basis where, numerically $h_{AB} = \left[\gamma^0\right]^A{}_B$, as

$$\pi = i\psi^\dagger$$

Undaunted by the peculiar lack of a time derivative in the momentum, we press on with the Hamiltonian:

$$\begin{aligned}
H &= \int d^3x\ \left(\pi\dot{\psi} - \mathcal{L}\right) \\
&= \int d^3x\ \left(i\bar{\psi}\gamma^0\partial_0\psi - \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi\right) \\
&= \int d^3x\ \left(i\bar{\psi}\gamma^0\partial_0\psi - i\bar{\psi}\gamma^0\partial_0\psi - i\bar{\psi}\gamma^i\partial_i\psi + m\bar{\psi}\psi\right)
\end{aligned}$$

The time derivatives *cancel*, leaving

$$\begin{aligned} H &= \int d^3x \, \left(-i\bar{\psi}\gamma^i\partial_i\psi + m\bar{\psi}\psi\right) \\ &= i\int d^3x \; i\bar{\psi}\left(i\gamma^i\partial_i - m\right)\psi \\ &= i\int d^3x \; \pi\gamma^0\left(i\gamma^i\partial_i - m\right)\psi \end{aligned} \tag{6.29}$$

Once again, we are struck by the absence of time derivatives in the energy. If we ust the field equation, we may rewrite H as

$$\begin{aligned} H &= i\int d^3x \; \pi\gamma^0\left(i\gamma^i\partial_i - m\right)\psi \\ &= \int d^3x \; \pi\gamma^0\left(\gamma^0\partial_0\right)\psi \\ &= \int d^3x \; \pi\partial_0\psi \end{aligned} \tag{6.30}$$

but since this assumes the field equation we cannot use this form to write Hamilton's equations. Indeed, computing $\partial_0\psi$ using Eq.(6.30) merely gives an identity,

$$\begin{aligned} \partial_0\psi &= \int d^3x' \left(\frac{\delta H(x)}{\delta\pi(x')}\frac{\delta\psi(x)}{\delta\psi(x')} - \frac{\delta H(x)}{\delta\psi(x')}\frac{\delta\psi(x)}{\delta\pi(x')}\right) \\ &= \int d^3x' \left(\partial_0\psi(x')\right)\delta^3(x-x') \\ &= \partial_0\psi(x) \end{aligned}$$

Still, eq.(6.30) is useful for computing the operator form of the Hamiltonian from solutions.

We may find the field equation using Eq.(6.29) for H,

$$\begin{aligned} \partial_0\psi &= \{H,\psi\} \\ &= \int d^3x' \left(\frac{\delta H(x)}{\delta\pi(x')}\frac{\delta\psi(x)}{\delta\psi(x')} - \frac{\delta H(x)}{\delta\psi(x')}\frac{\delta\psi(x)}{\delta\pi(x')}\right) \\ &= \int d^3x' \; i\gamma^0\left(i\gamma^i\partial_i - m\right)\psi(x') \; \delta^3(x-x') \\ &= i\gamma^0\left(i\gamma^i\partial_i - m\right)\psi(x) \end{aligned}$$

Multiplying by $i\gamma^0$ this becomes $i\gamma^0\partial_0\psi = -\left(i\gamma^i\partial_i - m\right)\psi(x)$ and combining the gamma terms as a sum, we recover the Dirac equation,

$$\left(i\gamma^\alpha\partial_\alpha - m\right)\psi(x) = 0$$

For the momentum Hamilton equation, we find the conjugate field equation:

$$\begin{aligned} \partial_0\pi &= \{H,\pi\} \\ &= \int d^3x' \left(\frac{\delta H(x)}{\delta\pi(x')}\frac{\delta\pi(x)}{\delta\psi(x')} - \frac{\delta H(x)}{\delta\psi(x')}\frac{\delta\pi(x)}{\delta\pi(x')}\right) \\ &= \int d^3x' \left(-\frac{\delta H(x)}{\delta\psi(x')}\frac{\delta\pi(x)}{\delta\pi(x')}\right) \end{aligned}$$

To find $\frac{\delta H(x)}{\delta\psi(x')}$ we integrate the Hamiltonian by parts,

$$\begin{aligned}
H &= i\int d^3x\ \pi\gamma^0\left(i\gamma^i\partial_i - m\right)\psi \\
&= \int d^3x\ \left(i\pi\gamma^0 i\gamma^i\partial_i\psi - mi\pi\gamma^0\psi\right) \\
&= \int d^3x\ \left(-i\partial_i\pi\gamma^0 i\gamma^i\psi - mi\pi\gamma^0\psi\right) \\
&= -\int d^3x\ i\left(i\partial_i\pi\gamma^0\gamma^i + m\pi\gamma^0\right)\psi
\end{aligned}$$

The first term is often written as

$$i\partial_i\pi\gamma^0\gamma^i = i\pi\gamma^0\gamma^i\overleftarrow{\partial}_i$$

to indicate that the derivative acts to the left on π, without having to re-order the terms. Then the Hamiltonian becomes

$$\begin{aligned}
H &= -i\int d^3x\ \left(i\pi\gamma^0\gamma^i\overleftarrow{\partial}_i + m\pi\gamma^0\right)\psi \\
&= -i\int d^3x\ \pi\gamma^0\left(i\gamma^i\overleftarrow{\partial}_i + m\right)\psi
\end{aligned}$$

an the functional derivative is

$$\frac{\delta H(x)}{\delta\psi(x')} = -i\pi\gamma^0\left(i\gamma^i\overleftarrow{\partial}_i + m\right)$$

The Hamilton equation for π is therefore

$$\begin{aligned}
\partial_0\pi &= i\pi\gamma^0\left(i\gamma^i\overleftarrow{\partial}_i + m\right) \\
i\bar{\psi}\gamma^0\overleftarrow{\partial}_0 &= -\bar{\psi}\left(i\gamma^i\overleftarrow{\partial}_i + m\right)
\end{aligned}$$

giving the conjugate Dirac equation,

$$\bar{\psi}\left(i\gamma^\alpha\overleftarrow{\partial}_\alpha + m\right) = 0 \qquad (6.31)$$

Finally, we write the fundamental Poisson brackets,

$$\begin{aligned}
\left\{\psi^A(\mathbf{x},t),\psi^B(\mathbf{x}',t)\right\}_{PB} &= 0 \\
\left\{\pi_A(\mathbf{x},t),\psi^B(\mathbf{x}',t)\right\}_{PB} &= \delta^B_A\delta^3(\mathbf{x}-\mathbf{x}') \\
\left\{\pi_A(\mathbf{x},t),\pi_B(\mathbf{x}',t)\right\}_{PB} &= 0
\end{aligned} \qquad (6.32)$$

We may now quantize the classical solution.

6.3.1 Quantization of the Dirac field

Following the rule for canonical quantization we used for scalar fields, we use Eqs.(6.32) to write the fundamental commutator of the spinor field as

$$\begin{aligned}
\left[\hat{\psi}^A(\mathbf{x},t),\hat{\psi}^B(\mathbf{x}',t)\right] &= 0 \\
\left[\hat{\pi}_A(\mathbf{x},t),\hat{\psi}^B(\mathbf{x}',t)\right] &= i\delta^B_A\delta^3(\mathbf{x}-\mathbf{x}') \\
\left[\hat{\pi}_A(\mathbf{x},t),\hat{\pi}_B(\mathbf{x}',t)\right] &= 0
\end{aligned} \qquad (6.33)$$

with the caution that these will be modified for reasons discussed below. We can immediately turn to our examination of the commutation relations of the mode amplitudes.

6.3.1.1 Solving for the mode amplitudes

From the classical solution given by Eq.(6.22),

$$\psi(\mathbf{x},t)=\frac{1}{(2\pi)^{3/2}}\sum_{a=1}^{2}\int d^3k\sqrt{\frac{m}{\omega}}\left(b_a(\mathbf{k})\,u_a(\mathbf{k})\,e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}+d_a^*(\mathbf{k}))v_a(\mathbf{k})\,e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)$$

we immetiately find the conjugate momentum, $\pi=i\bar{\psi}\gamma^0$,

$$\pi(\mathbf{x},t) = \frac{i}{(2\pi)^{3/2}}\sum_{a=1}^{2}\int d^3k\sqrt{\frac{m}{\omega}}\left(b_a^*(\mathbf{k})\bar{u}_a(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}+d_a(\mathbf{k})\bar{v}_a(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)\gamma^0$$

and we may solve for the amplitudes. In addition to the sort of linear combination we used for scalar fields, however, we have to eliminate the basis spinors to isolate the mode amplitudes.

Setting $t=t'=0$, we first invert the Fourier transform:

$$\begin{aligned}
\tilde{\psi}(\mathbf{k}) &\equiv \frac{1}{(2\pi)^{3/2}}\int\psi(\mathbf{x},0)e^{-i\mathbf{k}\cdot\mathbf{x}}d^3x\\
&= \frac{1}{(2\pi)^{3}}\sum_{a=1}^{2}\int d^3x\int d^3k'\sqrt{\frac{m}{\omega'}}\left(b_a(\mathbf{k}')\,u_a(\mathbf{k}')\,e^{i\left(\mathbf{k}'-\mathbf{k}\right)\cdot\mathbf{x}}+d_a^*(\mathbf{k}')\,v_a(\mathbf{k}')\,e^{-i\left(\mathbf{k}'+\mathbf{k}\right)\cdot\mathbf{x}}\right)\\
&= \sum_{a=1}^{2}\int d^3k'\sqrt{\frac{m}{\omega'}}\left(b_a(\mathbf{k}')\,u_a(\mathbf{k}')\,\delta^3(\mathbf{k}'-\mathbf{k})+d_a^*(\mathbf{k}')\,v_a(\mathbf{k}')\,\delta^3(\mathbf{k}'+\mathbf{k})\right)\\
&= \sum_{a=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_a(\mathbf{k})\,u_a(\mathbf{k})+d_a^*(-\mathbf{k})\,v_a(-\mathbf{k})\right)
\end{aligned}$$

and for π, we immediately find

$$\begin{aligned}
\tilde{\psi}^\dagger(\mathbf{k}) &\equiv \frac{1}{(2\pi)^{3/2}}\int\psi^\dagger(\mathbf{x},0)e^{i\mathbf{k}\cdot\mathbf{x}}d^3x\\
&= \sum_{a=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_a^*(\mathbf{k})u_a^\dagger(\mathbf{k})+d_a(-\mathbf{k})v_a^\dagger(-\mathbf{k})\right)
\end{aligned}\qquad(6.34)$$

so that

$$\begin{aligned}
\tilde{\pi}(\mathbf{k}) &= i\tilde{\psi}^\dagger(\mathbf{k})\,h\gamma^0\\
&= i\bar{\tilde{\psi}}(\mathbf{k})\,\gamma^0\\
&= i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_a^*(\mathbf{k})\,u_a^\dagger(\mathbf{k})+d_a(-\mathbf{k})\,v_a^\dagger(-\mathbf{k})\right)h\gamma^0\\
&= i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_a^*(\mathbf{k})\,\bar{u}_a(\mathbf{k})+d_a(-\mathbf{k})\,\bar{v}_a(-\mathbf{k})\right)\gamma^0
\end{aligned}$$

Now, we would like to use the spinor inner product to isolate b_a and d_a. However, since $\tilde{\psi}(\mathbf{k})$ involves $v_j(-\mathbf{k})$ instead of $v_j(\mathbf{k})$, we need a modified form of the orthonormality relation. From the form of our solution for $v_j(\mathbf{k})$, we immediately see that

$$
\begin{aligned}
v_1(-\mathbf{k}) &= \sqrt{\frac{m+\omega}{2m}}\begin{pmatrix} \frac{-k^z}{\omega+m} \\ -\frac{k^x+ik^y}{\omega+m} \\ 1 \\ 0 \end{pmatrix} = -\gamma^0 v_1(\mathbf{k}) \\
v_2(-\mathbf{k}) &= \sqrt{\frac{m+\omega}{2m}}\begin{pmatrix} -\frac{k^x-ik^y}{\omega+m} \\ \frac{k^z}{\omega+m} \\ 0 \\ 1 \end{pmatrix} = -\gamma^0 v_2(\mathbf{k}) \qquad (6.35)
\end{aligned}
$$

We also need

$$\bar{v}_i(-\mathbf{k}) = v_i^\dagger(-\mathbf{k})\,h = \left(-\gamma^0 v_i(\mathbf{k})\right)^\dagger h = -v_i^\dagger(\mathbf{k})\,\gamma^0 h$$

as well as two more identities to reach our goal.

Exercise: Show that

$$
\begin{aligned}
\bar{u}_a\gamma^0 u_b &= \frac{\omega}{m}\delta_{ab} \\
\bar{v}_a\gamma^0 v_b &= \frac{\omega}{m}\delta_{ab} \qquad (6.36)
\end{aligned}
$$

In explicit components, these are

$$
\begin{aligned}
\left[u_a^\dagger(\mathbf{k})\right]^C h_{CB}\left[\gamma^0\right]^B{}_A\, u_b^A(\mathbf{k}) &= \frac{\omega}{m}\delta_{ab} \\
\left[v_a^\dagger(\mathbf{k})\right]^C h_{CB}\left[\gamma^0\right]^B{}_A\, v_b^A(\mathbf{k}) &= \frac{\omega}{m}\delta_{ab}
\end{aligned}
$$

Returning to the Fourier transforms,

$$\tilde{\psi}^\dagger(\mathbf{k}) = \sum_{a=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_a^*(\mathbf{k})\,u_a^\dagger(\mathbf{k}) + d_a(-\mathbf{k})\,v_a^\dagger(-\mathbf{k})\right)$$

so that

$$
\begin{aligned}
\tilde{\pi}(\mathbf{k}) &= i\tilde{\psi}^\dagger(\mathbf{k})\,h\gamma^0 \\
&= i\bar{\tilde{\psi}}(\mathbf{k})\,\gamma^0 \\
&= i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_a^*(\mathbf{k})\,u_a^\dagger(\mathbf{k}) + d_a(-\mathbf{k})\,v_a^\dagger(-\mathbf{k})\right)h\gamma^0 \\
&= i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_a^*(\mathbf{k})\,\bar{u}_a(\mathbf{k}) + d_a(-\mathbf{k})\,\bar{v}_a(-\mathbf{k})\right)\gamma^0
\end{aligned}
$$

where we used $\gamma^0 h\gamma^0 = h$. As a result,

$$
\begin{aligned}
\bar{u}_i(\mathbf{k})\,\gamma^0\tilde{\psi}(\mathbf{k}) &= \sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_j(\mathbf{k})\bar{u}_i(\mathbf{k})\gamma^0 u_j(\mathbf{k}) - d_j^\dagger(-\mathbf{k})\bar{u}_i(\mathbf{k})v_j(\mathbf{k})\right) \\
&= \sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}b_j(\mathbf{k})\bar{u}_i(\mathbf{k})\,\gamma^0 u_j(\mathbf{k}) \\
&= \sqrt{\frac{\omega}{m}}b_i(\mathbf{k})
\end{aligned}
$$

and similarly

$$\begin{aligned}
\bar{v}_i(\mathbf{k})\,\tilde{\psi}(\mathbf{k}) &= \sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_j(\mathbf{k})\,\bar{v}_i(\mathbf{k})\,u_j(\mathbf{k}) - d_j^{\dagger}(-\mathbf{k})\,\bar{v}_i(\mathbf{k})\,\gamma^0 v_j(\mathbf{k})\right) \\
&= -\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}d_j^{*}(-\mathbf{k})\,\bar{v}_i(\mathbf{k})\,\gamma^0 v_j(\mathbf{k}) \\
&= -\sqrt{\frac{\omega}{m}}d_j^{*}(-\mathbf{k})
\end{aligned}$$

while for the momentum,

$$\begin{aligned}
\tilde{\pi}(\mathbf{k})\,u_i(\mathbf{k}) &= i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_j^{*}(\mathbf{k})u_j^{\dagger}(\mathbf{k})h\gamma^0 u_i(\mathbf{k}) - d_j(-\mathbf{k})\,v_j^{\dagger}(\mathbf{k})\,hu_i(\mathbf{k})\right) \\
&= i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}b_j^{*}(\mathbf{k})\,u_j^{\dagger}(\mathbf{k})\,h\gamma^0 u_i(\mathbf{k}) \\
&= i\sqrt{\frac{\omega}{m}}b_i^{*}(\mathbf{k})
\end{aligned}$$

and

$$\begin{aligned}
\tilde{\pi}(\mathbf{k})\,\gamma^0 v_i(\mathbf{k}) &= i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(b_j^{*}(\mathbf{k})\,u_j^{\dagger}(\mathbf{k})\,h\gamma^0\gamma^0 v_i(\mathbf{k}) - d_j(-\mathbf{k}))v_j^{\dagger}(\mathbf{k})\,h\gamma^0 v_i(\mathbf{k})\right) \\
&= -i\sum_{j=1}^{2}\sqrt{\frac{m}{\omega}}\left(d_j(-\mathbf{k})\,v_j^{\dagger}(\mathbf{k})\,h\gamma^0 v_i(\mathbf{k})\right) \\
&= -i\sqrt{\frac{\omega}{m}}d_j(-\mathbf{k})
\end{aligned}$$

Noting that

$$\begin{aligned}
\tilde{\psi}(\mathbf{k}) &\equiv \frac{1}{(2\pi)^{3/2}}\int \psi(\mathbf{x},0)\,e^{-i\mathbf{k}\cdot\mathbf{x}}d^3x \\
\tilde{\psi}^{\dagger}(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}}\int \psi^{\dagger}(\mathbf{x},0)\,e^{i\mathbf{k}\cdot\mathbf{x}}d^3x
\end{aligned}$$

we collect terms and replace the mode amplitudes by operators and the conjugates by adjoints. Then, using Eqs.(6.35)

$$\hat{b}_i(\mathbf{k}) = \sqrt{\frac{m}{\omega}}\frac{1}{(2\pi)^{3/2}}\int \bar{u}_i(\mathbf{k})\,\gamma^0\hat{\psi}(\mathbf{x},0)\,e^{-i\mathbf{k}\cdot\mathbf{x}}d^3x \tag{6.37}$$

$$\hat{b}_i^{\dagger}(\mathbf{k}) = \sqrt{\frac{m}{\omega}}\frac{1}{(2\pi)^{3/2}}\int \hat{\psi}^{\dagger}(\mathbf{x},0)\,h\gamma^0 u_i(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{x}}d^3x \tag{6.38}$$

$$\hat{d}_j(\mathbf{k}) = \sqrt{\frac{m}{\omega}}\frac{1}{(2\pi)^{3/2}}\int \hat{\psi}^{\dagger}(\mathbf{x},0)\,h\gamma^0 v_i(\mathbf{k})\,e^{-i\mathbf{k}\cdot\mathbf{x}}d^3x \tag{6.39}$$

$$\hat{d}_j^{\dagger}(\mathbf{k}) = \sqrt{\frac{m}{\omega}}\frac{1}{(2\pi)^{3/2}}\int \bar{v}_i(\mathbf{k})\,\gamma^0\hat{\psi}(\mathbf{x},0)\,e^{i\mathbf{k}\cdot\mathbf{x}}d^3x \tag{6.40}$$

Next we want to find the commutation relations satisfied by these mode amplitudes. We start from the fundamental commutators, Eqs.(6.33), rewriting the $\left(\hat{\pi}_A, \hat{\psi}^B\right)$ commutator by replacing $\hat{\pi}_A(\mathbf{x},t)$ with $i\hat{\psi}^\dagger(\mathbf{x},t)h\gamma^0$. We then have

$$
\begin{aligned}
i\left[\left[\hat{\psi}^\dagger(\mathbf{x},t)\right]^C h_{CD}\left[\gamma^0\right]^D{}_A, \left[\hat{\psi}(\mathbf{x}',t)\right]^B\right] &= i\delta^B_A\delta^3(\mathbf{x}-\mathbf{x}') \\
i\left[\left[\hat{\psi}^\dagger(\mathbf{x},t)\right]_D, \left[\hat{\psi}(\mathbf{x}',t)\right]^B\right]\left[\gamma^0\right]^D{}_A &= i\delta^B_A\delta^3(\mathbf{x}-\mathbf{x}') \\
\left[\left[\hat{\psi}^\dagger(\mathbf{x},t)\right]_C, \left[\hat{\psi}(\mathbf{x}',t)\right]^B\right] &= \left[\gamma^0\right]^B{}_C\delta^3(\mathbf{x}-\mathbf{x}')
\end{aligned}
$$

or simply

$$
\left[\hat{\psi}^\dagger(\mathbf{x},t)h, \hat{\psi}(\mathbf{x}',t)\right] = \gamma^0\delta^3(\mathbf{x}-\mathbf{x}')
$$

We are now in a position to compute the commutators of the mode operators

6.3.1.2 Anticommutation

Now consider the $\hat{b}_a(\mathbf{k}), \hat{b}^\dagger_b(\mathbf{k}')$ and $\hat{d}_j(\mathbf{k}), \hat{d}^\dagger_i(\mathbf{k}')$ commutators:

$$
\begin{aligned}
\left[\hat{b}_a(\mathbf{k}), \hat{b}^\dagger_b(\mathbf{k}')\right] &= \frac{m}{(2\pi)^3\omega}\int\int d^3x d^3x' e^{i\mathbf{k}'\cdot\mathbf{x}-i\mathbf{k}\cdot\mathbf{x}'}\bar{u}_{aC}(\mathbf{k})\left[\gamma^0\right]^C{}_D\left[\psi^D(\mathbf{x}',0), \left(\psi^\dagger(\mathbf{x},0)h\right)_A\right]\left[\gamma^0\right]^A{}_B\left[u_b(\mathbf{k}')\right]^B \\
&= -\frac{m}{(2\pi)^3\omega}\int\int d^3x d^3x' e^{i\mathbf{k}'\cdot\mathbf{x}-i\mathbf{k}\cdot\mathbf{x}'}\bar{u}_{aC}(\mathbf{k})\left[\gamma^0\right]^C{}_D\left[\gamma^0\right]^D{}_A\left[\gamma^0\right]^A{}_B\left[u_b(\mathbf{k}')\right]^B\delta^3(\mathbf{x}-\mathbf{x}') \\
&= -\frac{m}{(2\pi)^3\omega}\int d^3x e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}}\bar{u}_{aC}(\mathbf{k})\left[\gamma^0\right]^C{}_B\left[u_b(\mathbf{k}')\right]^B \\
&= -\frac{m}{\omega}\frac{1}{(2\pi)^3}\int d^3x\frac{\omega}{m}\delta_{ab}e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}} \\
&= -\delta^3(\mathbf{k}'-\mathbf{k})\delta_{ab}
\end{aligned}
$$

and

$$
\begin{aligned}
\left[\hat{d}_a(\mathbf{k}), \hat{d}^\dagger_b(\mathbf{k}')\right] &= \frac{m}{\omega}\frac{1}{(2\pi)^3}\int\int d^3x d^3x' e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot\mathbf{x}'}\left[\gamma^0\right]^D{}_E v^E_a(\mathbf{k})\left[\left[\psi^\dagger\right]^C(\mathbf{x}',0)h_{CD}, \psi^B(\mathbf{x},0)\right]\bar{v}_{bA}(\mathbf{k}')\left[\gamma^0\right]^A{}_B \\
&= \frac{m}{\omega}\frac{1}{(2\pi)^3}\int\int d^3x d^3x' e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot\mathbf{x}'}\left[\gamma^0\right]^D{}_E v^E_a(\mathbf{k})\left[\gamma^0\right]^B{}_D\delta^3(\mathbf{x}-\mathbf{x}')\bar{v}_{bA}(\mathbf{k}')\left[\gamma^0\right]^A{}_B \\
&= \frac{m}{\omega}\frac{1}{(2\pi)^3}\int\int d^3x d^3x' e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot\mathbf{x}'}\bar{v}_{bA}(\mathbf{k})\gamma^0 v_a(\mathbf{k}')\delta^3(\mathbf{x}-\mathbf{x}') \\
&= \frac{m}{\omega}\delta^3(\mathbf{k}-\mathbf{k}')\bar{v}_{bA}(\mathbf{k})\left[\gamma^0\right]^A{}_E v^E_b(\mathbf{k}') \\
&= \delta_{ab}\delta^3(\mathbf{k}-\mathbf{k}')
\end{aligned}
$$

This is just the relationship we expect for $\hat{d}_a(\mathbf{k})$ – the mode amplitudes $\hat{d}_a(\mathbf{k})$ and $\hat{d}^\dagger_a(\mathbf{k})$ act as annihilation and creation operators, respectively. However, commutator

$$
\left[\hat{b}_a(\mathbf{k}), \hat{b}^\dagger_b(\mathbf{k}')\right] = -\delta^3(\mathbf{k}-\mathbf{k}')\delta_{ab}
$$

has the wrong sign, with $\hat{b}_a(\mathbf{k})$ rather than $\hat{b}^\dagger_b(\mathbf{k})$ acting like the creation operator. However, $\hat{b}_a(\mathbf{k})$ multiplies $e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}$ while $\hat{d}^\dagger_a(\mathbf{k})$ multiplies the CPT conjugate of $e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}$. This is consistent with our identification of these modes as particles and antiparticles, respectively. As we shall see, this pairing of particle creation with antiparticle annihilation, and vice versa, is necessary for other reasons as well. The identification we have chosen is necessary for conservation of charge. In addition, particle-antiparticle annihilation would not work correctly – every interaction that created a particle would have to annihilate a particle. We do not observe this. What went wrong?

We have very little freedom for introducing a sign here. In particular, the bilinear form $v_i^\dagger \gamma^0 v_j$ is governed by the Lorentz invariance properties of the spinor products. An overall sign on the field or the momentum would change the sign of the $\hat{d}_a(\mathbf{k})$ commutator as well as the $\hat{b}_a(\mathbf{k})$ commutator, thereby merely displacing the problem. Moreover, since $\hat{b}_a(\mathbf{k})$ and $\hat{b}_b^\dagger(\mathbf{k})$ enter the commutator together, a relative sign in the definition of $\hat{b}_a(\mathbf{k})$ is cancelled by a corresponding sign from $\hat{b}_b^\dagger(\mathbf{k})$. The only place a sign enters in a way that we could change the outcome is in our use of the antisymmetry of the commutator. If this "bracket" of conjugate variables were symmetric instead of antisymmetric, the proper relationship would be restored. But recall that this bracket was imposed by fiat – it is simply a rule that says we should take Poisson brackets to field commutators to arrive at the quantum field theory from the classical field theory.

Of course, we know that using anticommutators for fermionic fields *is* the right answer – essentially all of the rigid structure of the world, from the discretely stacked energy levels of nucleons in the nucleus and electrons in atoms to the endstates of stars as white dwarfs and neutron stars, relies on the Pauli exclusion principle. This principle states that no two fermions can occupy the same state and it is enforced mathematically by requiring fermion fields to anticommute. Here, we see the principle emerging from field theory as a condition of CPT invariance. Below, we will see that the same conclusion follows from a consideration of energy.

Returning to the previous calculations, we see that nothing goes awry if we replace the canonical quantization rule with a sign change to an *anticommutator* in the case of fermions. Defining the anticommutator

$$\left\{\hat{A},\hat{B}\right\} \equiv \hat{A}\hat{B}+\hat{B}\hat{A}$$

the fundamental anticommutation relations for the Dirac field are:

$$\begin{aligned}\left\{\hat{\psi}^A(\mathbf{x},t),\hat{\psi}^B(\mathbf{x}',t)\right\} &= 0\\ \left\{\hat{\pi}_A(\mathbf{x},t),\hat{\psi}^B(\mathbf{x}',t)\right\} &= i\delta_A^B\delta^3(\mathbf{x}-\mathbf{x}')\\ \left\{\hat{\pi}_A(\mathbf{x},t),\hat{\pi}_B(\mathbf{x}',t)\right\} &= 0\end{aligned} \tag{6.41}$$

with the consequence

$$\left\{\hat{b}_a^\dagger(\mathbf{k}),\hat{b}_b(\mathbf{k}')\right\} = \delta_{ab}\delta^3(\mathbf{k}-\mathbf{k}') \tag{6.42}$$

$$\left\{\hat{d}_a^\dagger(\mathbf{k}),\hat{d}_b(\mathbf{k}')\right\} = \delta_{ab}\delta^3(\mathbf{k}-\mathbf{k}') \tag{6.43}$$

All other *anti*commutators vanish. This implies the exclusion principal, since the vanishing of

$$\left\{\hat{\psi}(\mathbf{x},t),\hat{\psi}(\mathbf{x}',t)\right\} = \hat{\psi}(\mathbf{x},t)\hat{\psi}(\mathbf{x}',t)+\hat{\psi}(\mathbf{x}',t)\hat{\psi}(\mathbf{x},t)$$

implies

$$\hat{\psi}(\mathbf{x},t)\hat{\psi}(\mathbf{x}',t) = -\hat{\psi}(\mathbf{x}',t)\hat{\psi}(\mathbf{x},t)$$

for any two spinor fields, and therefore for two identical spinors at the same point,

$$\hat{\psi}(\mathbf{x},t)\hat{\psi}(\mathbf{x},t) = -\hat{\psi}(\mathbf{x},t)\hat{\psi}(\mathbf{x},t) = 0$$

Identical fermions cannot exist in the same state.

6.3.1.3 The Dirac Hamiltonian

Next, consider the Hamiltonian. We wish to express it as a quantum operator in terms of the creation and annihilation operators. It is now convenient to use the simplified form of the Dirac Hamiltonian, eq.(6.30):

$$H = i \int d^3x \; \pi\gamma^0 \left(i\gamma^i \partial_i - m \right) \psi = \int d^3x \; \pi \partial_0 \psi$$

which assumes the field equation is satisfied.

We begin by substituting the field operator expansions,

$$\begin{aligned} \hat{\psi}(\mathbf{x}, t) &= \frac{1}{(2\pi)^{3/2}} \sum_{a=1}^{2} \int d^3k \sqrt{\frac{m}{\omega}} \left(\hat{b}_a(\mathbf{k}) u_a(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} + \hat{d}_a^\dagger(\mathbf{k})) v_a(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \\ \hat{\pi}(\mathbf{x}, t) = i\psi^\dagger(\mathbf{x}, t) h\gamma^0 &= \frac{i}{(2\pi)^{3/2}} \sum_{a=1}^{2} \int d^3k \sqrt{\frac{m}{\omega}} \left(\hat{b}_a^\dagger(\mathbf{k}) \bar{u}_a(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} + \hat{d}_a(\mathbf{k}) \bar{v}_a(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \gamma^0 \end{aligned}$$

into the integral for the Hamiltonian,

$$\begin{aligned} \hat{H} &= \int d^3x \; : \hat{\pi} \partial_0 \hat{\psi} : \\ &= \frac{i}{(2\pi)^3} \sum_{a=1}^{2} \sum_{b=1}^{2} \int d^3x \int d^3k \int d^3k' \; \frac{m}{\sqrt{\omega\omega'}} : \left(\hat{d}_a(\mathbf{k}) v_a^\dagger(\mathbf{k}) h\gamma^0 e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} + \hat{b}_a^\dagger(\mathbf{k}) u_a^\dagger(\mathbf{k}) h\gamma^0 e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \\ &\quad \times \left(-i\omega' \hat{b}_b(\mathbf{k}') u_b(\mathbf{k}') e^{-i\left(\omega' t - \mathbf{k}'\cdot\mathbf{x}\right)} + i\omega' \hat{d}_b^\dagger(\mathbf{k}') v_b(\mathbf{k}') e^{i\left(\omega' t - \mathbf{k}'\cdot\mathbf{x}\right)} \right) : \end{aligned}$$

Collecting terms we have

$$\begin{aligned} \hat{H} &= \frac{i}{(2\pi)^3} \sum_{a=1}^{2} \sum_{b=1}^{2} \int d^3x \int d^3k \int d^3k' \; \frac{i\omega' m}{\sqrt{\omega\omega'}} \\ &\quad \times : \left(-\hat{d}_a(\mathbf{k}) \hat{b}_b(\mathbf{k}') v_a^\dagger(\mathbf{k}) h\gamma^0 u_b(\mathbf{k}') e^{-i\left(\omega+\omega'\right)t + i\left(\mathbf{k}+\mathbf{k}'\right)\cdot\mathbf{x}} + \hat{d}_a(\mathbf{k}) \hat{d}_b^\dagger(\mathbf{k}') v_a^\dagger(\mathbf{k}) h\gamma^0 v_b(\mathbf{k}') e^{-i\left(\omega-\omega'\right)t + i\left(\mathbf{k}-\mathbf{k}'\right)\cdot\mathbf{x}} \right. \\ &\quad \left. + -\hat{b}_a^\dagger(\mathbf{k}) \hat{b}_b(\mathbf{k}') u_a^\dagger(\mathbf{k}) h\gamma^0 u_b(\mathbf{k}') e^{i\left(\omega-\omega'\right)t - i\left(\mathbf{k}-\mathbf{k}'\right)\cdot\mathbf{x}} + \hat{b}_a^\dagger(\mathbf{k}) \hat{d}_b^\dagger(\mathbf{k}') u_a^\dagger(\mathbf{k}) h\gamma^0 v_b(\mathbf{k}') e^{i\left(\omega+\omega'\right)t - i\left(\mathbf{k}+\mathbf{k}'\right)\cdot\mathbf{x}} \right) : \end{aligned}$$

Now, integrating over d^3x, we produce Dirac delta functions:

$$\begin{aligned} \hat{H} &= -\sum_{a=1}^{2} \sum_{b=1}^{2} \int d^3k \int d^3k' \; \frac{\omega' m}{\sqrt{\omega\omega'}} \\ &\quad \times : \left(-\hat{d}_a(\mathbf{k}) \hat{b}_b(\mathbf{k}') v_a^\dagger(\mathbf{k}) h\gamma^0 u_b(\mathbf{k}') \delta^3(\mathbf{k}+\mathbf{k}') e^{-2i\omega t} + \hat{d}_a(\mathbf{k}) \hat{d}_b^\dagger(\mathbf{k}') v_a^\dagger(\mathbf{k}) h\gamma^0 v_b(\mathbf{k}') \delta^3(\mathbf{k}-\mathbf{k}') \right. \\ &\quad \left. + -\hat{b}_a^\dagger(\mathbf{k}) \hat{b}_b(\mathbf{k}') u_a^\dagger(\mathbf{k}) h\gamma^0 u_b(\mathbf{k}') \delta^3(\mathbf{k}-\mathbf{k}') + \hat{b}_a^\dagger(\mathbf{k}) \hat{d}_b^\dagger(\mathbf{k}') u_a^\dagger(\mathbf{k}) h\gamma^0 v_b(\mathbf{k}') \delta^3(\mathbf{k}+\mathbf{k}') e^{2i\omega t} \right) : \end{aligned}$$

which immediately integrate to give

$$\begin{aligned} \hat{H} &= -m \sum_{a=1}^{2} \sum_{b=1}^{2} \int d^3k : \left(-\hat{d}_a(\mathbf{k}) \hat{b}_b(-\mathbf{k}) v_a^\dagger(\mathbf{k}) h\gamma^0 u_b(-\mathbf{k}) e^{-2i\omega t} + \hat{d}_a(\mathbf{k}) \hat{d}_b^\dagger(\mathbf{k}) v_a^\dagger(\mathbf{k}) h\gamma^0 v_b(\mathbf{k}) \right. \\ &\quad \left. + -\hat{b}_a^\dagger(\mathbf{k}) \hat{b}_b(\mathbf{k}) u_a^\dagger(\mathbf{k}) h\gamma^0 u_b(\mathbf{k}) + \hat{b}_a^\dagger(\mathbf{k}) \hat{d}_b^\dagger(-\mathbf{k}) u_a^\dagger(\mathbf{k}) h\gamma^0 v_b(-\mathbf{k}) e^{2i\omega t} \right) : \end{aligned}$$

Using Eqs.(6.35),

$$\begin{aligned} v_a(-\mathbf{k}) &= -\gamma^0 v_a(\mathbf{k}) \\ u_a(-\mathbf{k}) &= \gamma^0 u_a(\mathbf{k}) \end{aligned}$$

we have

$$\begin{aligned}
\hat{H} &= -m\sum_{a=1}^{2}\sum_{b=1}^{2}\int d^3k : \Big(-\hat{d}_a(\mathbf{k})\,\hat{b}_b(-\mathbf{k})\,v_a^\dagger(\mathbf{k})\,hu_b(\mathbf{k})\,e^{-2i\omega t} + \hat{d}_a(\mathbf{k})\,\hat{d}_b^\dagger(\mathbf{k})\,v_a^\dagger(\mathbf{k})\,h\gamma^0 v_b(\mathbf{k}) \\
&\qquad + -\hat{b}_a^\dagger(\mathbf{k})\,\hat{b}_b(\mathbf{k})\,u_a^\dagger(\mathbf{k})\,h\gamma^0 u_b(\mathbf{k}) - \hat{b}_a^\dagger(\mathbf{k})\,\hat{d}_b^\dagger(-\mathbf{k})\,u_a^\dagger(\mathbf{k})\,hv_b(\mathbf{k})e^{2i\omega t}\Big) : \\
&= -m\sum_{a=1}^{2}\sum_{b=1}^{2}\int d^3k : \Big(-\hat{d}_a(\mathbf{k})\,\hat{b}_b(-\mathbf{k})\,(\bar{v}_a(\mathbf{k})\,u_b(\mathbf{k}))\,e^{-2i\omega t} + \hat{d}_a(\mathbf{k})\,\hat{d}_b^\dagger(\mathbf{k})\,\big(\bar{v}_a(\mathbf{k})\,\gamma^0 v_b(\mathbf{k})\big) \\
&\qquad + -\hat{b}_a^\dagger(\mathbf{k})\,\hat{b}_b(\mathbf{k})\,\big(\bar{u}_a(\mathbf{k})\,\gamma^0 u_b(\mathbf{k})\big) - \hat{b}_a^\dagger(\mathbf{k})\,\hat{d}_b^\dagger(-\mathbf{k})\,\bar{u}_a(\mathbf{k})\,v_b(\mathbf{k})e^{2i\omega t}\Big) :
\end{aligned}$$

Finally, with the orthonormality we set $\bar{v}_a(\mathbf{k})\,u_b(\mathbf{k}) = 0 = \bar{u}_a(\mathbf{k})\,v_b(\mathbf{k})$ and use Eqs.(6.36) to write

$$\hat{H} = -m\sum_{a=1}^{2}\sum_{b=1}^{2}\int d^3k : \Big(\hat{d}_a(\mathbf{k})\,\hat{d}_b^\dagger(\mathbf{k})\,\frac{\omega}{m}\delta_{ab} - \hat{b}_a^\dagger(\mathbf{k})\,\hat{b}_b(\mathbf{k})\,\frac{\omega}{m}\delta_{ab}\Big) :$$

and finally

$$\hat{H} = \sum_{a=1}^{2}\int d^3k\,\omega : \Big(\hat{b}_a^\dagger(\mathbf{k})\,\hat{b}_a(\mathbf{k}) - \hat{d}_a(\mathbf{k})\,\hat{d}_a^\dagger(\mathbf{k})\Big) :$$

This would be a troubling result if it weren't for the anticommutation relations. If we simply used the normal ordering procedure, the second term would be negative and the energy indefinite. However, using the anticommutator, Eq.(6.43), the normal ordering prescription for fermions is taken to mean

$$: \hat{d}_b(\mathbf{k}')\hat{d}_a^\dagger(\mathbf{k}) : \;=\; -\hat{d}_a^\dagger(\mathbf{k})\hat{d}_b(\mathbf{k}') \tag{6.44}$$

We therefore write the normal ordered Hamiltonian operator as

$$\hat{H} = \sum_{a=1}^{2}\int d^3k\,\omega\Big(\hat{b}_a^\dagger(\mathbf{k})\,\hat{b}_a(\mathbf{k}) + \hat{d}_a^\dagger(\mathbf{k})\,\hat{d}_a(\mathbf{k})\Big) \tag{6.45}$$

This convention preserves the anticommutativity, while still eliminating the infinite delta function contribution to the vacuum energy.

6.3.2 Symmetries of the Dirac field

We'd now like to find the conserved currents of the Dirac field. There are two kinds – the spacetime symmetries, including Lorentz transformations and translations, and a $U(1)$ phase symmetry. We'll discuss the spacetime symmetries first. We put off our study of the phase symmetry to the next chapter, where it leads us systematically to Quantum Electrodynamics: *QED*.

6.3.2.1 Translations

Under a translation, $x^\alpha \to x^\alpha + a^\alpha$, the Dirac field changes by

$$\psi(x^\alpha) \to \psi(x^\alpha + a^\alpha) = \psi(x^\alpha) + \frac{\partial\psi(x^\alpha)}{\partial x^\beta}a^\beta$$

so we identify Δ of eq.(1.35) as

$$\Delta = (\partial_\beta\psi)\,a^\beta$$

The four conserved currents form the energy-momentum tensor, given by eq.(1.37):

$$\begin{aligned} T^{\alpha\beta} &= \frac{\delta\mathcal{L}}{\delta\left(\partial_\alpha\psi\right)}\partial^\beta\psi - \mathcal{L}\eta^{\mu\beta} \\ &= i\bar{\psi}\gamma^\alpha\partial^\beta\psi - \eta^{\alpha\beta}\bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi \\ &= i\bar{\psi}\gamma^\alpha\partial^\beta\psi \end{aligned}$$

since the Lagrangian density vanishes when the field equation is satisfied. This form is called the canonical energy-momentum tensor, but it is not symmetric. Instead, write the action in the symmetric form

$$\begin{aligned} S &= \int \overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi \\ &= \frac{1}{2}\int\left(\overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi - \overline{\psi}\left(i\overleftarrow{\partial\!\!\!/} + m\right)\psi\right) \end{aligned}$$

then the energy-momentum takes the Belinfante form

$$\begin{aligned} T^{\alpha\beta} &= \frac{\delta\mathcal{L}}{\delta\left(\partial_\alpha\psi\right)}\partial^\beta\psi + \partial^\beta\overline{\psi}\frac{\delta\mathcal{L}}{\delta\left(\partial_\alpha\overline{\psi}\right)} - \mathcal{L}\eta^{\mu\beta} \\ &= \frac{1}{2}\left(\overline{\psi}i\gamma^\alpha\right)\partial^\beta\psi + \frac{1}{2}\partial^\beta\overline{\psi}\left(-i\gamma^\alpha\psi\right) - \frac{1}{2}\left(\overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi - \overline{\psi}\left(i\overleftarrow{\partial\!\!\!/} + m\right)\psi\right)\eta^{\mu\beta} \\ &= \frac{i}{2}\left(\overline{\psi}\gamma^\alpha\partial^\beta\psi - \partial^\beta\overline{\psi}\gamma^\alpha\psi\right) \\ &= \frac{i}{2}\left(\overline{\psi}\gamma^\alpha\partial^\beta\psi - \partial^\beta\overline{\psi}\gamma^\alpha\psi\right) \end{aligned}$$

We check that this is divergence free, $\partial\!\!\!/\psi = -im\psi$

$$\begin{aligned} \partial_\alpha T^{\alpha\beta} &= \frac{i}{2}\left(\partial_\alpha\overline{\psi}\gamma^\alpha\partial^\beta\psi - \partial_\alpha\partial^\beta\overline{\psi}\gamma^\alpha\psi + \overline{\psi}\gamma^\alpha\partial^\beta\partial_\alpha\psi - \partial^\beta\overline{\psi}\gamma^\alpha\partial_\alpha\psi\right) \\ &= \frac{i}{2}\left(\overline{\psi}\overleftarrow{\partial\!\!\!/}\partial^\beta\psi - \partial^\beta\overline{\psi}\overleftarrow{\partial\!\!\!/}\psi + \overline{\psi}\partial^\beta\partial\!\!\!/\psi - \partial^\beta\overline{\psi}\partial\!\!\!/\psi\right) \\ &= \frac{i}{2}\left(im\overline{\psi}\partial^\beta\psi - im\partial^\beta\overline{\psi}\psi - im\overline{\psi}\partial^\beta\psi + im\partial^\beta\overline{\psi}\psi\right) \\ &= -\frac{m}{2}\left(\overline{\psi}\partial^\beta\psi - \partial^\beta\overline{\psi}\psi - \overline{\psi}\partial^\beta\psi + \partial^\beta\overline{\psi}\psi\right) \\ &= 0 \end{aligned}$$

For the conserved charges, we therefore find that the conserved energy is the Hamiltonian,

$$\hat{P}^0 = i\int d^3x\bar{\psi}\gamma^0\partial^0\psi = \hat{H}$$

while the conserved momentum is

$$\begin{aligned} \hat{P}^i &= -i:\int d^3x\bar{\psi}\gamma^0\partial_i\psi: \\ &= \sum_{a=1}^{2}\sum_{b=1}^{2}\int d^3k\frac{mk'^i}{\omega}\left(\hat{b}_a^\dagger(\mathbf{k})\hat{b}_b(\mathbf{k})u_a^\dagger(\mathbf{k})h\gamma^0u_b(\mathbf{k}) - \hat{b}_a^\dagger(\mathbf{k})\hat{d}_b^\dagger(-\mathbf{k})u_a^\dagger(\mathbf{k})h\gamma^0e^{2i\omega t}v_b(-\mathbf{k})\right. \\ &\qquad\left. +\hat{d}_a(\mathbf{k})\hat{b}_b(-\mathbf{k})v_a^\dagger(\mathbf{k})h\gamma^0e^{-2i\omega t}u_b(-\mathbf{k}) - :\hat{d}_a(\mathbf{k})\hat{d}_b^\dagger(\mathbf{k}):v_a^\dagger(\mathbf{k})h\gamma^0v_b(\mathbf{k})\right) \\ &= \sum_{a=1}^{2}\int d^3k\ k'^i\left(\hat{b}_a^\dagger(\mathbf{k})\hat{b}_a(\mathbf{k}) - :\hat{d}_a(\mathbf{k})\hat{d}_a^\dagger(\mathbf{k}):\right) \\ &= \sum_{a=1}^{2}\int d^3k\ k'^i\left(\hat{b}_a^\dagger(\mathbf{k})\hat{b}_a(\mathbf{k}) + \hat{d}_a^\dagger(\mathbf{k})\hat{d}_a(\mathbf{k})\right) \end{aligned}$$

so that

$$\hat{\mathbf{P}} = \sum_{a=1}^{2} \int d^3k \; \mathbf{k} \left(\hat{b}_a^\dagger(\mathbf{k}) \hat{b}_a(\mathbf{k}) + \hat{d}_a^\dagger(\mathbf{k}) \hat{d}_a(\mathbf{k}) \right)$$

This is just what we expect.

6.4 Gauging the Dirac action

Starting with the action for the Dirac equation,

$$S = \int d^4x \; \bar{\psi} \left(i\gamma^\mu \partial_\mu - m \right) \psi$$

we note that in addition to the Poincaré symmetry, it has a global phase symmetry. That is, if we replace

$$\begin{aligned} \psi &\rightarrow \psi e^{i\eta} \\ \bar{\psi} &\rightarrow \bar{\psi} e^{-i\eta} \end{aligned} \tag{6.46}$$

for any constant phase η, the action S remains unchanged. This leads immediately to a conserved current. For an infinitesimal phase change,

$$\begin{aligned} \Delta &= \psi (1 + i\eta) - \psi = i\eta\psi \\ \bar{\Delta} &= \bar{\psi} (1 + i\eta) - \bar{\psi} = -i\eta\bar{\psi} \end{aligned}$$

so computing $J^\alpha \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \psi)} (i\eta\psi)$, and writing the result with operators, we find the Dirac $U(1)$ current,

$$\hat{J}^\alpha = -\eta \hat{\bar{\psi}} \gamma^\alpha \hat{\psi} \tag{6.47}$$

In terms of creation and annihilation operators the conserved charge operator is thererfore

$$\begin{aligned} \hat{Q} &= \int \hat{J}^0 d^3x \\ &= -\eta \int d^3x \; : \hat{\bar{\psi}} \gamma^0 \hat{\psi} : \\ &= -\eta \sum_{a=1}^{2} \sum_{b=1}^{2} \int d^3k \frac{m}{\omega} \Big(: \hat{b}_a^\dagger(\mathbf{k}) u_a^\dagger(\mathbf{k}) h \gamma^0 \hat{b}_b(\mathbf{k}) u_b(\mathbf{k}) + \hat{b}_a^\dagger(\mathbf{k}) u_a^\dagger(\mathbf{k}) e^{2i\omega t} h \gamma^0 \hat{d}_b^\dagger(-\mathbf{k}) v_b(-\mathbf{k}) \\ &\qquad + \hat{d}_a(\mathbf{k}) v_a^\dagger(\mathbf{k}) e^{-2i\omega t} h \gamma^0 \hat{b}_b(-\mathbf{k}) u_b(-\mathbf{k}) + \hat{d}_a(\mathbf{k}) v_a^\dagger(\mathbf{k}) h \gamma^0 \hat{d}_b^\dagger(\mathbf{k}) v_b(\mathbf{k}) : \Big) \\ &= -\eta \sum_{a=1}^{2} \int d^3k \left(\hat{b}_a^\dagger(\mathbf{k}) \hat{b}_a(\mathbf{k}) + : \hat{d}_a(\mathbf{k}) \hat{d}_a^\dagger(\mathbf{k}) : \right) \end{aligned}$$

Applying the fermionic normal ordering, the charge operator is therefore

$$\hat{Q} = -\eta \sum_{a=1}^{2} \int d^3k \left(\hat{b}_a^\dagger(\mathbf{k}) \hat{b}_a(\mathbf{k}) - \hat{d}_a^\dagger(\mathbf{k}) \hat{d}_a(\mathbf{k}) \right) \tag{6.48}$$

Therefore, the total conserved charge Q is proportional to the difference between the number of particles and the number of antiparticles. The most straightforward interpretation of this conservation law is a conservation of *electrical* charge and (with some slight modification for the electroweak theory) this interpretation is correct.

6.4.1 The covariant derivative

We could now take the electromagnetic current of the spinor field, $J^\alpha = -\eta\bar{\psi}\gamma^\alpha\psi$ as the source for the Maxwell field by including $J^\alpha A_\alpha$ with the Maxwell action. However, gauging provides a more systematic way to come up with the same action in a way that immediately generalizes to other types of interactions. The procedure is as follows.

Suppose we try to write a revised version of the Dirac action which is invariant under *local* phase transformations,

$$\begin{aligned} \psi &\rightarrow \psi' = \psi e^{i\varphi(t,\mathbf{x})} \\ \bar{\psi} &\rightarrow \bar{\psi}' = \bar{\psi} e^{-i\varphi(t,\mathbf{x})} \end{aligned} \tag{6.49}$$

Clearly, this must be a different action because of the derivative acting on the phase. If we substitute these expressions into the Dirac action we find

$$\begin{aligned} S' &= \int d^4x\ \bar{\psi}e^{-i\varphi(t,\mathbf{x})}\left(i\gamma^\mu\partial_\mu - m\right)\psi e^{i\varphi(t,\mathbf{x})} \\ &= S - \bar{\psi}\gamma^\mu\psi\left(\partial_\mu\varphi\left(t,\mathbf{x}\right)\right) \end{aligned}$$

In order to build a new action which is invariant, we somehow need to cancel this extra term.

The key to a solution is that an extra, undesired piece occurs whenever we take a derivative. We can fix the problem by introducing a different kind of derivative, D_α, called a *covariant* derivative. For local phase symmetry, we say that the derivative must be made covariant with respect to phase transformations. All this means is that it should commute with the phase change, in the sense that

$$D'_\alpha\psi' = e^{i\varphi(t,\mathbf{x})}\left(D_\alpha\psi\right) \tag{6.50}$$

We just demand that the derivative $D_\alpha\psi$ should transform in the same way as ψ itself. If we can find such a covariant derivative, then we may set

$$S_{local} = \int d^4x\ \bar{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi \tag{6.51}$$

This is the action we need because then

$$\begin{aligned} S'_{local} &= \int d^4x\ \left(i\bar{\psi}'\gamma^\mu D'_\mu\psi' - m\bar{\psi}'\psi'\right) \\ &= \int d^4x\ \left(i\bar{\psi}e^{-i\varphi(t,\mathbf{x})}\gamma^\mu e^{i\varphi(t,\mathbf{x})}D_\mu\psi - m\bar{\psi}\psi\right) \\ &= S_{local} \end{aligned}$$

The only trick is to find a suitable generalization of the derivative.

6.4.1.1 Derivations

To generalize the derivative, we need to know the properties that make an operator a derivation. We define

Define: A *derivation* is an operator D which is linear and Leibnitz:

1. Linear: $D\left(\alpha f + \beta g\right) = \alpha Df + \beta Dg$
2. Leibnitz: $D\left(fg\right) = \left(Df\right)g + f\left(Dg\right)$

Notice that these two conditions together require D to vanish when acting on constants. The Leibnitz property gives $D\left(\alpha f\right) = \left(D\alpha\right)f + \alpha Df$, while linearity requires $D\left(\alpha f\right) = \alpha Df$. These are consistent only if $\left(D\alpha\right)f = 0$ for any function f. Choosing $f(x) = 1$ gives the result.

Next, consider how two derivations may differ. If D_1 is a derivation and we define

$$D_2 = D_1 + F(x)$$

then D_2 is also linear

$$\begin{aligned} D_2(\alpha f + \beta g) &= (D_1 + F)(\alpha f + \beta g) \\ &= D_1(\alpha f + \beta g) + F(\alpha f + \beta g) \\ &= \alpha D_1 f + \beta D_1 g + \alpha F f + \beta F g \\ &= \alpha (D_1 + F) f + \beta (D_1 + F) g \\ &= \alpha (D_2 f) + \beta (D_2 g) \end{aligned}$$

but the Leibnitz property fails:

$$\begin{aligned} D_2(fg) &= D_1(fg) + Ffg \\ &= (D_1 f) g + f (D_1 g) + Ffg \\ &\neq (D_2 f) g + f (D_2 g) \end{aligned}$$

We can fix this problem by introducting additive *weights* for functions. If f_n has weight n, and g_m has weight m, then we require the product, $h_{m+n} = f_n g_m$ to have weight $m + n$. Now we can define

$$D_2 g_m = D_1 g_m + mF g_m$$

and the Leibnitz rule is satisfied:

$$\begin{aligned} D_2(f_n g_m) &= D_1(f_n g_m) + (n+m) F f_n g_m \\ &= (D_1 f_n) g_m + f_n (D_1 g_m) + (n+m) F f_n g_m \\ &= (D_2 f_n) g_m + f_n (D_2 g_m) \end{aligned}$$

The use of weights is consistent with phase transformations, because if we have a product of two spinors and each changes by a phase, we get a doubled phase factor:

$$\chi\psi \to \left(\chi e^{i\varphi}\right)\left(\psi e^{i\varphi}\right) = \chi\psi e^{2i\varphi}$$

Thus, each spinor would be assigned a weight of one.

The additive term in a covariant derivative is called a *connection*.

6.4.1.2 The $U(1)$-covariant derivative

We are now in a position to find a suitable covariant derivative. Since the partial derivative, ∂_α, is a derivation, we may add a single vector field to form another derivation,

$$D_\alpha = \partial_\alpha - i\eta A_\alpha$$

where the factor of $-i$ is simply a convenient convention. This definition is sufficient. The condition we require, Eq.(6.50), becomes

$$D'_\alpha \psi' = e^{i\varphi(t,\mathbf{x})}(D_\alpha \psi)$$

where not only $\psi' = e^{i\varphi(t,\mathbf{x})}\psi$ but also,

$$D'_\alpha = \partial_\alpha - iA'_\alpha$$

Combining these expressions,

$$\begin{aligned} D'_\alpha \psi' &= (\partial_\alpha - iA'_\alpha)\left(e^{i\varphi}\psi\right) \\ &= i\left(\partial_\alpha \varphi\right)\psi + e^{i\varphi}\partial_\alpha \psi - iA'_\alpha e^{i\varphi(t,\mathbf{x})}\psi \end{aligned}$$

This must reduce to a phase times the original covariant derivative,

$$e^{i\varphi}\left(D_\alpha \psi\right) = e^{i\varphi}\partial_\alpha \psi - ie^{i\varphi}A_\alpha \psi$$

so equating,

$$\begin{aligned} ie^{i\varphi}\left(\partial_\alpha \varphi\right)\psi + e^{i\varphi}\partial_\alpha \psi - iA'_\alpha e^{i\varphi}\psi &= e^{i\varphi}\partial_\alpha \psi - ie^{i\varphi}A_\alpha \psi \\ ie^{i\varphi}\left(\partial_\alpha \varphi\right)\psi - iA'_\alpha e^{i\varphi}\psi &= -ie^{i\varphi}A_\alpha \psi \end{aligned}$$

Since this must hold for all ψ we see that A_α must change according to

$$A'_\alpha = A_\alpha + \partial_\alpha \varphi \tag{6.52}$$

Other than this necessary transformation property, A_α is an arbitrary vector field. The new action, S_{local} is now invariant under the combined transformations,

$$\begin{aligned} \psi \to \psi' &= \psi e^{i\varphi(t,\mathbf{x})} \\ \bar{\psi} \to \bar{\psi}' &= \bar{\psi} e^{-i\varphi(t,\mathbf{x})} \\ A_\alpha \to A'_\alpha &= A_\alpha + \partial_\alpha \varphi\left(t,\mathbf{x}\right) \end{aligned} \tag{6.53}$$

6.4.2 Gauging

Given the preceeding construction, the action

$$S_{local} = \int d^4x \; \bar{\psi}\left(i\gamma^\alpha D_\alpha - m\right)\psi \tag{6.54}$$

where $D_\alpha = \partial_\alpha - iA_\alpha$, is invariant under local phase transformations.

Exercise: Demonstrate the invariance of S_{local} under the simultaneous transformations of Eqs.(6.53) by explicit substitution.

This procedure, of making a global symmetry into a local symmetry by introducing a covariant derivative, is called *gauging* the symmetry.

We are left with a question: What is the new field A_α? As it stands, it does not matter much because A_α has no interesting physical properties. Since no derivatives of A_α appear in S_{local}, A_α cannot propagate. In fact, we can't even vary S_{local} with respect to A_α because it forces the current to vanish:

$$\frac{\delta S_{local}}{\delta A_\alpha} = \bar{\psi}\gamma^\alpha \psi$$

We can fix this by adding a term built from derivatives of the connection A_α, but because A_α obeys an inhomogeneous transformation property we need some way to tell what parts of A_α are physical. For example, we could not just write

$$\Box A_\alpha = 0$$

as a field equation for A_α because under a phase transformation the simple wave equation changes to

$$\Box A_\alpha + \Box\left(\partial_\alpha \varphi\right) = 0$$

Which equation would we solve?

Fortunately, there is a standard way to find physical fields associated with a connection. It depends on the fact that, unlike partial derivatives, covariant derivatives do not commute. For the $U(1)$ case, the commutator of two covariant derivatives on an arbitrary spinor gives

$$\begin{aligned}
[D_\alpha, D_\beta]\psi &= D_\alpha(\partial_\beta\psi - iA_\beta\psi) - D_\beta(\partial_\alpha\psi - iA_\alpha\psi) \\
&= \partial_\alpha\partial_\beta\psi - i(\partial_\alpha A_\beta)\psi - iA_\beta(\partial_\alpha\psi) - iA_\alpha(\partial_\beta\psi - iA_\beta\psi) \\
&\quad -\partial_\beta\partial_\alpha\psi + i(\partial_\beta A_\alpha)\psi + iA_\alpha(\partial_\beta\psi) + iA_\beta(\partial_\alpha\psi - iA_\alpha) \\
&= -i(\partial_\alpha A_\beta - \partial_\beta A_\alpha + i[A_\alpha, A_\beta])\psi
\end{aligned}$$

Several important things have happened here. First, because the covariant expression on the right, $[D_\alpha, D_\beta]\psi$ is $U(1)$-covariant, it transforms the same way as ψ, that is,

$$\left[D'_\alpha, D'_\beta\right]\psi' = e^{i\varphi(\mathbf{x},t)}[D_\alpha, D_\beta]\psi$$

The left side must transform in the same way. Since ψ on the right is convariant changes to $e^{i\varphi(\mathbf{x},t)}\psi$ (that is, ψ transforms linearly and homogeneously under a $U(1)$ transformation) the remaining factor

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha + i[A_\alpha, A_\beta] \tag{6.55}$$

must be invariant when we transform the fields according to Eq.(6.53). For more general - non-Abelian - group symmetries, the connection is more involved and the commutator term does not vanish, but in the present case the $U(1)$-*field strength* is simply

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$$

and we immediately check that

$$\begin{aligned}
F'_{\alpha\beta} &= \partial_\alpha A'_\beta - \partial_\beta A'_\alpha \\
&= \partial_\alpha(A_\beta + \partial_\beta\varphi) - \partial_\beta(A_\alpha + \partial_\alpha\varphi) \\
&= \partial_\alpha A_\beta - \partial_\beta A_\alpha \\
&= F_{\alpha\beta}
\end{aligned}$$

as expected.

This is also a characteristic property – the curvature of a connection is always a *tensor* : it transforms linearly and homogeneously under a gauge transformation.

A relationship of this same form may be familiar from the Ricci identity of general relativity, where the more complicated general-coordinate-covariant derivative,

$$D_\alpha v^\beta = \partial_\alpha v^\beta + v^\mu \Gamma^\beta_{\ \mu\alpha}$$

yields the Riemann curvature tensor as the field strength of the Christoffel connection,

$$\begin{aligned}
[D_\alpha, D_\beta]v^\mu &= D_\alpha\left(\partial_\beta v^\mu + v^\nu\Gamma^\mu_{\ \nu\beta}\right) - D_\beta(\partial_\alpha v^\mu + v^\nu\Gamma^\mu_{\ \nu\alpha}) \\
&= \partial_\alpha\left(\partial_\beta v^\mu + v^\nu\Gamma^\mu_{\ \nu\beta}\right) + \left(\partial_\beta v^\rho + v^\nu\Gamma^\rho_{\ \nu\beta}\right)\Gamma^\mu_{\ \rho\alpha} - \left(\partial_\rho v^\mu + v^\nu\Gamma^\mu_{\ \nu\rho}\right)\Gamma^\rho_{\ \beta\alpha} \\
&\quad -\partial_\beta(\partial_\alpha v^\mu + v^\nu\Gamma^\mu_{\ \nu\alpha}) - (\partial_\alpha v^\rho + v^\nu\Gamma^\rho_{\ \nu\alpha})\Gamma^\mu_{\ \rho\beta} + \left(\partial_\rho v^\mu + v^\nu\Gamma^\mu_{\ \nu\rho}\right)\Gamma^\rho_{\ \beta\alpha} \\
&= \partial_\alpha\partial_\beta v^\mu + \partial_\alpha v^\nu\Gamma^\mu_{\ \nu\beta} + v^\nu\partial_\alpha\Gamma^\mu_{\ \nu\beta} + \partial_\beta v^\rho\Gamma^\mu_{\ \rho\alpha} + v^\nu\Gamma^\rho_{\ \nu\beta}\Gamma^\mu_{\ \rho\alpha} - \left(\partial_\rho v^\mu + v^\nu\Gamma^\mu_{\ \nu\rho}\right)\Gamma^\rho_{\ \beta\alpha} \\
&\quad -\partial_\beta\partial_\alpha v^\mu - \partial_\beta v^\nu\Gamma^\mu_{\ \nu\alpha} - v^\nu\partial_\beta\Gamma^\mu_{\ \nu\alpha} - \partial_\alpha v^\rho\Gamma^\mu_{\ \rho\beta} - v^\nu\Gamma^\rho_{\ \nu\alpha}\Gamma^\mu_{\ \rho\beta} + \left(\partial_\rho v^\mu + v^\nu\Gamma^\mu_{\ \nu\rho}\right)\Gamma^\rho_{\ \beta\alpha}
\end{aligned}$$

Cancelling like terms and using the symmetry of mixed partials and the symmetry of the Christoffel connection, $\Gamma^\rho_{\ \beta\alpha} = \Gamma^\rho_{\ \alpha\beta}$, we have the Riemann curvature tensor as the invariant

$$[D_\alpha, D_\beta]v^\mu = v^\nu\left(\partial_\alpha\Gamma^\mu_{\ \nu\beta} + \Gamma^\rho_{\ \nu\beta}\Gamma^\mu_{\ \rho\alpha} - \partial_\beta\Gamma^\mu_{\ \nu\alpha} - \Gamma^\rho_{\ \nu\alpha}\Gamma^\mu_{\ \rho\beta}\right) \equiv v^\nu R^\mu_{\ \nu\alpha\beta}$$

The computation is similar for arbitrary Lie groups.

6.4.2.1 Gauged action

We can use the curvature to write an action for A_α. Any Lorentz invariant quantity built purely from $F_{\alpha\beta}$ is a possible term in the action. There are two possible terms:

$$F_{\alpha\beta}F^{\alpha\beta}$$
$$\varepsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}$$

However, the second of these does not contribute to the field equations for A_α because it is a total divergence:

$$\begin{aligned}
\varepsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu} &= \varepsilon^{\alpha\beta\mu\nu}\left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right)F_{\mu\nu} \\
&= 2\varepsilon^{\alpha\beta\mu\nu}\partial_\alpha A_\beta F_{\mu\nu} \\
&= \partial_\alpha\left(2\varepsilon^{\alpha\beta\mu\nu}A_\beta F_{\mu\nu}\right) - 2\varepsilon^{\alpha\beta\mu\nu}A_\beta\partial_\alpha F_{\mu\nu} \\
&= \partial_\alpha\left(2\varepsilon^{\alpha\beta\mu\nu}A_\beta F_{\mu\nu}\right)
\end{aligned}$$

where the last term vanishes because

$$\varepsilon^{\alpha\beta\mu\nu}A_\beta\partial_\alpha F_{\mu\nu} = \frac{1}{3}\varepsilon^{\alpha\beta\mu\nu}A_\beta\left(\partial_\alpha F_{\mu\nu} + \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu}\right)$$

and

$$\begin{aligned}
\partial_\alpha F_{\mu\nu} + \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} &= \partial_\alpha\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) + \partial_\mu\left(\partial_\nu A_\alpha - \partial_\alpha A_\nu\right) + \partial_\nu\left(\partial_\alpha A_\mu - \partial_\mu A_\alpha\right) \\
&\equiv 0
\end{aligned}$$

Therefore, the only gauge invariant action up to second order in the field strength $F_{\alpha\beta}$ is

$$S_{local} = \int d^4x\,\left(\bar\psi\left(i\gamma^\alpha D_\alpha - m\right)\psi - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\right) \tag{6.56}$$

This is the result of $U(1)$ gauge theory, since $U(1)$ is the group of possible phase transformations. The procedure is readily generalized to other symmetry groups.

Notice that if we restore the charge η in the covariant derivative, $D_\alpha = \partial_\alpha - i\eta A_\alpha$, and expand the covariant derivative in S_{local},

$$\begin{aligned}
S_{local} &= \int d^4x\,\left(\bar\psi\left(i\gamma^\alpha\partial_\alpha - m\right)\psi + \bar\psi\gamma^\alpha\eta A_\alpha\psi - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\right) \\
&= \int d^4x\,\left(\bar\psi\left(i\gamma^\alpha\partial_\alpha - m\right)\psi - J^\alpha A_\alpha - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\right)
\end{aligned}$$

we get exactly the $J^\alpha A_\alpha$ coupling to the Dirac current $J^\alpha = -\eta\bar\psi\gamma^\alpha\psi$ that we hypothesized in the beginning, but now automatically from the gauge construction.

Chapter 7

Quantizing the Maxwell field

We have quantized spin zero and spin 1/2 fields; we now come to the most important spin 1 case, the Maxwell field. The free Maxwell theory is described in terms of the Faraday tensor $F_{\alpha\beta}$ by the action

$$S = -\frac{1}{4}\int d^4x\ F_{\alpha\beta}F^{\alpha\beta} \tag{7.1}$$

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha \tag{7.2}$$

Notice that the field A_α is necessarily massless, because mass term in the action would be of the form $m^2 A^\alpha A_\alpha$, and this term is not gauge invariant.

Before starting the Hamiltonian formulation and quantization, we carry out the classical theory to establish relations between the fields and sign conventions.

7.1 The Maxwell equations

7.1.1 The Maxwell equations in vector notation

The vector form of the homogeneous Maxwell equations is

$$\begin{aligned} \boldsymbol{\nabla}\times\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t} &= 0 \\ \boldsymbol{\nabla}\cdot\mathbf{B} &= 0 \end{aligned}$$

while the sourced equations are

$$\begin{aligned} \nabla\cdot\mathbf{E} &= \rho \\ \boldsymbol{\nabla}\times\mathbf{B}-\frac{\partial\mathbf{E}}{\partial t} &= \mathbf{J} \end{aligned}$$

These are written so that the electromagnetic fields are written on the left, the sources on the right.

The homogeneous equation for the magnetic field, $\boldsymbol{\nabla}\cdot\mathbf{B}=0$, immediately gives $\mathbf{B}$ as the curl of the vector potential,

$$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A} \tag{7.3}$$

Substituting this into the remaining homogeneous equation shows that

$$\begin{aligned} 0 &= \boldsymbol{\nabla}\times\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t} \\ &= \boldsymbol{\nabla}\times\left(\mathbf{E}+\frac{\partial\mathbf{A}}{\partial t}\right) \end{aligned}$$

so that $\mathbf{E}+\frac{\partial \mathbf{A}}{\partial t}$ must be the gradient of a scalar. Therefore, we may write

$$\mathbf{E} = -\boldsymbol{\nabla}\varphi - \frac{\partial}{\partial t}\mathbf{A} \tag{7.4}$$

7.1.2 Relations to the Faraday tensor

Now consider the Faraday tensor, related to the vector potential as in Eq.(7.2). With the 4-vector potential $A^\alpha = \left(A^0, \mathbf{A}\right)$ defined as

$$\begin{aligned} A^\alpha &= (\varphi, \mathbf{A}) \\ A_\alpha &= (\varphi, A_i) = \left(\varphi, -A^i\right) \end{aligned}$$

we may find the components of $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$.

Writing in components with $\eta_{\alpha\beta} = diag\,(1,-1,-1,-1)$ and ε_{ijk} as usual but $\varepsilon^{ijk} = -\varepsilon_{ijk}$, the magnetic field is

$$\begin{aligned} B^i &= \varepsilon_{ijk}\partial_j A^k \\ &= \varepsilon^{ij}{}_k \partial_j A^k \\ &= \varepsilon^{ijk}\partial_j A_k \\ &= \frac{1}{2}\varepsilon^{ijk}\left(\partial_j A_k - \partial_k A_j\right) \\ &= \frac{1}{2}\varepsilon^{ijk}F_{jk} \end{aligned}$$

where raising and lowering any single spatial index introduces a sign. We can invert this, giving F_{mn} as:

$$\begin{aligned} B^i &= \frac{1}{2}\varepsilon^{ijk}F_{jk} \\ \varepsilon_{imn}B^i &= \frac{1}{2}\varepsilon_{imn}\varepsilon^{ijk}F_{jk} \\ \varepsilon_{imn}B^i &= -\frac{1}{2}\left(\delta^j_m\delta^k_n - \delta^k_m\delta^j_n\right)F_{jk} \\ \varepsilon_{imn}B^i &= -F_{mn} \\ F_{mn} &= -\varepsilon_{imn}B^i \end{aligned}$$

Now, for the electric field, Eq.(7.4) becomes

$$\begin{aligned} E^i &= -\partial_i\varphi - \frac{\partial}{\partial t}A^i \\ &= -\partial_i A^0 - \partial_0 A^i \\ &= -\partial_i A_0 + \partial_0 A_i \\ &= F_{0i} \end{aligned}$$

Therefore,

$$F_{\alpha\beta} = \begin{pmatrix} 0 & E^x & E^y & E^z \\ -E^x & 0 & -B^z & B^y \\ -E^y & B^x & 0 & -B^z \\ -E^z & -B^y & B^z & 0 \end{pmatrix} \tag{7.5}$$

and

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E^x & -E^y & -E^z \\ E^x & 0 & -B^z & B^y \\ E^y & B^x & 0 & -B^z \\ E^z & -B^y & B^z & 0 \end{pmatrix} \tag{7.6}$$

7.1.3 The Maxwell equations in terms of the Faraday tensor

Now we examine the Maxwell equations.

First, notice that from the definition of the Faraday tensor, Eq.(7.2) we have

$$\begin{aligned} F_{\alpha\beta,\mu} + F_{\beta\mu,\alpha} + F_{\mu\alpha,\beta} &= \partial_\mu (\partial_\alpha A_\beta - \partial_\beta A_\alpha) + \partial_\alpha (\partial_\beta A_\mu - \partial_\mu A_\beta) + \partial_\beta (\partial_\mu A_\alpha - \partial_\alpha A_\mu) \\ & \quad (\partial_\mu \partial_\alpha A_\beta - \partial_\alpha \partial_\mu A_\beta) - (\partial_\mu \partial_\beta A_\alpha - \partial_\beta \partial_\mu A_\alpha) + (\partial_\alpha \partial_\beta A_\mu - \partial_\beta \partial_\alpha A_\mu) \\ &\equiv 0 \end{aligned}$$

Then, including a source term, $J^\beta = (\rho, J^i)$, to the Maxwell action, Eq.(7.1), and varying,

$$\begin{aligned} 0 &= \delta S \\ &= \delta \int d^4x \left(-\frac{1}{4} F_{\alpha\beta} F^{\alpha\beta} - A_\alpha J^\alpha \right) \\ &= \int d^4x \left(-\frac{1}{2} F^{\alpha\beta} (\partial_\alpha \delta A_\beta - \partial_\beta \delta A_\alpha) - \delta A_\alpha J^\alpha \right) \\ &= \int d^4x \left(-F^{\alpha\beta} \partial_\alpha \delta A_\beta - \delta A_\alpha J^\alpha \right) \\ &= \int d^4x \left(\partial_\alpha F^{\alpha\beta} - J^\beta \right) \delta A_\beta \end{aligned}$$

so

$$\partial_\beta F^{\beta\alpha} = J^\alpha$$

In terms of the Faraday tensor, the Maxwell equations are therefore

$$\partial_\mu F_{\alpha\beta} + \partial_\alpha F_{\beta\mu} + \partial_\beta F_{\mu\alpha} = 0 \tag{7.7}$$

$$\partial_\beta F^{\beta\alpha} = J^\alpha \tag{7.8}$$

We combine the the homogeneous Eq.(7.7) with the explicit covariant form of the Faraday tensor, Eq.(7.5), by expanding into time and space parts and checking one combination of indices at a time. Since Eq.(7.7) is totally antisymmetric, each index must be distinct – setting $\alpha = 0, \beta = 0, \mu = 0$ or $\alpha = i, \beta = 0, \mu = 0$ simply gives zero.

Now let $\alpha = i, \beta = j, \mu = 0$. We find

$$F_{ij,0} + F_{j0,i} + F_{0i,j} = 0$$

This is

$$\begin{aligned} -\varepsilon_{ijk} \frac{\partial}{\partial t} B^k - \frac{\partial}{\partial x^i} E^j + \frac{\partial}{\partial x^j} E^i &= 0 \\ -\varepsilon_{ijk} \left(\frac{\partial}{\partial t} B^k + (\boldsymbol{\nabla} \times \mathbf{E})^k \right) &= 0 \\ \frac{\partial}{\partial t} \mathbf{B} + \boldsymbol{\nabla} \times \mathbf{E} &= 0 \end{aligned}$$

as expected. Finally, with all three indices spacelike,

$$\begin{aligned} F_{ij,k} + F_{jk,i} + F_{ki,j} &= 0 \\ -\varepsilon_{ijm} \partial_k B^m - \varepsilon_{jkm} \partial_i B^m - \varepsilon_{kim} \partial_j B^m &= 0 \\ \varepsilon_{ijk} \left(-\varepsilon_{ijm} \partial_k B^m - \varepsilon_{jkm} \partial_i B^m - \varepsilon_{kim} \partial_j B^m \right) &= 0 \\ -2\delta_k^m \partial_k B^m - 2\delta_i^m \partial_i B^m - 2\delta_j^m \partial_j B^m &= 0 \\ -6 \partial_m B^m &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned}$$

Now, for Eq.(7.8), we have

$$\partial_\beta F^{\beta\alpha} = J^\alpha$$

Let $\alpha = 0$ and we find Gauss' law,

$$\begin{aligned}
\partial_\beta F^{\beta 0} &= J^0 \\
\partial_i E^i &= \rho \\
\nabla \cdot \mathbf{E} &= \rho
\end{aligned}$$

Finally, with $\alpha = k$,

$$\begin{aligned}
\partial_\beta F^{\beta k} &= J^k \\
\partial_0 F^{0k} + \partial_j F^{jk} &= J^k \\
-\partial_0 E^k - \varepsilon_{jkm}\partial_j B^m &= J^k \\
-\partial_0 E^k + (\boldsymbol{\nabla} \times \mathbf{B})^k &= J^k \\
\boldsymbol{\nabla} \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} &= \mathbf{J}
\end{aligned}$$

Note that the Lagrange density is

$$\begin{aligned}
F_{\alpha\beta}F^{\alpha\beta} &= 2F_{0i}F^{0i} + F_{ij}F^{ij} \\
&= -2\mathbf{E}^2 + \left(-\varepsilon_{kij}B^k\right)\left(-\varepsilon_m{}^{ij}B^m\right) \\
&= -2\mathbf{E}^2 + \varepsilon_{kij}\varepsilon_m{}^{ij}B^k B^m \\
&= -2\mathbf{E}^2 - \varepsilon_{kij}\varepsilon^{mij}B^k B^m \\
&= -2\mathbf{E}^2 + 2\delta_k^m B^k B^m \\
&= -2\left(\mathbf{E}^2 - \mathbf{B}^2\right)
\end{aligned}$$

We may now turn to the Hamiltonian formulation and quantization.

7.2 Hamiltonian formulation of the Maxwell equations

Now we know that

$$\begin{aligned}
S &= -\frac{1}{4}\int d^4x \; F_{\alpha\beta}F^{\alpha\beta} \\
F_{\alpha\beta} &= \partial_\alpha A_\beta - \partial_\beta A_\alpha
\end{aligned}$$

We immediately hit a problem when we try to write the Hamiltonian formulation, because

$$\begin{aligned}
\pi_\mu(x) &= \frac{\delta S}{\delta \partial_0 A^\mu(x)} \\
&= \eta_{\mu\nu}\,\frac{\delta S}{\delta \partial_0 A_\nu(x)} \\
&= -\frac{1}{2}\eta_{\mu\nu}\int d^3x' \; F^{\alpha\beta}(x')\,\frac{\delta}{\delta\partial_0 A_\mu(x)}\left(\partial_\alpha A_\beta(x') - \partial_\beta A_\alpha(x')\right) \\
&= -\frac{1}{2}\eta_{\mu\nu}\int d^3x' \; F^{\alpha\beta}(x')\left(\delta^0_\alpha\delta^\mu_\beta - \delta^0_\beta\delta^\mu_\alpha\right)\delta^3(x-x') \\
&= -\frac{1}{2}\eta_{\mu\nu}\int d^3x' \; \left(F^{0\mu}(x') - F^{\mu 0}(x')\right)\delta^3(x-x') \\
&= -\eta_{\mu\nu}F^{0\nu}(x)
\end{aligned}$$

and therefore the conjugate momentum to A_0 vanishes:

$$\begin{aligned}
\pi_0 &= F^{00} = 0 \\
\pi_i &= F^{i0} = E^i
\end{aligned} \tag{7.9}$$

We should expect something like this. Since A_α is gauge dependent, not all of its components are physical.

There are several ways to deal with this problem. First, let's see what happens if we just ignore it. Then the Hamiltonian is

$$\begin{aligned}
H &= \int d^3x \left(\pi_0 \partial_0 A^0 + \pi_i \partial_0 A^i + \frac{1}{4} F_{\alpha\beta} F^{\alpha\beta}\right) \\
&= \int d^3x \left(0 + \pi_i \partial_0 A^i - \frac{1}{2} F_{i0} F^{i0} + \frac{1}{4} F_{ij} F^{ij}\right) \\
&= \int d^3x \left(E^i \partial_0 A_i - \frac{1}{2} E^i E^i + \frac{1}{4} F_{ij} F^{ij}\right) \\
&= \int d^3x \left(\pi^i A_{i,0} - \frac{1}{2} \pi^i \pi^i - \frac{1}{4} F_{ij} F^{ij}\right)
\end{aligned}$$

Remembering that $\varepsilon^{ijk} = -\varepsilon_{ijk}$,

$$\begin{aligned}
\mathbf{B} &= \nabla \times \mathbf{A} \\
B^m &= -\varepsilon^{mij} \left(A_{i,j} - A_{j,i}\right) \\
-\frac{1}{2} \varepsilon_{mij} B^m &= \frac{1}{2} \varepsilon_{mij} \varepsilon^{mkn} \left(A_{k,n} - A_{n,k}\right) \\
&= \frac{1}{2} \left(\delta_i^k \delta_j^n - \delta_i^n \delta_j^k\right) \left(A_{k,n} - A_{n,k}\right) \\
&= A_{i,j} - A_{j,i} \\
&= F_{ij}
\end{aligned}$$

$$\begin{aligned}
H &= \int d^3x \left(E^i A_{i,0} + \frac{1}{2} \pi^i \pi_i - \frac{1}{4} F_{ij} F^{ij}\right) \\
&= \int d^3x \left(E^i A_{i,0} + \frac{1}{2} \pi^i \pi_i - \frac{1}{4} F_{ij} F^{ij}\right)
\end{aligned}$$

$$\begin{aligned}
H &= -\int d^3x \left(\pi_\alpha \partial_0 A^\alpha - \frac{1}{2} F^{0i} \left(A_{0,i} - A_{i,0}\right) - \frac{1}{4} \left(A_{i,j} - A_{j,i}\right) \left(A^{i,j} - A^{j,i}\right)\right) \\
&= -\int d^3x \left(F_{i0} \partial_0 A^i - \frac{1}{2} F^{0i} \left(A_{0,i} - A_{i,0}\right) - \frac{1}{4} \left(A_{i,j} - A_{j,i}\right) \left(A^{i,j} - A^{j,i}\right)\right) \\
&= -\int d^3x \left(F^{0i} \partial_0 A^i - \frac{1}{2} F^{0i} \left(A_{0,i} + \partial_0 A^i\right) - \frac{1}{4} \left(A_{i,j} - A_{j,i}\right) \left(A^{i,j} - A^{j,i}\right)\right) \\
&= -\int d^3x \left(\frac{1}{2} F^{0i} \partial_0 A^i - \frac{1}{2} F^{0i} A_{0,i} - \frac{1}{4} \left(A_{i,j} - A_{j,i}\right) \left(A^{i,j} - A^{j,i}\right)\right) \\
&= -\int d^3x \left(\frac{1}{2} \pi^i \left(-A_{i,0} - A_{0,i}\right) - \frac{1}{4} \left(A_{i,j} - A_{j,i}\right) \left(A^{i,j} - A^{j,i}\right)\right) \\
&= \int d^3x \left(\frac{1}{2} \pi^i \pi^i + \pi^i A_{0,i} + \frac{1}{4} \left(-\varepsilon_{ijk} B^k\right) \left(-\varepsilon^{ijm} B_m\right)\right) \\
&= \int d^3x \left(\frac{1}{2} \pi^i \pi^i + \pi^i A_{0,i} + \frac{1}{2} B^i B^i\right)
\end{aligned}$$

where we have defined the magnetic field as

$$\begin{aligned} B^m &= \varepsilon^{mij}\left(A_{i,j} - A_{j,i}\right) \\ \frac{1}{2}B^m\varepsilon_{mij} &= A_{i,j} - A_{j,i} \\ \mathbf{B} &= \nabla \times \mathbf{A} \end{aligned}$$

Check

$$\begin{aligned} \varepsilon^{jnm}\partial_n B_m &= \varepsilon^{jnm}\partial_n\left(-\varepsilon_{mij}\left(A^{i,j} - A^{j,i}\right)\right) \\ &= -\varepsilon^{jnm}\varepsilon_{mij}\partial_n\left(A^{i,j} - A^{j,i}\right) \\ &= -\left(-\varepsilon^{njm}\right)\left(\varepsilon_{ijm}\right)\partial_n\left(A^{i,j} - A^{j,i}\right) \\ &= 2\delta^n_i\partial_n\left(A^{i,j} - A^{j,i}\right) \\ &= 2\partial_k\left(A^{i,j} - A^{j,i}\right) \end{aligned}$$

When the field equations are satisfied, we have $\nabla \cdot E = 0$. Then the middle term becomes a surface term which does not contribute to the field equations,

$$\begin{aligned} \int d^3x\ \left(\pi^i A_{0,i}\right) &= \int d^3x\ \left(E^i A_{0,i}\right) \\ &= \int d^3x\ \left(\partial_i\left(E^i A_0\right) - \left(\nabla \cdot E\right)A_0\right) \\ &= E^i A_0\big|_{boundary} \end{aligned}$$

and final expression is simply

$$H = \frac{1}{2}\int d^3x\ \left(\mathbf{E}^2 + \mathbf{B}^2\right)$$

In fact, throwing out the surface term, we can write the Hamiltonian in general as

$$H = \int d^3x\ \left(\frac{1}{2}\pi^i\pi^i + \frac{1}{2}B^iB^i - A_0\left(\partial_i\pi^i\right)\right) \tag{7.10}$$

Then A_0 appears as a Lagrange multiplier, enforcing $\nabla \cdot E = 0$ as a constraint. Since H is independent of π_0, we have $\dot{A}_0 = \{H, A_0\} = 0$.

Now we check Hamilton's equations:

$$\begin{aligned} \partial_0 A^0 &= \frac{\delta H}{\delta \pi_0} = 0 \\ \partial_0 A^i &= \frac{\delta H}{\delta \pi_i} \\ &= -\frac{\delta}{\delta \pi^i}\int d^3x\ \left(\frac{1}{2}\pi^i\pi^i + \pi^i A_{0,i} + \frac{1}{2}B^iB^i\right) \\ &= -\left(\pi^i + A_{0,i}\right) \end{aligned}$$

where $\frac{\delta}{\delta\pi_i} = -\frac{\delta}{\delta\pi^i}$. The first expression $\partial_0 A^0 = 0$, is a possible gauge choice but not necessary. Something about the formalism has forced this upon us. The second expression gives the usual expression for the electric field,

$$E^i = -\partial_0 A^i - \partial_i A^0$$

For the momentum we have

$$\begin{aligned}
\partial_0 \pi^0 &= -\frac{\delta H}{\delta A_0} \\
&= -\frac{\delta}{\delta A_0} \int d^3x' \left(\frac{1}{2}\pi^i \pi^i + \pi^i A_{0,i} + \frac{1}{2} B^i B^i \right) \\
&= \partial_i \pi^i
\end{aligned} \tag{7.11}$$

and

$$\begin{aligned}
\partial_0 \pi^j &= -\frac{\delta H}{\delta A_j} \\
&= -\frac{\delta}{\delta A_j} \int d^3x' \left(\frac{1}{2}\pi^i \pi^i + \pi^i A_{0,i} + \frac{1}{2} B^i B^i \right) \\
&= -\int d^3x' \left(\frac{\delta}{\delta A_j} \frac{1}{4} \left(A_{m,n} - A_{n,m}\right)\left(A^{m,n} - A^{n,m}\right) \right) \\
&= -\frac{1}{2} \int d^3x' \left(A^{m,n} - A^{n,m}\right) \frac{\delta}{\delta A_j} \left(A_{m,n} - A_{n,m}\right) \\
&= -\frac{1}{2} \int d^3x' \left(A^{m,n} - A^{n,m}\right) \left(\partial_n \frac{\delta}{\delta A_j} A_m - \partial_m \frac{\delta}{\delta A_j} A_n \right) \\
&= \frac{1}{2} \int d^3x' \left(\partial_n \left(A^{m,n} - A^{n,m}\right) \delta^j_m - \partial_m \left(A^{m,n} - A^{n,m}\right) \delta^j_n \right) \delta^3(x - x') \\
&= \partial_n \left(A^{j,n} - A^{n,j} \right) \\
&= \left(-\varepsilon^{jnm} \partial_n B_m \right) \\
&= -\left(\nabla \times \mathbf{B}\right)^j
\end{aligned}$$

which we may write as

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= 0 = 0 \\
\frac{\partial \mathbf{E}}{\partial t} + \nabla \times \mathbf{B} &= 0 = 0
\end{aligned}$$

using $\pi^0 = 0$. The final Maxwell equation follows automatically from our definition of the magnetic field as the curl of the potential, $\mathbf{B} = \nabla \times \mathbf{A}$. Notice that in order to get the complete set of equations we had to use all four conjugate momenta even though $\pi^0 \equiv 0$. So far, the only thing that has gone wrong is the emergence of the condition $\partial_0 A^0 = 0$, which is not a necessary consequence of Maxwell theory.

Now let's check the fundamental Poisson brackets. Normally we would expect

$$\left\{ \pi_\alpha, A^\beta \right\}_{P.B.} = \delta^\beta_\alpha \delta^3 \left(\mathbf{x} - \mathbf{x}' \right)$$

but we immediately see an inconsistency with $\pi_0 = 0$. Setting $\alpha = \beta = 0$, the right side does not vanish, but the left side does. Let's check it explicitly:

$$\begin{aligned}
\left\{ \pi_0, A^0 \right\}_{P.B.} &= \int d^3x' \left(\frac{\delta \pi_0}{\delta \pi_\alpha} \frac{\delta A^0}{\delta A^\alpha} - \frac{\delta \pi_0}{\delta A^\alpha} \frac{\delta A^0}{\delta \pi_\alpha} \right) \\
&= \int d^3x' \left(\frac{\delta (0)}{\delta \pi_\alpha} \frac{\delta A^0}{\delta A^\alpha} \right) \\
&= 0
\end{aligned}$$

While some ambiguity might be claimed for $\frac{\delta\pi_0}{\delta\pi_\alpha}$, consider the Poisson bracket of A^0 with anything else, computed in the (A^α, π_β) basis:

$$\begin{aligned}
\left\{f\left(A,\pi\right), A^0\right\}_{P.B.} &= \int d^3x' \left(\frac{\delta f}{\delta\pi_\alpha}\frac{\delta A^0}{\delta A^\alpha} - \frac{\delta f}{\delta A^\alpha}\frac{\delta A^0}{\delta\pi_\alpha}\right) \\
&= \int d^3x' \left(\frac{\delta f}{\delta\pi_\alpha}\delta^0_\alpha\delta^3\left(x-x'\right)\right) \\
&= \frac{\delta f}{\delta\pi_0}\left(x\right) \\
&= 0
\end{aligned}$$

since f will not ever depend on $\pi_0 = 0$.

One resolution of the dilemma is to choose a gauge in which A^0 also vanishes. Though such a gauge choice breaks manifest Lorentz covariance of the formulation, it is always possible. Suppose we begin with a generic form of the 4-potential $\tilde{A}_\alpha$. Then performing a gauge transformation to a new potential A_α we have

$$A_\alpha = \tilde{A}_\alpha + \partial_\alpha\varphi\left(t,\mathbf{x}\right)$$

and in particular we demand

$$0 = A_0 = \tilde{A}_0 + \partial_0\varphi\left(t,\mathbf{x}\right)$$

Therefore, we need only choose

$$\varphi\left(t,\mathbf{x}\right) = -\int\limits_{t'}^{t}\tilde{A}_0\left(t',\mathbf{x}\right)dt'$$

to eliminate A_0. Notice that this does not use all of the gauge freedom. If we choose, we can make another gauge transformation (say, by a function φ') as long as $\partial_0\varphi' = 0$. This just means that we can still adjust the gauge using an arbitrary function of the spatial coordinates, $\varphi\left(\mathbf{x}\right)$.

Now the problem has been shifted to a different location. By eliminating A^0 and π_0 from our list of independent variables, we have lost the ability to derive one of the Maxwell equations, $\nabla\cdot\mathbf{E} = 0$, since this follows from Eq.(7.11) and H no longer depends on A^0. This equation remains as a constraint that must be satisfied by hand. The remaining (equal time) Poisson brackets are

$$\left\{\pi_j\left(t,\mathbf{x}\right), A^i\left(t,\mathbf{x}'\right)\right\}_{P.B.} = \delta^i_j\delta^3\left(\mathbf{x}-\mathbf{x}'\right) \tag{7.12}$$

This still gives problems. Since momentum is given by the electric field, $\pi^i = E^i$, the divergence constraint requires

$$\frac{\partial}{\partial x^i}\left\{\pi^i\left(t,\mathbf{x}\right), A_j\left(t,\mathbf{x}'\right)\right\}_{P.B.} = \delta^i_j\frac{\partial}{\partial x^i}\delta^3\left(\mathbf{x}-\mathbf{x}'\right)$$

so that

$$0 = \left\{\nabla\cdot\mathbf{E}\left(\mathbf{x}\right), A_j\left(\mathbf{x}'\right)\right\}_{P.B.} = \delta^i_j\frac{\partial}{\partial x^i}\delta^3\left(\mathbf{x}-\mathbf{x}'\right)$$

and once again we have an inconsistency.

7.2.1 Handling the constraint

The differential condition that the divergence of the electric field vanish, $\nabla\cdot\mathbf{E} = \mathbf{0}$, may be turned into an algebraic condition by parameterizing our fields by wave number rather than position. Thus, our fields

$A^i(\mathbf{x})$ and $\pi_j(\mathbf{x})$ at any time t may be recast as Fourier transforms,

$$A^i(\mathbf{x},t) = \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2\omega}}\left(\tilde{A}^i(\mathbf{k})e^{ik_\alpha x^\alpha} + \tilde{A}^{i*}(\mathbf{k})e^{-ik_\alpha x^\alpha}\right) \tag{7.13}$$

$$\begin{aligned}\pi^i(\mathbf{x},t) &= -\partial_0 A^i \\ &= \frac{-i}{(2\pi)^{3/2}}\int d^3k\,\sqrt{\frac{\omega}{2}}\left(\tilde{A}^i(\mathbf{k})e^{ik_\alpha x^\alpha} - \tilde{A}^{i*}(\mathbf{k})e^{-ik_\alpha x^\alpha}\right)\end{aligned} \tag{7.14}$$

where ω is an as yet unspecified function of $\mathbf{k}^2$. We also easily find the inverse transforms,

$$\begin{aligned}\frac{1}{(2\pi)^{3/2}}\int d^3x A^i(\mathbf{x},t)e^{-ik_m x^m} &= \frac{1}{(2\pi)^3}\int d^3x\int\frac{d^3k'}{\sqrt{2\omega'}}\left(\tilde{A}^i(\mathbf{k}')e^{ik'_\alpha x^\alpha}e^{-ik_m x^m} + \tilde{A}^{i*}(\mathbf{k}')e^{-ik'_\alpha x^\alpha}e^{-ik_m x^m}\right) \\ &= \int\frac{d^3k'}{\sqrt{2\omega'}}\left(\tilde{A}^i(\mathbf{k}')e^{i\omega' t}\delta^3(\mathbf{k}-\mathbf{k}') + \tilde{A}^{i*}(\mathbf{k}')e^{-i\omega' t}\delta^3(\mathbf{k}+\mathbf{k}')\right) \\ &= \frac{1}{\sqrt{2\omega}}\left(\tilde{A}^i(\mathbf{k})e^{i\omega t} + \tilde{A}^{i*}(-\mathbf{k})e^{-i\omega t}\right) \\ \frac{1}{(2\pi)^{3/2}}\int d^3x\pi^i(\mathbf{x},t)e^{-ik_m x^m} &= \frac{-i}{(2\pi)^3}\int d^3x\int d^3k\,\sqrt{\frac{\omega}{2}}\left(\tilde{A}^i(\mathbf{k})e^{ik_\alpha x^\alpha} - \tilde{A}^{i*}(\mathbf{k})e^{-ik_\alpha x^\alpha}\right)e^{-ik_m x^m} \\ &= -i\sqrt{\frac{\omega}{2}}\left(\tilde{A}^i(\mathbf{k})e^{i\omega t} - \tilde{A}^{i*}(-\mathbf{k})e^{-i\omega t}\right)\end{aligned}$$

Solving for the transforms,

$$\begin{aligned}\tilde{A}^i(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega}{2}}\int d^3x\left(A^i(\mathbf{x},t) + \frac{i}{\omega}\pi^i(\mathbf{x},t)\right)e^{-ik_\alpha x^\alpha} \\ \tilde{A}^{i*}(-\mathbf{k}) &= \frac{1}{2(2\pi)^{3/2}}\int d^3x\left(\sqrt{2\omega}A^i(\mathbf{x},t) - i\sqrt{\frac{2}{\omega}}\pi^i(\mathbf{x},t)\right)e^{-ik_m x^m}e^{i\omega t} \\ \tilde{A}^{i*}(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega}{2}}\int d^3x\left(A^i(\mathbf{x},t) - \frac{i}{\omega}\pi^i(\mathbf{x},t)\right)e^{ik_\alpha x^\alpha}\end{aligned}$$

from which can directly show that the change to new variables, $\tilde{A}^i(\mathbf{k})$ and $-i\tilde{A}^{i*}(\mathbf{k})$, is canonical. To see this we compute the Poisson bracket,

$$\begin{aligned}\left\{\tilde{A}^i(\mathbf{k}), -i\tilde{A}^{i\dagger}(\mathbf{k}')\right\}_{A,\pi} &= -i\int d^3y\left(\frac{\delta\tilde{A}^i(\mathbf{k})}{\delta\pi_k(\mathbf{y})}\frac{\delta\tilde{A}^{j\dagger}(\mathbf{k}')}{\delta A^k(\mathbf{y})} - \frac{\delta\tilde{A}^i(\mathbf{k})}{\delta A^k(\mathbf{y})}\frac{\delta\tilde{A}^{j\dagger}(\mathbf{k}')}{\delta\pi_k(\mathbf{y})}\right) \\ &= \frac{-i}{(2\pi)^3}\int d^3y\int d^3x\int d^3x'\left[\frac{\delta}{\delta\pi_k(\mathbf{y})}\left(\sqrt{\frac{\omega}{2}}\left(\frac{i}{\omega}\pi^i(\mathbf{x},t)\right)e^{-ik_\alpha x^\alpha}\right)\frac{\delta}{\delta A^k(\mathbf{y})}\left(\sqrt{\frac{\omega'}{2}}\left(A^j(\mathbf{x}',t)\right)e^{ik'_\alpha}\right.\right. \\ &\qquad - \frac{\delta}{\delta A^k(\mathbf{y})}\left(\sqrt{\frac{\omega}{2}}\left(A^i(\mathbf{x},t)\right)e^{-ik_\alpha x^\alpha}\right)\frac{\delta}{\delta\pi_k(\mathbf{y})}\left(\sqrt{\frac{\omega'}{2}}\left(-\frac{i}{\omega'}\pi^j(\mathbf{x}',t)\right)\right.\end{aligned}$$

Carrying out the functional derivatives and integrating over the resulting delta functions, we have

$$\begin{aligned}\left\{\tilde{A}^i(\mathbf{k}), -i\tilde{A}^{i*}(\mathbf{k}')\right\}_{A,\pi} &= \frac{1}{2(2\pi)^3}\int d^3y\int d^3x\int d^3x'\sqrt{\frac{\omega}{\omega'}}\left(\delta^3(\mathbf{x}-\mathbf{y})\eta^{ij}\delta^3(\mathbf{x}'-\mathbf{y}) + \delta^3(\mathbf{x}-\mathbf{y})\delta^i_k\delta^3(\mathbf{x}'-\mathbf{y})\eta^{jk}\right)e^{-ik_\alpha x} \\ &= \frac{1}{(2\pi)^3}\int d^3x\sqrt{\frac{\omega'}{\omega}}\eta^{ij}e^{-i\left(k_\alpha-k'_\alpha\right)x^\alpha} \\ &= \eta^{ij}\sqrt{\frac{\omega'}{\omega}}e^{-i\left(\omega-\omega'\right)t}\delta^3(\mathbf{k}-\mathbf{k}') \\ &= \eta^{ij}\delta^3(\mathbf{k}-\mathbf{k}')\end{aligned}$$

Therefore, $\tilde{A}^i(\mathbf{k})$ and $-i\tilde{A}^{i*}(\mathbf{k})$ are just as good as $A^i(\mathbf{x},t)$ and $\pi_i(\mathbf{x},t)$ for describing the fields. We can equally well think of $\mathbf{x}$ or $\mathbf{k}$ as a continuous index for the "coordinates" $A^i(\mathbf{x},t)$, and we can write our Poisson brackets in terms of either set.

Now consider the awkward constraint, $\nabla\cdot\mathbf{E}(\mathbf{x})=0$. In the momentum basis this becomes

$$\begin{aligned} 0 &= \nabla_i \pi^i(\mathbf{x},t) \\ &= -\frac{i}{(2\pi)^{3/2}}\int d^3k\sqrt{\frac{\omega}{2}}\left(ik_i\tilde{A}^i(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}+k_i\tilde{A}^{i*}(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right) \end{aligned}$$

Inverting the Fourier transform shows that we must have the *algebraic* contraints,

$$\begin{aligned} k_i\tilde{A}^i(\mathbf{k}) &= 0 \\ k_i\tilde{A}^{i*}(\mathbf{k}) &= 0 \end{aligned}$$

We have therefore succeeded in finding a set of canonical variables in which the constraint equation is algebraic. The algebraic constraint simply says the the field A^i and its momentum are transverse, a fact which we already knew about electromagnetic waves.

Defining the transverse projection operator

$$P^i{}_j = \delta^i_j - \frac{k^ik_j}{\mathbf{k}^2} \tag{7.15}$$

satisfying $P^i{}_jk_i=0$. Then we can finally isolate the physical degrees of freedom as two polarization vectors,

$$\begin{aligned} \varepsilon^i &\equiv P^i{}_j\tilde{A}^j(\mathbf{k}) \\ \varepsilon^{i\dagger} &\equiv P^i{}_j\tilde{A}^{j\dagger}(\mathbf{k}) \end{aligned} \tag{7.16}$$

These automatically satisfy

$$\begin{aligned} k_i\varepsilon^i(\mathbf{k}) &= 0 \\ k_i\varepsilon^{i\dagger}(\mathbf{k}) &= 0 \end{aligned}$$

Finally, we compute the Poisson bracket of $\varepsilon^i(\mathbf{k})$ and $-i\varepsilon^{i\dagger}(\mathbf{k})$. Since $\varepsilon^i(\mathbf{k})$ and $-i\varepsilon^{i\dagger}(\mathbf{k})$ span the physical subspace, these are our fundamental Poisson brackets. To do this, we simply project the brackets we have already found for the transforms of the fields:

$$\begin{aligned} \left\{\varepsilon^i(\mathbf{k}),\varepsilon^{j\dagger}(\mathbf{k}')\right\}_{A,\pi} &= \left\{P^i_k\tilde{A}^k(\mathbf{k}),-iP^j_m\tilde{A}^{m\dagger}(\mathbf{k}')\right\}_{A,\pi} \\ &= P^i_kP^j_m\eta^{km}\delta^3(\mathbf{k}-\mathbf{k}') \\ &= P^{ij}\delta^3(\mathbf{k}-\mathbf{k}') \end{aligned}$$

Now the bracket is consistent: if we dot k^i into both sides we get zero, while on the two dimensional subspace spanned by ε^i, the bracket, P^{ij} reduces to a Kronecker delta.

Notice that if we transform back to the original, position-dependent variables, we get a projective Dirac delta,

$$P^i{}_j(\mathbf{x}-\mathbf{x}') \equiv \frac{1}{(2\pi)^3}\int d^3k\left(\delta^i_j-\frac{k^ik_j}{\mathbf{k}^2}\right)e^{ik_i(x-x')^i} \tag{7.17}$$

Then we have

$$\begin{aligned} \frac{\partial}{\partial x^i}P^i{}_j(\mathbf{x}-\mathbf{x}') &\equiv \frac{1}{(2\pi)^3}\int d^3k\left(\delta^i_j-\frac{k^ik_j}{\mathbf{k}^2}\right)\frac{\partial}{\partial x^i}e^{ik_i(x-x')^i} \\ &\equiv \frac{i}{(2\pi)^3}\int d^3k\left(k_j-k_j\frac{k^ik_i}{\mathbf{k}^2}\right)e^{ik_i(x-x')^i} \\ &= 0 \end{aligned}$$

and the corresponding Poisson bracket,

$$\left\{\pi_j(\mathbf{x}), A^i(\mathbf{x}')\right\}_{P.B.} = P^i{}_j(\mathbf{x}-\mathbf{x}')$$

is consistent, with one further caveat. Since P^{ij} is symmetric in i and j, we have not only

$$\left\{\partial^j\pi_j(\mathbf{x}), A^i(\mathbf{x}')\right\}_{P.B.} = \partial^j P^i{}_j(\mathbf{x}-\mathbf{x}') = 0$$

but also must have

$$\left\{\partial^j\pi_j(\mathbf{x}), \partial_i A^i(\mathbf{x}')\right\}_{P.B.} = \frac{\partial}{\partial x'^i} P^i{}_j(\mathbf{x}-\mathbf{x}') = 0$$

and therefore we need an additional gauge condition,

$$\nabla\cdot\mathbf{A} = 0 \tag{7.18}$$

From the Fourier expansion of A^i, we see that the condition already follows from $k_i\varepsilon^i = 0$, but we must also check that the condition is consistent with the gauge freedom of the potential.

To check the consistency, recall that we have some residual gauge freedom beyond what was required to set $A^0 = 0$. Now suppose we have imposed $A^0 = 0$, and in that gauge,

$$\nabla\cdot\mathbf{A} = f(\mathbf{x},t)$$

for some function $f(\mathbf{x},t)$. Then changing the gauge again by $\varphi(\mathbf{x},t)$, we have $\mathbf{A}' = \mathbf{A} + \nabla\varphi(\mathbf{x},t)$ so that the divergence of the new $\mathbf{A}'$ is given by

$$\begin{aligned}\nabla\cdot\mathbf{A}' &= \nabla\cdot(\mathbf{A}+\nabla\varphi(\mathbf{x},t)) \\ &= f(\mathbf{x},t) + \nabla^2\varphi(\mathbf{x},t)\end{aligned}$$

Demanding $\nabla\cdot\mathbf{A}' = 0$ is always possible since we know how to solve the Poisson equation, $\nabla^2\varphi(\mathbf{x},t) = -f$ using Green functions. In the case of vanishng boundary conditions at infinity, the solution is

$$\varphi(\mathbf{x},t) = -\frac{1}{4\pi}\int d^3x'\frac{\nabla\cdot\mathbf{A}(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|}$$

However, we also have to maintain $A^0 = 0$, which, in general, will change by the time derivative of φ :

$$\begin{aligned}A'^0 &= A^0 + \partial^0\varphi(\mathbf{x},t) \\ &= 0 + \frac{1}{4\pi}\int d^3x'\partial^0\frac{\nabla\cdot\mathbf{A}(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|}\end{aligned}$$

We are saved here by the constraint, Gauss' law, since interchanging the time derivative and the divergence,

$$-\partial^0\nabla\cdot\mathbf{A} = \partial_i\left(-\partial_0 A^i\right) = \partial_i E^i = 0$$

Therefore, $A'^0 = A^0 = 0$, and we have simultaneously imposed the pair of gauge conditions,

$$\begin{aligned}A^0 &= 0 \\ \nabla\cdot\mathbf{A} &= 0\end{aligned} \tag{7.19}$$

Notice that, as a consequence of these, A^α also satisfies the Lorentz gauge condition

$$\partial_\alpha A^\alpha = 0$$

With these two gauge conditions we have reduced the vector potential to two independent components. These correspond to the two polarization states of light. We now turn to the free solution and quantization.

7.2.2 Vacuum solution to classical E&M

First, we need the solutions to the classical theory. The homogeneous field equations, Eqs.(7.7) are automatically satisfied by writing $\mathbf{E}$ and $\mathbf{B}$ in terms of the potentials, and Gauss's law $\nabla \cdot \mathbf{E} = \rho = 0$ is satisfied by the gauge conditions Eqs.(7.19)

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \nabla \cdot \left(-\nabla A^0 - \frac{\partial}{\partial t}\mathbf{A} \right) \\ &= \nabla \cdot \left(-\frac{\partial}{\partial t}\mathbf{A} \right) \\ &= -\frac{\partial}{\partial t} \left(\nabla \cdot \mathbf{A} \right) \\ &= 0 \end{aligned}$$

The only remaining field equation is

$$\boldsymbol{\nabla} \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = \mathbf{J} = 0 \qquad (7.20)$$

Substituting the expressions for $\mathbf{B}$ and $\mathbf{E}$ in terms of the potentials given by Eqs.(7.3) and (7.4), into Eq.(7.20) gives:

$$\begin{aligned} 0 &= \boldsymbol{\nabla} \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} \\ &= \nabla \times (\nabla \times \mathbf{A}) - \frac{\partial}{\partial t} \left(-\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\varphi \right) \\ &= -\nabla^2 \mathbf{A} + \nabla (\nabla \cdot \mathbf{A}) - \frac{\partial^2 \mathbf{A}}{\partial t^2} - \frac{\partial}{\partial t} \boldsymbol{\nabla}\varphi \end{aligned}$$

so with $\varphi = 0$ and $\nabla \cdot \mathbf{A} = 0$ we have the wave equation for the vector potential,

$$0 = -\nabla^2 \mathbf{A} + \frac{\partial^2 \mathbf{A}}{\partial t^2} = \Box \mathbf{A}$$

Starting with the Fourier integral for $\mathbf{A}$, Eq.(7.13)

$$A^i(\mathbf{x}, t) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega}} \left(\tilde{A}^i(\mathbf{k}) e^{ik_\alpha x^\alpha} + \tilde{A}^{i*}(\mathbf{k}) e^{-ik_\alpha x^\alpha} \right)$$

we apply the required conditions. First, the wave equation,

$$\begin{aligned} 0 &= \Box A^i(\mathbf{x}, t) \\ &= \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega}} \left(\tilde{A}^i(\mathbf{k}) \Box e^{ik_\alpha x^\alpha} + \tilde{A}^{i*}(\mathbf{k}) \Box e^{-ik_\alpha x^\alpha} \right) \\ &= \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega}} k_\alpha k^\alpha \left(\tilde{A}^i(\mathbf{k}) e^{ik_\alpha x^\alpha} + \tilde{A}^{i*}(\mathbf{k}) e^{-ik_\alpha x^\alpha} \right) \end{aligned}$$

so we require

$$k_\alpha k^\alpha = \omega^2 - \mathbf{k}^2 = 0$$

This gives the angular frequency in terms of the wave vector. For positive energies, $\omega = \sqrt{\mathbf{k}^2}$.

Next, we have the gauge condition

$$\begin{aligned}
0 &= \boldsymbol{\nabla}\cdot\mathbf{A}(\mathbf{x},t)\\
&= \frac{1}{(2\pi)^{3/2}}\int\frac{d^3k}{\sqrt{2\omega}}\left(\boldsymbol{\nabla}\cdot\left(e^{ik_\alpha x^\alpha}\tilde{\mathbf{A}}(\mathbf{k})\right)+\boldsymbol{\nabla}\cdot\left(e^{-ik_\alpha x^\alpha}\tilde{\mathbf{A}}^*(\mathbf{k})\right)\right)\\
&= \frac{1}{(2\pi)^{3/2}}\int\frac{d^3k}{\sqrt{2\omega}}\left(ie^{ik_\alpha x^\alpha}\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k})-ie^{-ik_\alpha x^\alpha}\mathbf{k}\cdot\tilde{\mathbf{A}}^*(\mathbf{k})\right)
\end{aligned}$$

Thus $\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k})=0$ and $\mathbf{k}\cdot\tilde{\mathbf{A}}^*(\mathbf{k})=0$ leaving the transverse projections,

$$A^i(\mathbf{x},t)=\frac{1}{(2\pi)^{3/2}}\int\frac{d^3k}{\sqrt{2\omega}}\left(\varepsilon^i(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}+\varepsilon^{i*}(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)\tag{7.21}$$

The only differences from the scalar field is that here we have a different expansion for each polarization of the potential and this time the field equation gives us a simpler expression for the frequency in terms of the wave vector, $\omega=\sqrt{\mathbf{k}^2}$ because the photon has zero mass.

The conjugate momentum, $\pi_i=-\pi^i=-\partial_0 A^i$, follows immediately,

$$\pi_i(\mathbf{x},t) = -\frac{i}{(2\pi)^{3/2}}\int d^3k\sqrt{\frac{\omega}{2}}\left(\varepsilon^i(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}-\varepsilon^{i*}(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)\tag{7.22}$$

so we are ready to quantize.

Now we invert the transforms to find the solution for the mode amplitudes in terms of the fields,

$$\begin{aligned}
\frac{1}{(2\pi)^{3/2}}\int d^3xA^i(\mathbf{x},t)e^{i\mathbf{k}'\cdot\mathbf{x}} &= \frac{1}{(2\pi)^3}\int d^3x\int\frac{d^3k}{\sqrt{2\omega}}\left(\varepsilon^i(\mathbf{k})e^{i(\omega t-\mathbf{k}\cdot\mathbf{x})}+\varepsilon^{i*}(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)e^{i\mathbf{k}'\cdot\mathbf{x}}\\
&= \frac{1}{(2\pi)^3}\int\frac{d^3k}{\sqrt{2\omega}}\int d^3x\left(\varepsilon^i(\mathbf{k})e^{i\omega t}e^{-i\left(\mathbf{k}-\mathbf{k}'\right)\cdot\mathbf{x}}+\varepsilon^{i*}(\mathbf{k})e^{-i\omega t}e^{i\left(\mathbf{k}+\mathbf{k}'\right)\cdot\mathbf{x}}\right)\\
&= \int\frac{d^3k}{\sqrt{2\omega}}\left(\varepsilon^i(\mathbf{k})e^{i\omega t}\delta^3(\mathbf{k}-\mathbf{k}')+\varepsilon^{i*}(\mathbf{k})e^{-i\omega t}\delta^3(\mathbf{k}+\mathbf{k}')\right)\\
&= \frac{1}{\sqrt{2\omega'}}\left(\varepsilon^i(\mathbf{k}')e^{i\omega t}+\varepsilon^{i\dagger}(-\mathbf{k}')e^{-i\omega t}\right)
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{(2\pi)^{3/2}}\int d^3x\pi_i(\mathbf{x},t)e^{i\mathbf{k}'\cdot\mathbf{x}} &= -\frac{i}{(2\pi)^3}\int d^3x\int d^3k\sqrt{\frac{\omega}{2}}\left(\varepsilon^i(\mathbf{k})e^{i(\omega t-i\mathbf{k}\cdot\mathbf{x})}-\varepsilon^{i*}(\mathbf{k})e^{-i(\omega t-\mathbf{k}\cdot\mathbf{x})}\right)e^{i\mathbf{k}'\cdot\mathbf{x}}\\
&= -i\int d^3k\sqrt{\frac{\omega}{2}}\left(\varepsilon^i(\mathbf{k})e^{i\omega t}\delta^3(\mathbf{k}-\mathbf{k}')-\varepsilon^{i*}(\mathbf{k})e^{-i\omega t}\delta^3(\mathbf{k}+\mathbf{k}')\right)\\
&= -i\sqrt{\frac{\omega'}{2}}\left(\varepsilon^i(\mathbf{k}')e^{i\omega t}-\varepsilon^{i*}(-\mathbf{k}')e^{-i\omega t}\right)
\end{aligned}$$

At $t=0$, we solve

$$\begin{aligned}
\frac{1}{(2\pi)^{3/2}}\int d^3x\sqrt{2\omega}A^i(\mathbf{x},0)e^{i\mathbf{k}\cdot\mathbf{x}} &= \varepsilon^i(\mathbf{k})+\varepsilon^{i*}(-\mathbf{k})\\
\frac{i}{(2\pi)^{3/2}}\int d^3x\sqrt{\frac{2}{\omega'}}\pi_i(\mathbf{x},0)e^{i\mathbf{k}\cdot\mathbf{x}} &= \varepsilon^i(\mathbf{k})-\varepsilon^{i*}(-\mathbf{k})
\end{aligned}$$

so adding

$$\begin{aligned}
\varepsilon^i(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega}{2}}\int d^3x\left(A^i(\mathbf{x},0)+\frac{i}{\omega}\pi_i(\mathbf{x},0)\right)e^{i\mathbf{k}\cdot\mathbf{x}}\\
\varepsilon^{i*}(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega}{2}}\int d^3x\left(A^i(\mathbf{x},0)-\frac{i}{\omega}\pi_i(\mathbf{x},0)\right)e^{-i\mathbf{k}\cdot\mathbf{x}}
\end{aligned}$$

7.2.3 Quantization

Returning to our fundamental Poisson brackets for A^i and π_j at $t=0$,

$$\left\{\pi_j(\mathbf{x}), A^i(\mathbf{y})\right\}_{P.B.} = P^i{}_j(\mathbf{x}-\mathbf{y})$$

we introduce the commutators as

$$\left[\hat{\pi}_j(\mathbf{x}), \hat{A}^i(\mathbf{y})\right] = i\hbar P^i{}_j(\mathbf{x}-\mathbf{y})$$

From this we compute the commutator of the mode amplitudes,

$$\begin{aligned}
\left[\hat{\varepsilon}^i(\mathbf{k}), \hat{\varepsilon}^{j\dagger}(\mathbf{k}')\right] &= \left[\frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega}{2}}\int d^3x\left(A^i(\mathbf{x},0)+\frac{i}{\omega}\pi_i(\mathbf{x},0)\right)e^{i\mathbf{k}\cdot\mathbf{x}}, \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\omega'}{2}}\int d^3y\left(A^j(\mathbf{y},0)-\frac{i}{\omega}\pi_j(\mathbf{y},0)\right)e^{-i\mathbf{k}'\cdot\mathbf{y}}\right] \\
&= \frac{1}{(2\pi)^3}\frac{\sqrt{\omega\omega'}}{2}\int d^3x\int d^3y e^{i\mathbf{k}\cdot\mathbf{x}-i\mathbf{k}'\cdot\mathbf{y}}\left[\hat{A}^i(\mathbf{x},0)+\frac{i}{\omega}\hat{\pi}_i(\mathbf{x},0), A^j(\mathbf{y},0)-\frac{i}{\omega}\pi_j(\mathbf{y},0)\right] \\
&= \frac{1}{(2\pi)^3}\frac{\sqrt{\omega\omega'}}{2}\int d^3x\int d^3y e^{i\mathbf{k}\cdot\mathbf{x}-i\mathbf{k}'\cdot\mathbf{y}}\left(-\frac{i}{\omega}\left[\hat{A}^i(\mathbf{x},0), \pi_j(\mathbf{y},0)\right]+\frac{i}{\omega}\left[\hat{\pi}_i(\mathbf{x},0), A^j(\mathbf{y},0)\right]\right) \\
&= \frac{1}{(2\pi)^3}\frac{\sqrt{\omega\omega'}}{2}\frac{i}{\omega}\int d^3x\int d^3y e^{i\mathbf{k}\cdot\mathbf{x}-i\mathbf{k}'\cdot\mathbf{y}}\left(i\hbar P^i{}_j(\mathbf{x}-\mathbf{y})+i\hbar P^j{}_i(\mathbf{x}-\mathbf{y})\right) \\
&= \frac{1}{(2\pi)^3}\frac{\sqrt{\omega\omega'}}{\omega}\hbar\int d^3x\int d^3y e^{i\mathbf{k}\cdot\mathbf{x}-i\mathbf{k}'\cdot\mathbf{y}}P^{ij}(\mathbf{x}-\mathbf{y})
\end{aligned}$$

Now, substituting from Eq.(7.17),

$$\begin{aligned}
\left[\hat{\varepsilon}^i(\mathbf{k}), \hat{\varepsilon}^{j\dagger}(\mathbf{k}')\right] &= \frac{1}{(2\pi)^6}\sqrt{\frac{\omega'}{\omega}}\int d^3k''\int d^3x\int d^3y\left(\delta^{ij}-\frac{k^ik^j}{\mathbf{k}^2}\right)e^{i\mathbf{k}\cdot\mathbf{x}-i\mathbf{k}'\cdot\mathbf{y}}e^{i\mathbf{k}''\cdot(\mathbf{x}-\mathbf{y})} \\
&= \frac{1}{(2\pi)^6}\sqrt{\frac{\omega'}{\omega}}\int d^3k''\int d^3x\int d^3y\left(\delta^{ij}-\frac{k^ik^j}{\mathbf{k}^2}\right)e^{i(\mathbf{k}+\mathbf{k}'')\cdot\mathbf{x}}e^{-i(\mathbf{k}'+\mathbf{k}')\cdot\mathbf{y}} \\
&= \sqrt{\frac{\omega'}{\omega}}\int d^3k''\left(\delta^{ij}-\frac{k^ik^j}{\mathbf{k}^2}\right)\delta^3(\mathbf{k}+\mathbf{k}'')\delta^3(\mathbf{k}'+\mathbf{k}'')
\end{aligned}$$

and therefore

$$\left[\hat{\varepsilon}^i(\mathbf{k}), \hat{\varepsilon}^{j\dagger}(\mathbf{k}')\right] = P^{ij}\delta^3(\mathbf{k}-\mathbf{k}') \qquad (7.23)$$

The mode amplitudes are therefore raising and lowering operators. As with scalar and spinor fields, we could go on and define a complete set of energy eigenstates using these raising and lowering operators.

Exercise: Define the space of states of the electromagnetic field.

It is most convenient to rewrite each mode amplitude as a product of an operator with polarization 4-vector.

$$\hat{\varepsilon}^i(\mathbf{k}) = \varepsilon^{\alpha}_{(i)}(\mathbf{k})\,\hat{a}_{(i)}(\mathbf{k}) \qquad (7.24)$$

The operators $\hat{a}^{\dagger}_{(i)}(\mathbf{k})$ and $\hat{a}_{(i)}(\mathbf{k})$ now create or anihilate one quantum of light with wave vector $\mathbf{k}$ and polarization vector $\varepsilon^{\alpha}_{(i)}$. For each $i=1,2$, $\varepsilon^{\alpha}_{(i)}(\mathbf{k})$ is a spacelike 4-vector giving one of the two polarization directions.

The two vectors $\varepsilon^{\alpha}_{(i)}(\mathbf{k})$ satisfy a pair of covariant constraints, in place of $k_i\varepsilon^i(\mathbf{k})=0$. One of the pair of constraints expresses the transverse condition, while the additional constraint gives $\varepsilon^{\alpha}_{(i)}(\mathbf{k})$ a vanishing time component in the current frame of reference, i.e., the Lorentz reference frame in which we fixed $A^0 = 0$. In this frame, let t^α be the unit timelike vector, $t^\alpha = (1, \mathbf{0})$. This allows us to rewrite our gauge conditions in a Lorentz invariant way,

$$\begin{aligned} t^\alpha A_\alpha &= 0 \\ \partial_\alpha A^\alpha &= 0 \end{aligned}$$

Exercise: Show that the two conditions

$$\begin{aligned} t^\alpha A_\alpha &= 0 \\ \partial_\alpha A^\alpha &= 0 \end{aligned}$$

are equivalent to the gauge conditions

$$\begin{aligned} A^0 &= 0 \\ \nabla \cdot \mathbf{A} &= 0 \end{aligned}$$

Now, noting that $k^\alpha = (\omega, \mathbf{k})$, demand

$$\begin{aligned} t_\alpha \varepsilon^\alpha_{(i)}(\mathbf{k}) &= 0 \\ k_\alpha \varepsilon^\alpha_{(i)}(\mathbf{k}) &= 0 \end{aligned}$$

The first equation reduces each $\varepsilon^\alpha_{(i)}(\mathbf{k})$ to a purely spatial vector, $\varepsilon^\alpha_{(i)}(\mathbf{k}) = \left(0, \varepsilon_{(i)}(\mathbf{k})\right)$, and the second then reduces to $\mathbf{k} \cdot \varepsilon_{(i)}(\mathbf{k}) = 0$, as required. We may also choose the two polarizations $\varepsilon^\alpha_{(i)}(\mathbf{k})$ to be orthonormal

$$\varepsilon^\alpha_{(i)}(\mathbf{k}) \varepsilon_{(j)\alpha}(\mathbf{k}) = -\delta_{ij}$$

Exercise: In a frame of reference where $t^\alpha = (1, 0, 0, 0)$ for an electromagnetic wave in the z direction (i.e., $k^\alpha = (0, 0, 0, 1)$), find expressions for $\varepsilon^\alpha_{(1)}(\mathbf{k})$ and $\varepsilon^\alpha_{(2)}(\mathbf{k})$.

We can now write the field operator in final form:

$$\hat{A}^\alpha (\mathbf{x}, t) = \frac{1}{(2\pi)^{3/2}} \sum_{i=1}^{2} \int \frac{d^3k}{\sqrt{2\omega}} \left(\varepsilon^\alpha_{(i)} (\mathbf{k})\, \hat{a}_{(i)}(\mathbf{k}) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} + \varepsilon^{\alpha\dagger}_{(i)} (\mathbf{k})\, \hat{a}^\dagger_{(i)}(\mathbf{k}) e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \right) \tag{7.25}$$

Chapter 8

Appendices

8.1 Appendix A: The Casimir operators of the Poincaré group.

The Lie algebra of the Poincaré group is:

$$\begin{aligned}
[M_{\alpha\beta}, M_{\rho\sigma}] &= \eta_{\beta\rho}M_{\alpha\sigma} - \eta_{\beta\sigma}M_{\alpha\rho} - \eta_{\alpha\rho}M_{\beta\sigma} + \eta_{\alpha\sigma}M_{\beta\rho} \\
\left[M^{\alpha}{}_{\beta}, P_{\nu}\right] &= \eta_{\nu\beta}P^{\alpha} - \delta^{\alpha}_{\nu}P_{\beta} \\
[P_{\alpha}, P_{\beta}] &= 0
\end{aligned}$$

Exercise: Prove that P^2 and W^2 are Casimir operators of the Poincaré group, where

$$\begin{aligned}
P^2 &= \eta^{\alpha\beta}P_{\alpha}P_{\beta} \\
W^2 &= \eta_{\alpha\beta}W^{\alpha}W^{\beta}
\end{aligned}$$

and W^{α} is given by

$$W^{\mu} = \frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}P_{\nu}M_{\alpha\beta}$$

It is easy to show that P^2 is a Casimir operator. Just compute

$$\begin{aligned}
\left[P_{\mu}, P^2\right] &= \eta^{\alpha\beta}\left[P_{\mu}, P_{\alpha}P_{\beta}\right] \\
&= \eta^{\alpha\beta}P_{\alpha}\left[P_{\mu}, P_{\beta}\right] + \eta^{\alpha\beta}\left[P_{\mu}, P_{\alpha}\right]P_{\beta} \\
&= 0
\end{aligned}$$

and (with $M_{\mu\nu} = \eta_{\mu\alpha}M^{\alpha}{}_{\nu}$),

$$\begin{aligned}
\left[M_{\mu\nu}, P^2\right] &= \eta^{\alpha\beta}\left[M_{\mu\nu}, P_{\alpha}P_{\beta}\right] \\
&= \eta^{\alpha\beta}P_{\alpha}\left[M_{\mu\nu}, P_{\beta}\right] + \eta^{\alpha\beta}\left[M_{\mu\nu}, P_{\alpha}\right]P_{\beta} \\
&= \eta^{\alpha\beta}P_{\alpha}\left(\eta_{\nu\beta}P_{\mu} - \eta_{\mu\beta}P_{\nu}\right) \\
&\quad + \eta^{\alpha\beta}\left(\eta_{\nu\alpha}P_{\mu} - \eta_{\mu\alpha}P_{\nu}\right)P_{\beta} \\
&= P_{\nu}P_{\mu} - P_{\mu}P_{\nu} + P_{\mu}P_{\nu} - P_{\nu}P_{\mu} \\
&= 0
\end{aligned}$$

Now we turn to W^2. First find the commutator of P_μ with a single W_β,

$$\begin{aligned}
[P_\mu, W_\beta] &= \left[P_\mu, \varepsilon_\beta^{\ \ \nu\rho\sigma} P_\nu M_{\rho\sigma}\right] \\
&= \varepsilon_\beta^{\ \ \nu\rho\sigma} P_\nu \left[P_\mu, M_{\rho\sigma}\right] \\
&= -\varepsilon_\beta^{\ \ \nu\rho\sigma} P_\nu \left[M_{\rho\sigma}, P_\mu\right] \\
&= -\varepsilon_\beta^{\ \ \nu\rho\sigma} P_\nu \left(\eta_{\sigma\mu} P_\rho - \eta_{\rho\mu} P_\sigma\right) \\
&= -\varepsilon_\beta^{\ \ \nu\rho\sigma} \left(\eta_{\sigma\mu} P_\nu P_\rho - \eta_{\rho\mu} P_\nu P_\sigma\right) \\
&= 0
\end{aligned}$$

and therefore,

$$\begin{aligned}
\left[P_\mu, W^2\right] &= \eta^{\alpha\beta} W_\alpha \left[P_\mu, W_\beta\right] + \eta^{\alpha\beta} \left[P_\mu, W_\alpha\right] W_\beta \\
&= 0
\end{aligned}$$

Finally, consider the co mmutator of W^2 with $M_{\alpha\beta}$. This is automatic since W^2 is a scalar; alternatively we may use the Lorentz invariance to boost to $P_\mu = m\,(1,0,0,0)$. Then

$$\begin{aligned}
W_\mu &= \frac{1}{2}\varepsilon_\mu^{\ \ \nu\alpha\beta} P_\mu M_{\alpha\beta} \\
&= \frac{m}{2}\varepsilon_\mu^{\ \ 0\alpha\beta} M_{\alpha\beta} \\
W^\mu &= -\frac{m}{2}\varepsilon^{0\mu\alpha\beta} M_{\alpha\beta}
\end{aligned}$$

so in this frame,

$$\begin{aligned}
W^0 &= 0 \\
W^i &= -\frac{m}{2}\varepsilon^{0ijk} M_{jk} \\
&= -\frac{m}{2}\varepsilon^{ijk} M_{jk}
\end{aligned}$$

Now, contracting a pair of Levi-Civita tensors with the spatial part of the Lorentz commutator,

$$\begin{aligned}
\left[M_{\alpha\beta}, M_{\rho\sigma}\right] &= \eta_{\beta\rho} M_{\alpha\sigma} - \eta_{\beta\sigma} M_{\alpha\rho} - \eta_{\alpha\rho} M_{\beta\sigma} + \eta_{\alpha\sigma} M_{\beta\rho} \\
\left[M_{jk}, M_{mn}\right] &= \eta_{km} M_{jn} - \eta_{jm} M_{kn} - \eta_{kn} M_{jm} + \eta_{jn} M_{km} \\
&= -\delta_{km} M_{jn} + \delta_{jm} M_{kn} + \delta_{kn} M_{jm} - \delta_{jn} M_{km} \\
\left[\varepsilon^{ijk} M_{jk}, \varepsilon^{lmn} M_{mn}\right] &= \varepsilon^{ijk}\varepsilon^{lmn}\left(-\delta_{km} M_{jn} + \delta_{jm} M_{kn} + \delta_{kn} M_{jm} - \delta_{jn} M_{km}\right) \\
&= \varepsilon^{ijk}\varepsilon^{lnk} M_{jn} + \varepsilon^{ikm}\varepsilon^{lnm} M_{kn} + \varepsilon^{ijk}\varepsilon^{lmk} M_{jm} + \varepsilon^{ikn}\varepsilon^{lmn} M_{km} \\
&= \varepsilon^{ijk}\varepsilon^{lnk} M_{jn} + \varepsilon^{ijm}\varepsilon^{lnm} M_{jn} + \varepsilon^{ijk}\varepsilon^{lnk} M_{jn} + \varepsilon^{ijm}\varepsilon^{lnm} M_{jn} \\
&= 4\varepsilon^{ijk}\varepsilon^{lnk} M_{jn}
\end{aligned}$$

so that

$$\begin{aligned}
\frac{4}{m^2}\left[W^i, W^l\right] &= 4\varepsilon^{ijk}\varepsilon^{lnk} M_{jn} \\
&= 4\varepsilon^{ijk}\varepsilon^{lnk} M_{jn}
\end{aligned}$$

Now using

$$\begin{aligned}
\varepsilon_{kij}\varepsilon^{kmn} M_{mn} &= \left(\delta_i^m \delta_j^n - \delta_i^n \delta_j^m\right) M_{mn} \\
&= 2M_{ij}
\end{aligned}$$

we have

$$\begin{aligned} \frac{4}{m^2}\left[W^i, W^l\right] &= 2\varepsilon^{ijk}\varepsilon^{lnk}\varepsilon_{rjn}\varepsilon^{rst}M_{st} \\ &= 2\left(\delta^i_r\delta^k_n - \delta^k_r\delta^i_n\right)\varepsilon^{lnk}\varepsilon^{rst}M_{st} \\ &= -2\left(\varepsilon^{li}{}_r\varepsilon^{rst}\right)M_{st} \\ &= -4\varepsilon^{li}{}_r W^r \end{aligned}$$

Now compute

$$\begin{aligned} \varepsilon^{ijk}\left[M_{ij}, W^2\right] &= 2\left[W^k, W^2\right] \\ &= 2\delta_{mn}\left(\left[W^k, W^m\right]W^n + W^m\left[W^k, W^n\right]\right) \\ &= -2m^2\delta_{mn}\left(\varepsilon^{km}{}_s W^s W^n + W^m\varepsilon^{kn}{}_s W^s\right) \\ &= -2m^2\varepsilon^k{}_{ms}\left(W^s W^m + W^m W^s\right) \\ &= 0 \end{aligned}$$

so in this Lorentz frame, $\left[M_{ij}, W^2\right] = 0$. But W^2 is Lorentz invariant, so the commutastor vanishes in every Lorentz frame.

8.1.1 Appendix B: Completeness relation for Dirac solutions

Exercise: Prove the completeness relation,

$$\sum_{a=1}^{2}\left(\left[u_a(p^\alpha)\right]^A\left[\bar{u}_a(p^\alpha)\right]_B - \left[v_a(p^\alpha)\right]^A\left[\bar{v}_a(p^\alpha)\right]_B\right) = \delta^A_B$$

where $A, B = 1, \ldots, 4$ label the components of the basis spinors, and the spinors are given by

$$\begin{aligned} \left[u_1(p^\alpha)\right]^A &= \sqrt{\frac{E+m}{2m}}\begin{pmatrix} 1 \\ 0 \\ \frac{p^z}{E+m} \\ \frac{p^x+ip^y}{E+m} \end{pmatrix}; \left[u_2(p^\alpha)\right]^A = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} 0 \\ 1 \\ \frac{p^x-ip^y}{E+m} \\ \frac{-p^z}{E+m} \end{pmatrix} \\ \left[v_1(p^\alpha)\right]^A &= \sqrt{\frac{m+E}{2m}}\begin{pmatrix} \frac{p^z}{E+m} \\ \frac{p^x+ip^y}{E+m} \\ 1 \\ 0 \end{pmatrix}; \left[v_2(p^\alpha)\right]^A = \sqrt{\frac{m+E}{2m}}\begin{pmatrix} \frac{p^x-ip^y}{E+m} \\ \frac{-p^z}{E+m} \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

We also have the barred spinors given by $\bar{u}_a = u_a^\dagger h$ and $\bar{v}_a = v_a^\dagger h$:

$$\begin{aligned} \left[\bar{u}_1(p^\alpha)\right]_A &= \sqrt{\frac{E+m}{2m}}\begin{pmatrix} 1 \\ 0 \\ -\frac{p^z}{E+m} \\ -\frac{p^x-ip^y}{E+m} \end{pmatrix}; \left[\bar{u}_2(p^\alpha)\right]_A = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} 0 \\ 1 \\ -\frac{p^x+ip^y}{E+m} \\ \frac{p^z}{E+m} \end{pmatrix} \\ \left[\bar{v}_1(p^\alpha)\right]_A &= \sqrt{\frac{m+E}{2m}}\begin{pmatrix} \frac{p^z}{E+m} \\ \frac{p^x-ip^y}{E+m} \\ -1 \\ 0 \end{pmatrix}; \left[\bar{v}_2(p^\alpha)\right]_A = \sqrt{\frac{m+E}{2m}}\begin{pmatrix} \frac{p^x+ip^y}{E+m} \\ \frac{-p^z}{E+m} \\ 0 \\ -1 \end{pmatrix} \end{aligned}$$

First, compute the individual products. For the u-type spinors,

$$[u_1]^A [\bar{u}_1]_B = \frac{E+m}{2m} \begin{pmatrix} 1 & 0 & -\frac{p^z}{E+m} & -\frac{p^x-ip^y}{E+m} \\ 0 & 0 & 0 & 0 \\ \frac{p^z}{E+m} & 0 & -\left(\frac{p^z}{E+m}\right)^2 & -\frac{p^z(p^x-ip^y)}{(E+m)^2} \\ \frac{p^x+ip^y}{E+m} & 0 & -\frac{p^z(p^x+ip^y)}{(E+m)^2} & -\frac{(p^x)^2+(p^y)^2}{(E+m)^2} \end{pmatrix}$$

$$[u_2]^A [\bar{u}_2]_B = \frac{E+m}{2m} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{p^x+ip^y}{E+m} & \frac{p^z}{E+m} \\ 0 & \frac{p^x-ip^y}{E+m} & -\frac{(p^x)^2+(p^y)^2}{(E+m)^2} & \frac{p^z(p^x-ip^y)}{(E+m)^2} \\ 0 & \frac{-p^z}{E+m} & \frac{p^z(p^x+ip^y)}{(E+m)^2} & -\left(\frac{p^z}{E+m}\right)^2 \end{pmatrix}$$

so the sum is

$$[u_1]^A [\bar{u}_1]_B + [u_2]^A [\bar{u}_2]_B = \frac{E+m}{2m} \begin{pmatrix} 1 & 0 & -\frac{p^z}{E+m} & -\frac{p^x-ip^y}{E+m} \\ 0 & 1 & -\frac{p^x+ip^y}{E+m} & \frac{p^z}{E+m} \\ \frac{p^z}{E+m} & \frac{p^x-ip^y}{E+m} & -\frac{\mathbf{p}^2}{(E+m)^2} & 0 \\ \frac{p^x+ip^y}{E+m} & \frac{-p^z}{E+m} & 0 & -\frac{\mathbf{p}^2}{(E+m)^2} \end{pmatrix}$$

For the v-type spinors, we find

$$[v_1]^A [\bar{v}_1]_B = \frac{m+E}{2m} \begin{pmatrix} \left(\frac{p^z}{E+m}\right)^2 & \frac{p^z(p^x-ip^y)}{(E+m)^2} & \frac{-p^z}{E+m} & 0 \\ \frac{p^z(p^x+ip^y)}{(E+m)^2} & \frac{(p^x)^2+(p^y)^2}{(E+m)^2} & -\frac{p^x+ip^y}{E+m} & 0 \\ \frac{p^z}{E+m} & \frac{p^x-ip^y}{E+m} & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$[v_2]^A [\bar{v}_2]_B = \frac{m+E}{2m} \begin{pmatrix} \frac{(p^x)^2+(p^y)^2}{(E+m)^2} & -\frac{p^z(p^x-ip^y)}{(E+m)^2} & 0 & -\frac{p^x-ip^y}{E+m} \\ -\frac{p^z(p^x+ip^y)}{(E+m)^2} & \left(\frac{p^z}{E+m}\right)^2 & 0 & \frac{p^z}{E+m} \\ 0 & 0 & 0 & 0 \\ \frac{p^x+ip^y}{E+m} & \frac{-p^z}{E+m} & 0 & -1 \end{pmatrix}$$

with sum

$$[v_1]^A [\bar{v}_1]_B + [v_2]^A [\bar{v}_2]_B = \frac{m+E}{2m} \begin{pmatrix} \frac{\mathbf{p}^2}{(E+m)^2} & 0 & \frac{-p^z}{E+m} & -\frac{p^x-ip^y}{E+m} \\ 0 & \frac{\mathbf{p}^2}{(E+m)^2} & -\frac{p^x+ip^y}{E+m} & \frac{p^z}{E+m} \\ \frac{p^z}{E+m} & \frac{p^x-ip^y}{E+m} & -1 & 0 \\ \frac{p^x+ip^y}{E+m} & \frac{-p^z}{E+m} & 0 & -1 \end{pmatrix}$$

The difference between the sum of the u-type and the sum of the v-type matrices is

$$\frac{1}{2m}\left(1 - \frac{\mathbf{p}^2}{(E+m)^2}\right) \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} = \delta^A_B$$

so the full completeness relation is

$$\sum_{a=1}^{2} \left([u_a(p^\alpha)]^A [\bar{u}_a(p^\alpha)]_B - [v_a(p^\alpha)]^A [\bar{v}_a(p^\alpha)]_B \right) = \delta^A_B$$

as claimed.

Bibliography

[1] Antonio N. Bernal and Miguel Sánchez, "On smooth Cauchy hypersurfaces and Geroch's splitting theorem", Communications in Mathematical Physics 243 (2003), no. 3, 461–470.