

# Class Notes for Quantum Field Theory: Section I

Introduction to 2nd Quantization, Lagrangian and Equations of Motion, Conservation Laws, the Klein Gordon field, the Dirac field, Spin-Statistics connection, Feynman Propagators,

Electromagnetic fields

Notes based on Mandl & Shaw, “Quantum Field Theory”

Best text for later use: Peskin and Schroeder, “Quantum Field Theory”

Jack Gunion  
U.C. Davis

230A, U.C. Davis, Fall Quarter

# Outline and Course Goals

1. Become familiar with fields — both classical and quantized.

By “field” we refer to some quantity defined as a function of  $(\vec{x}, t)$  over all of space-time. An example you are familiar with would be the vector potential field  $A(\vec{x}, t)$  of electromagnetic theory.

2. Learn why we wish to quantize a field and what it means to quantize a field.
3. Learn about the relation between field quantization and particles (something that applies when fields are defined on the space-time continuum) and how this is a natural extension of the relationship between quantization of ions on a lattice and phonons.
4. Understand the connection: symmetries of the Lagrangian  $\Leftrightarrow$  conserved quantities.

5. Learn why every particle must have an antiparticle in a quantized relativistic theory (and the CPT theorem).
6. Understand why a **local, Lorentz invariant, causal, 2nd quantized** relativistic field theory must have the observed connection between spin and statistics.

## Field Theory

1. Learn about free-particle propagators, especially the difference between Feynman, retarded and advanced propagators.
2. Learn about the association between fields and interactions, e.g. E&M field  $\Rightarrow$  E&M force.

Also, the association of particles with fields:  $\Rightarrow$  new force requires a new particle.

3. Develop calculation techniques, i.e **Feynman Rules** for transition amplitudes and cross sections:
  - Perturbation theory.

- Wick's Theorem.
- Time ordered products.
- A simple example of a Feynman diagram is that for  $e^-e^- \rightarrow e^-e^-$  via single photon exchange. The photon gives rise to the E&M force between the two charged electrons (photons couple to charge).

The photon in this diagram is “virtual” since it has  $q^2 < 0$  rather than  $q^2 = 0$ . Computing this diagram gives an interaction between the two electrons that varies as  $1/q^2$ .

This is only the simplest of many very complicated diagrams that should all be summed together at the amplitude level to obtain the full interaction. Because the charge describing the photon-electron coupling is small, the simple diagram is a pretty good approximation, but all the other diagrams have some influence.

In particular, the “higher-order” diagrams cause the strength of the inter-electron force to vary with  $q^2$  as  $e^2(q^2)/q^2$  where  $e^2(q^2)$  increases as  $|q^2|$  increases in magnitude.

That is, we can codify the behavior of the diagram sum in terms of a “moving coupling constant” — the momentum dependence of the strength

of the inter-electron force is summarized by an effective coupling strength (effective charge) that varies with momentum.

- For each type of force and matter particle, we will write down a Lagrangian describing the interaction of the force particle with the matter particle.

As new forces/particles are discovered, we add new pieces to our Lagrangian and try to organize the Lagrangian so that all the symmetries ... are transparently displayed.

### Ultimate Goal

Write down a single fundamental Lagrangian that describes all the fundamental particles and interactions (after “2nd quantization”) that are now relevant or that were relevant in the past.

- We currently imagine that it was at the time of the big-bang that the largest number of fundamental particles and forces were “active” and in some kind of short-lived “equilibrium”. Since then, many of the particles have decayed or annihilated and many of the forces have become irrelevant, leaving us with those observed in every day life and those we can probe using the accelerators that we have built for this purpose.

- If there is (or was) a “God”, a possible view of his/her role was to choose this ultimate Lagrangian.
- We currently believe that there might be other causally disconnected universes whose evolution was controlled by a different Lagrangian (or possibly a version of the same Lagrangian with different “charges” and other parameters).

Theorists even discuss the “anthropic principle” according to which the Lagrangian that we “see” is the only one that our form of life could see, but that there are other forms of life in other universes that “see” a completely different Lagrangian.

# System of units and Conventions

## Units

- It will be convenient to use a system of units in which  $\hbar = c = 1$ .

- In this system of units:

$$L = T = E^{-1} = M^{-1}.$$

$$\text{e.g. } m = mc^2 = \frac{mc}{\hbar}$$

$$\text{e.g. } m_e = 9.109 \times 10^{-28} \text{ g} = 0.511 \text{ MeV (where MeV} = 10^6 \text{ eV)}.$$

- Some convenient conversion factors are:

- $(1 \text{ GeV})^{-1} = 0.197 \times 10^{-13} \text{ cm} = 0.197 \text{ fermi, where } 1 \text{ GeV} \equiv 10^9 \text{ eV.}$
- $(1 \text{ GeV})^{-2} = 0.389 \times 10^{-27} \text{ cm}^2 = 0.389 \text{ mb, where mb is short for milli-barn.}$

- In Electrodynamics: use so-called Heaviside-Lorentz conventions in which the factors of  $4\pi$  appear in Coulomb's Law and the fine structure constant rather than in Maxwell's equations.

i.e. the Coulomb potential for a point charge  $Q$  is

$$\Phi = \frac{Q}{4\pi r} \quad (1)$$

and the fine structure constant is

$$\alpha = \frac{e^2}{4\pi} = \frac{e^2}{4\pi\hbar c} \simeq \frac{1}{137} \quad (2)$$

where the latter value actually only applies in the long-wavelength (low-momentum-transfer) limit.

**Metric and Lorentz Notation Conventions**



- We raise and lower indices and construct Lorentz 4-invariants using

$$g^{\mu\nu} = g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (3)$$

Note that the consistency requirement of  $g^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} g_{\alpha\beta}$  is satisfied and that  $g_{\nu}^{\mu} = \delta_{\nu}^{\mu}$ .

- $x^{\mu} = (x^0, \vec{x})$  (careful if you have used Chau's notes that used to have this for  $x_{\mu}$  contrary to usual definition).

Vectors with raised indices are called “contravariant” and those with lowered indices are “covariant”.  $x_{\mu} = g_{\mu\nu} x^{\nu} = (x^0, -\vec{x})$ .

- $p \cdot x = g_{\mu\nu} p^{\mu} x^{\nu} = p^0 x^0 - \vec{p} \cdot \vec{x}$
- $p^2 = E^2 - |\vec{p}|^2 = m^2$
- $\partial_{\mu} = \frac{\partial}{\partial x^{\mu}} = \left( \frac{\partial}{\partial x^0}, \vec{\nabla} \right)$

- $\epsilon^{0123} = +1, \epsilon_{0123} = -1, \epsilon^{1230} = -1, \dots$

- For old “1st quantization”,

$E = i\frac{\partial}{\partial x^0}, \vec{p} = -i\vec{\nabla}$  which is summarized in the form  $p^\mu = i\partial^\mu$ , where the raised index accounts for the  $-$  sign in  $\vec{p} = -i\vec{\nabla}$ .

- Using the notation

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu \quad (4)$$

for a Lorentz transformation,  $x'^\mu x'_\mu = x^\mu x_\mu$  requires

$$\Lambda^{\lambda\mu} \Lambda_{\lambda\nu} = \delta^\mu_\nu \quad (5)$$

- If  $\phi(x)$  is a scalar function, then so is  $\delta\phi = \frac{\partial\phi}{\partial x^\mu} \delta x^\mu$ .

Hence,  $\frac{\partial\phi}{\partial x^\mu} \equiv \partial_\mu\phi \equiv \phi_{,\mu}$  is a covariant four-vector.

This is because, a scalar function is such that it is invariant under the Lorentz transform, just like the general dot product,  $ab = a^\mu b_\mu$ , so that since  $\delta x^\mu$  is a contravariant vector like  $a^\mu$ , then the object multiplying it to create a scalar function must be a covariant vector.

## General Description of what we will be doing and motivations

Our goal will be to develop Feynman Rules for the calculation of fundamental processes involving elementary particles. A prototype theory for which we wish to develop Feynman rules is QED.

In using Feynman rules for QED, we take a fundamental process involving photons and electrons and draw diagrams that can contribute to the process in question. We associate rules for writing down a mathematical expression for the Quantum Mechanical amplitude associated with each diagram.

As usual in QM, we sum the amplitudes for all the diagrams and only take the square of the net amplitude (which means that different diagrams interfere with one another).

An example, after extending QED to include muons as well as electrons, would be  $e^+e^- \rightarrow \mu^+\mu^-$ .

The relevant diagram is

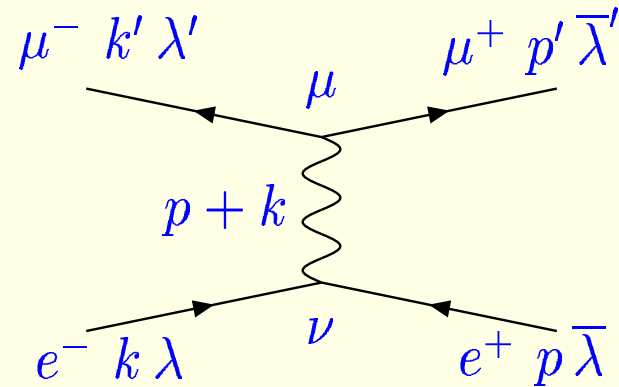


Figure 1: The one Feynman diagram contributing to  $e^+e^- \rightarrow \mu^+\mu^-$ .

- The presence of spins and different momenta for the particles must all be taken into account in some way.
- The photon connecting the  $e^+e^-$  to the  $\mu^+\mu^-$  cannot be a real photon since it must have  $(p+k)^2 > 0$ ; in fact  $(p+k)^2$  could be very large at a high energy collider.

This kind of photon is called “off-shell” or “virtual”.

In quantum mechanics, it is ok for a particle to be off its mass-shell so long as it doesn't hang around too long (uncertainty principle).

## Particles $\Leftrightarrow$ Fields

- The association between particles and fields has a long history beginning with the association of photons with the electromagnetic field.
- In the diagram we just drew, it is the virtual photon that is responsible for the interaction between the  $e^+e^-$  and the  $\mu^+\mu^-$  and the virtual photon is a quantum of the electromagnetic field.
- The electrons and muons, and their anti-particles, will turn out to be the quanta of the electron and muon fields.
- Only by using this kind of description in terms of the quanta of the fields can we account for processes where particles can be created or annihilated (as in our Feynman diagram).

Relativistic quantum mechanics does not do the job — the Dirac wave function for the electron, for example, only describes what a single electron

does in interaction with a potential. Creation or annihilation of that electron is not possible.

## Problems for relativistic QM

Some you have hopefully already encountered.

- negative energy states
- Klein Paradox
- probability not easily defined
- probability can “disappear”

Such failures and the need for a field viewpoint could have been anticipated.

- $E = mc^2$  clearly allows for pair creation processes.

Even when energy is inadequate for real pair creation, pair states can appear in 2nd order perturbation theory so long as  $\Delta E \Delta t \lesssim \hbar$ . (again “virtual” processes)

- Causality is violated.

To see this, let us write transition amplitude for particle propagation from  $\vec{x}_0$  at  $t = 0$  to  $\vec{x}$  at time  $t$ . We employ the unitary time translation operator which is given in terms of the Hamiltonian for the particle:

$$\begin{aligned}
 U(t) &= \langle \vec{x} | e^{-iHt} | \vec{x}_0 \rangle \\
 &= \langle \vec{x} | e^{-it\sqrt{\vec{P}_{\text{op}}^2 + m^2}} | \vec{x}_0 \rangle \\
 &= \int d^3\vec{p} \langle \vec{x} | e^{-it\sqrt{\vec{P}_{\text{op}}^2 + m^2}} | \vec{p} \rangle \langle \vec{p} | \vec{x}_0 \rangle \\
 &= \frac{1}{(2\pi)^3} \int d^3\vec{p} e^{-it\sqrt{p^2 + m^2}} e^{i\vec{p}\cdot(\vec{x} - \vec{x}_0)}, \tag{6}
 \end{aligned}$$

Note how we inserted a complete set of plane-wave states  $I = \int d^3\vec{p} |\vec{p}\rangle \langle \vec{p}|$  in the momentum basis and used  $\langle \vec{x} | \vec{p} \rangle = \frac{1}{(2\pi)^{3/2}} e^{-i\vec{p}\cdot\vec{x}}$ .

Now take  $\vec{x}_0 = \vec{0}$  and consider  $|\vec{x}| \gg t$  (i.e. well outside the light cone), go to one dimension for simplicity, and use stationary phase techniques for the exponent  $ipx - it\sqrt{p^2 + m^2}$  (with stationary point in  $p$  at  $p =$

$imx/\sqrt{x^2 - t^2}$ ) to get

$$\begin{aligned} U(t) &\sim \exp \left[ -it \sqrt{-\frac{x^2 m^2}{x^2 - t^2} + m^2} + i \left( \frac{ixm}{\sqrt{x^2 - t^2}} \right) x \right] \\ &= \exp \left[ -it \sqrt{\frac{-t^2 m^2}{x^2 - t^2}} - \frac{x^2 m}{\sqrt{x^2 - t^2}} \right] \\ &= \exp \left[ -m \sqrt{x^2 - t^2} \right] \end{aligned} \tag{7}$$

which is exponentially small, but **non-zero**, implying that causality is violated.

- Quantum field theory solves this problem by virtue of the fact that there are both particles and antiparticles propagating across the space-like interval, and their amplitudes cancel one another, thereby preserving causality.

For processes that do not violate causality (i.e. involve a time-like separation), this cancellation does not occur (although both particle and anti-particle propagation is occurring).



## The Electromagnetic Field – no sources, i.e. “free” field

We will first 2nd quantize this field, since it is the one with which you are most familiar. We will be glossing over various subtleties in order to develop some intuition.

First let us summarize classical electromagnetic theory. Presumably you all are familiar with:

- Maxwell’s equations
- the choice of a gauge such as the Coulomb (radiation) gauge  $\vec{\nabla} \cdot \vec{A} = 0$ , which for a plane wave state  $\vec{A}(x, t) = \vec{A}_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)}$  becomes  $\vec{k} \cdot \vec{A} = 0$ .

In this gauge, the equation of motion for  $\vec{A}$  becomes  $\square \vec{A} = 0$  ( $\square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2$ ) and

$$\vec{B} = \vec{\nabla} \times \vec{A}, \quad \vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} \quad (8)$$

for which the Hamiltonian or energy is given by

$$H_{\text{rad}} = \frac{1}{2} \int (\vec{E}^2 + \vec{B}^2) d^3 \vec{x} \quad (9)$$

- We will want to formulate things in Fourier space. In order to keep things most comprehensible, we will go to a finite volume in space, rather than the continuum.

We will use periodic boundary conditions in each direction of length  $L$ .

Ultimately, the volume  $V$  must cancel out of all calculations.

For periodic b.c., the momenta  $\vec{k}$  in the box are multiples of  $2\pi/L$  in each direction,  $\vec{k} = \frac{2\pi}{L}(n_1, n_2, n_3)$ ,  $n_i = 0, \pm 1, \dots$ , so that in the infinite volume limit

$$\sum_{\text{states}} = \sum_{n_1 n_2 n_3} = \frac{L^3}{(2\pi)^3} \sum_{\vec{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3\vec{k}. \quad (10)$$

- In our finite volume, we use periodic boundary conditions in the form  $\vec{A}(0, y, z, t) = \vec{A}(L, y, z, t), \dots$

In this case the functions  $\frac{1}{\sqrt{V}}\vec{\epsilon}_r(\vec{k})e^{i\vec{k}\cdot\vec{x}}$ ,  $r = 1, 2$  form a complete set of transverse orthonormal vector fields.

Here,

$$\vec{\epsilon}_r(\vec{k}) \cdot \vec{\epsilon}_s(\vec{k}) = \delta_{rs}, \quad \vec{\epsilon}_r(\vec{k}) \cdot \vec{k} = 0, \quad r, s = 1, 2 \quad (11)$$

where the latter equation is required by the radiation gauge  $\vec{k} \cdot \vec{A} = 0$ .

The above setup corresponds to a linear polarization basis with polarizations orthogonal to  $\vec{k}$ . We could also employ circular polarizations, but that would be less convenient for the moment, since it is easiest to keep everything explicitly real.

- This allows us to expand  $\vec{A}(\vec{x}, t)$  as a Fourier series:

$$\vec{A}(\vec{x}, t) = \sum_{\vec{k}} \sum_r \left( \frac{\hbar c^2}{2V \omega_{\vec{k}}} \right)^{1/2} \vec{\epsilon}_r(\vec{k}) \left[ a_r(\vec{k}, t) e^{i\vec{k} \cdot \vec{x}} + a_r(\vec{k}, t)^* e^{-i\vec{k} \cdot \vec{x}} \right] \quad (12)$$

where  $\omega_{\vec{k}} = c|\vec{k}|$ . Note that  $\omega_{\vec{k}}$  has dimensions of  $T^{-1}$ . The normalization factor is chosen quite purposefully as we shall later discuss. The construction is such that  $\vec{A}$  is explicitly real.

- From  $\square \vec{A} = 0$ , we find

$$\frac{\partial^2}{\partial t^2} a_r(\vec{k}, t) = -\omega_k^2 a_r(\vec{k}, t). \quad (13)$$

This is the equation of motion for a harmonic oscillator **for each of the**

**infinitely large number of  $(r, \vec{k})$  choices.**

**The solution is:  $a_r(\vec{k}, t) = a_r(\vec{k}) \exp(-i\omega_{\vec{k}}t)$ .**

- **If we work out  $H_{\text{rad}}$  we find**

$$H_{\text{rad}} = \sum_{\vec{k}} \sum_r \hbar \omega_{\vec{k}} a_r^*(\vec{k}) a_r(\vec{k}), \quad (14)$$

**where  $\hbar$  appeared because of the normalization factor appearing in the mode expansion of the  $A$  field. Note that  $H_{\text{rad}}$  is time-independent as expected for the total electromagnetic energy inside the box (in the absence of interactions).**

- **Net result: a perfectly fine reformulation of the classical electromagnetic field.  $a_r(\vec{k})$  represents a set of numbers and nothing more.**

## 2nd Quantization

- **We will quantize the electromagnetic field by hypothesizing that each of these infinite number of harmonic oscillator equations should be quantized in exactly the way we quantize a single harmonic oscillator system, and see what happens.**

- The way it is actually done, and we will eventually do it this way, is to:
  1. treat the fields themselves (at each  $\vec{x}, t$  location) as coordinates;
  2. use the Lagrangian/Hamiltonian of the field theory to determine the momentum conjugate to the fields when treated as coordinates;
  3. introduce the analogue of the usual QM commutation relation between “coordinates” and “momenta”; this is where that  $\sqrt{\hbar}$  and other factors introduced in the normalization of the  $\vec{A}$  expansion given in Eq. (12) comes into play. There is an  $\hbar$  that we require for the commutator of the “coordinate”  $\vec{A}$  field with its conjugate “momentum” ( $\propto \frac{\partial}{\partial t} \vec{A}$ ).
  4. and then reformulate that QM CR in terms of operators in the Fourier transform space.

The above operators are precisely the  $a_r(\vec{k})$  in the case of the E&M field, and so the  $a_r(\vec{k})$  and  $a_r^*(\vec{k})$  objects are no longer simple numbers but rather operators that work exactly like the  $a$  and  $a^\dagger$  operators of the harmonic oscillator. It is just that there is one such operator for every  $r, \vec{k}$  choice.

Harmonic Oscillator Reminder

- $H_{\text{osc}} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2$  with  $[q, p] = i\hbar$  is reinterpreted by writing

$$a, a^\dagger = \frac{1}{\sqrt{2\hbar m\omega}}(m\omega q \pm ip), \quad \Rightarrow \quad [a, a^\dagger] = 1 \quad (15)$$

and simple computation gives  $H_{\text{osc}} = \hbar\omega(a^\dagger a + \frac{1}{2})$ .

- One defines a ground state  $|0\rangle$  with energy  $\frac{1}{2}\hbar\omega$  and excited states given by

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle, \quad E_n = \hbar\omega(n + \frac{1}{2}). \quad (16)$$

- Further, we go from one occupation number to another via the raising and lowering operators, e.g.

$$a|n\rangle = n^{1/2}|n-1\rangle, \quad a^\dagger|n\rangle = (n+1)^{1/2}|n+1\rangle. \quad (17)$$

- By direct computation, for an operator such as  $a$  that is time-independent in the Schroedinger picture, its equation of motion in the Heisenberg picture

(see Appendix to Chapter 1 of Mandl-Shaw for review) is

$$i\hbar \frac{da(t)}{dt} = [a(t), H_{\text{osc}}] = \hbar\omega a(t), \quad \Rightarrow \quad a(t) = ae^{-i\omega t}. \quad (18)$$

### Back to 2nd quantization of the E&M field

By analogy, we promote  $a_r(\vec{k})$  and  $a_r^*(\vec{k})$  to operators  $a_r(\vec{k})$  and  $a_r^\dagger(\vec{k})$  obeying the commutation relations

$$\begin{aligned} [a_r(\vec{k}), a_s^\dagger(\vec{k}')] &= \delta_{rs} \delta_{\vec{k}\vec{k}'} \\ [a_r(\vec{k}), a_s(\vec{k}')] &= [a_r^\dagger(\vec{k}), a_s^\dagger(\vec{k}')] = 0, \end{aligned} \quad (19)$$

and write the quantized Hamiltonian **operating in the space of states defined below** as

$$H_{\text{op}} = \sum_{\vec{k}} \sum_r \hbar\omega_{\vec{k}} \left( a_r^\dagger(\vec{k}) a_r(\vec{k}) + \frac{1}{2} \right). \quad (20)$$

Note that the commutation relations assume that each “mode”  $(\vec{k}, r)$  is completely independent of every other mode; i.e. all these pseudo-harmonic oscillator systems are independent of one another.

**Note:** We can now crudely understand a bit better about the  $\sqrt{\hbar}$  in the normalization factor of the  $\vec{A}$  expansion. Roughly, the commutator between the “coordinate”  $\vec{A}$  and its conjugate “momentum” takes the form

$$[A^i(\vec{x}, t), \frac{\partial}{\partial t} A^j(\vec{x}', t)] = i\hbar \delta^3(\vec{x} - \vec{x}') \delta^{ij}, \quad (21)$$

implying that the  $\sqrt{\hbar}$  is needed when we have harmonic oscillator type normalization for the  $a$  and  $a^\dagger$  commutators, as given above in Eq. (19).

Continuing on, clearly the states

$$|n_r(\vec{k})\rangle = \frac{[a_r^\dagger(\vec{k})]^{n_r(\vec{k})}}{\sqrt{n_r(\vec{k})!}} |0\rangle \quad (22)$$

obey

$$N_r(\vec{k}) |n_r(\vec{k})\rangle = n_r(\vec{k}) |n_r(\vec{k})\rangle \quad (23)$$

where  $N_r(\vec{k}) = a_r^\dagger(\vec{k}) a_r(\vec{k})$ . The eigenfunctions of the radiation Hamiltonian



take the form (the  $|0\rangle$  really only appears once far to the right)

$$\prod_{\vec{k}_i} \prod_{r_i} |n_{r_i}(\vec{k}_i)\rangle = |n_1(\vec{k}_1)\rangle |n_2(\vec{k}_1)\rangle |n_1(\vec{k}_2)\rangle |n_2(\vec{k}_2)\rangle \dots \quad (24)$$

and have energy (computed as  $H_{\text{op}} \prod_{\vec{k}_i} \prod_{r_i} |n_{r_i}(\vec{k}_i)\rangle \equiv E \prod_{\vec{k}_i} \prod_{r_i} |n_{r_i}(\vec{k}_i)\rangle$ )

$$E = \sum_{\vec{k}} \sum_r \hbar\omega_{\vec{k}} \left( n_r(\vec{k}) + \frac{1}{2} \right). \quad (25)$$

Raising and lowering works as before: e.g.

$$a_r(\vec{k}) | \dots, n_r(\vec{k}), \dots \rangle = [n_r(\vec{k})]^{1/2} | \dots, n_r(\vec{k}) - 1, \dots \rangle. \quad (26)$$

Correspondingly, the eigenvalue of the energy **operator**,  $H_{\text{op}}$ , is reduced by  $\hbar\omega_{\vec{k}} = \hbar c|\vec{k}|$ .

If we write for the momentum operator the analogue of what we wrote for the energy operator,  $H_{\text{op}}$ , something that we will justify later (it simply

follows from the E&M flux vector  $\vec{S}$ ), we have

$$\vec{P}_{\text{op}} = \sum_{\vec{k}} \sum_r \hbar \vec{k} \left( N_r(\vec{k}) + \frac{1}{2} \right) = \sum_{\vec{k}} \sum_r \hbar \vec{k} N_r(\vec{k}) \quad (27)$$

and then  $a_r(\vec{k})$  also reduces the momentum eigenvalue of the state by  $\hbar \vec{k}$ . We could also get an expression for the angular momentum (using the circular polarization basis) and by computation show that a single photon state was one with a single unit of spin.

## General Process

- Write down the classical expression for energy, momentum, angular momentum.
- Rewrite this expression by substituting in the “Fourier” expansion of the  $A$  field (but treating  $a$ 's and  $a^\dagger$ 's as operators) and performing the  $\int d^3\vec{x}$  to get an *operator* form of the classical expression.
- Compute what happens when the operator form of the quantity acts upon a photon state to see what interpretation is appropriate.

The results we have obtained lead to ....

## Interpretation

- $a_r(\vec{k})$  is an annihilation operator that removes one photon in the mode  $(\vec{k}, r)$  of energy  $\hbar\omega_{\vec{k}}$  and momentum  $\hbar\vec{k}$  from the system state.

Thus, our little game gives a result in which the  $a$  and  $a^\dagger$  operators “look, feel, taste, ...” like they annihilate and create a single photon. If further investigation reveals no conflict with this interpretation, then probably our guess of how to formulate multi-photon states is correct.

- Our states can have an arbitrary number of photons. In fact, we can have any number of photons of the same  $(\vec{k}, r)$  value  $\Rightarrow$  Bose statistics is implicit.

$\Rightarrow$  something in all this will have to change when we go to fermions.

- You might worry that the ground state (with no photons) has infinite energy given by

$$\frac{1}{2} \sum_{\vec{k}} \sum_r, \quad (28)$$

i.e. by the sum of all the “zero-point” energies of the infinite number of harmonic oscillators.

However, this energy is unobservable in the sense that all we can detect are excitations relative to the ground state.

- Fortunately, the ground state does have zero momentum. Non-zero momentum would be observable in a very real sense.
- We have not yet checked causality, and we will delay this for a while.

How should we check causality in the context of the operator field  $\vec{A}(\vec{x}, t)$ ?  
We should require

$$[\vec{A}(\vec{x}, t), \vec{A}(\vec{x}', t')] = 0 \quad (29)$$

whenever  $(t - t')^2 - (\vec{x} - \vec{x}')^2 < 0$ , i.e. when the two *operator* fields are located at a space-like separation they should commute.

By simply carrying out the computation of the commutator using the algebraic expressions in hand and the commutators for the  $a$ 's and  $a^\dagger$ 's, we find that this is true.

- However, if we do the same thing for a spin-1/2 Dirac field  $\psi(\vec{x}, t)$ , we encounter a problem.

First, we note that the analogues of the  $\epsilon_r(\vec{k})e^{i\vec{k}\cdot\vec{x}-i\omega_{\vec{k}}t}$  “plane waves” are the Dirac spinor forms,  $u_s(\vec{k})e^{i\vec{k}\cdot\vec{x}-iE_{\vec{k}}t}$ , and similarly for the complex

**conjugate.** (Note that the algebraic form of  $\mathbf{u}_s(\vec{k})$  is completely fixed by the requirement that  $\psi(\vec{x}, t)$  obey the Dirac equation.) The result is that  $\psi(\vec{x}, t)$  takes the form

$$\psi(\vec{x}, t) = \sum_{\vec{k}} \sum_s \left( \frac{1}{2V E_{\vec{k}}} \right)^{1/2} \left[ b_s(\vec{k}) \mathbf{u}_s(\vec{k}) e^{i\vec{k} \cdot \vec{x} - iE_{\vec{k}} t} + d_s^\dagger(\vec{k}) \mathbf{v}_s(\vec{k}) e^{-i\vec{k} \cdot \vec{x} + iE_{\vec{k}} t} \right]. \quad (30)$$

Employing this form, and assuming commutation relations like those for the E&M  $a$  for the  $b$  and  $d$  operators we don't get zero when computing  $[\psi(\vec{x}, t), \psi(\vec{x}', t')]$  for space-like separations.

We must switch everything from commutation to anticommutation and get 0 anticommutator for space-like separation of two Dirac fields.

**Is this acceptable?** We will see that it is since the Dirac fields themselves are not observable. The observable operators [e.g. bilinear forms like  $\psi^\dagger(\vec{x}, t)\psi(\vec{x}, t)$ ] constructed from them will *commute* for space-like separation.

However, because we have fundamental anticommutation relations among the creation and annihilation operators, we will have Fermi-Dirac statistics.

- A bit of notation.

It is often convenient to divide the operator field  $\vec{A}$  into its positive and

negative frequency components which correspond to the part with the annihilation operator and the part with the creation operator, respectively.

$$\vec{A}(\vec{x}, t) = \vec{A}^+(\vec{x}, t) + \vec{A}^-(\vec{x}, t) \quad (31)$$

with

$$\begin{aligned} \vec{A}^+(\vec{x}, t) &= \sum_{\vec{k}} \sum_r \left( \frac{\hbar c^2}{2V \omega_{\vec{k}}} \right)^{1/2} \vec{e}_r(\vec{k}) a_r(\vec{k}) e^{i(\vec{k} \cdot \vec{x} - \omega_{\vec{k}} t)} \\ \vec{A}^-(\vec{x}, t) &= \sum_{\vec{k}} \sum_r \left( \frac{\hbar c^2}{2V \omega_{\vec{k}}} \right)^{1/2} \vec{e}_r(\vec{k}) a_r^\dagger(\vec{k}) e^{-i(\vec{k} \cdot \vec{x} - \omega_{\vec{k}} t)} \end{aligned} \quad (32)$$

Note:  $e^{-i\omega t}$  when operated on by  $E_{\text{op}} = i\hbar \frac{\partial}{\partial t}$  gives  $E = +\hbar\omega$ .

### Summary of differences between 2nd and 1st quantization

2nd quantization refers to the quantization of fields, in which the fields themselves become operators. The coordinates  $(\vec{x}, t)$  remain simple numbers.

1st quantization refers to non-relativistic or relativistic quantum mechanics in which we made the  $\vec{x}$  coordinate into an operator and obtained a wave equation for a non-operator wave function.

The 2nd-quantized theory, of course, gives the same results as the 1st quantized approach when energies and such are small enough that only 1 particle needs to be considered, with creation and annihilation processes being negligible.

One final note

By direct calculation

$$\langle \text{state with definite number of photons} | \vec{E} | \text{same state} \rangle = 0 \quad (33)$$

since the  $\vec{E}$  field, like the  $\vec{A}$  field from which it is computed, either annihilates or creates an extra photon and then the resulting state does not overlap with the starting state.

To get a non-zero expectation value for  $\vec{E}$  requires that the state in question be a superposition of states with different numbers of photons. (Kind of like a wave-packet state in QM can only have a non-zero expectation value for both  $\vec{x}$  and  $\vec{p}$  if it is a superposition of many plane waves.)

Thus, the usual classical  $\vec{E}$  fields seen in undergraduate laboratory experiments result from a coherent superposition of states with different, but large numbers (in order to have small fluctuation), of photons.

**Homework:** Do problem 1.2 of Mandl-Shaw.



## Appendix: Derivation of e.o.m. for $\vec{A}$ in Coulomb Gauge

- Maxwell's equations:

$$\begin{aligned}\vec{\nabla} \cdot \vec{E} &= 0 & \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} & \vec{\nabla} \times \vec{B} &= \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t}\end{aligned}\quad (34)$$

- Move to vector potential.

1. Write

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad (35)$$

then  $\vec{\nabla} \cdot \vec{B} = 0$  is automatic.

2. Write

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} \quad (36)$$

then  $\vec{\nabla} \cdot \vec{E} = 0$  provided that we are in Coulomb gauge of  $\vec{\nabla} \cdot \vec{A} = 0$ .

3.  $\vec{\nabla} \times \vec{E} = \vec{\nabla} \times -\frac{\partial \vec{A}}{\partial t} = -\frac{\partial \vec{\nabla} \times \vec{A}}{\partial t} = -\frac{\partial \vec{B}}{\partial t}$  where we employed Eq. (36) and then Eq. (35). This result implies that the 3rd Maxwell equation is automatically satisfied.
4. So, it is only the final Maxwell equation that will give a non-trivial e.o.m. for the  $\vec{A}$  field. The left hand side of this last ME is:

$$\vec{\nabla} \times \vec{B} = \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = -\vec{\nabla} \cdot \vec{\nabla} \vec{A} + \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) = -\vec{\nabla} \cdot \vec{\nabla} \vec{A} \quad (37)$$

where we used the Coulomb gauge condition  $\vec{\nabla} \cdot \vec{A} = 0$  in the last step. Using Eq. (36), the right hand side is

$$\frac{1}{c^2} \frac{\partial}{\partial t} \left( -\frac{\partial \vec{A}}{\partial t} \right) = -\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} \quad (38)$$

Setting these equal and moving all to one side gives the equation of motion

$$\square \vec{A} = 0, \quad \text{where} \quad \square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \quad (39)$$

# The Lagrangian Approach

## Some intuition from solid state

You could look at Itzykson and Zuber, p. 108 and following for some details on this.

- Imagine for the moment that you have a discretized glob of jelly or a solid state lattice. Each particle in the jelly or ion on the lattice will have its own coordinate and its own momentum, and, in addition, these particles or ions will interact with one another in some fashion.
- How would you quantize this system?

You would introduce a  $q_i$  and  $p_i$ ,  $i = 1, N$  ( $N =$  very large) for each particle, and some potential  $\sum_{ij} V(x_i, x_j)$  and require

$$[q_i, p_j] = i\hbar\delta_{ij}. \quad (40)$$

Here,  $i$  and  $j$  denote the location of the ions within the lattice, while  $q_i$  is a displacement coordinate of the ion relative to its central location, and  $p_i$  is the conjugate momentum for this displacement coordinate.

- If you carried this through, then you would find it easiest to deduce the excitations of this system by defining creation and annihilation operators for (in the lattice example) phonon excitations on the lattice.

You would have a “vacuum” state  $|0\rangle$  in which all the ions simply had their “zero-point” energy (like a single harmonic oscillator) and then there would be creation operators  $a_{\vec{k}}^\dagger$  that would excite the lattice **as a whole** to contain a phonon described by momentum  $\vec{k}$ :  $|\vec{k}\rangle = a_{\vec{k}}^\dagger|0\rangle$ .

This state would be a state describing a coherent “wave-like” motion of the lattice as a whole. The precise energy of the phonon state would be determined by such things as the restoring force keeping each ion at its lattice site location and on the potential describing interactions between different ions on the lattice.

- Now imagine going to the continuum limit of the ion lattice, which would be sort of like a jelly. At every  $(\vec{x}, t)$  location there would be a particle, whose coordinate location we could denote by  $\phi(\vec{x}, t)$ . This latter coordinate location can be thought of as the coordinate describing *the displacement* of the ion relative to its central location  $(\vec{x}, t)$ . It is this latter “displacement” coordinate that is the one that should be quantized (just like the simple harmonic oscillator coordinate is really a displacement coordinate).

We would then want to quantize in a continuum sort of way by defining the momentum  $\pi(\vec{x}, t)$  of this particle (actually  $\pi$  should be a momentum density) and require

$$[\phi(\vec{x}, t), \pi(\vec{x}', t)] = i\delta^3(\vec{x} - \vec{x}') \quad (41)$$

(I am going to start setting  $\hbar = 1$ .)

Note that the  $\delta$  function makes sense in that if we integrate over a small volume in  $\vec{x}'$  and think of  $\phi(\vec{x})$  as the coordinate of another small volume, and think of each little volume as being a pseudo particle (pp), then the above would be equivalent to

$$[\text{coordinate of pp centered at } \vec{x}, \text{conj. mom of pp centered at } \vec{x}'] = i\delta_{\vec{x}\vec{x}'}. \quad (42)$$

If there was more than one degree of freedom at each  $(\vec{x}, t)$  (such as spin degrees of freedom or the like), we would attach some appropriate index to  $\phi$  and  $\pi$  and require

$$[\phi_\alpha(\vec{x}, t), \pi_\beta(\vec{x}', t)] = i\delta_{\alpha\beta}\delta^3(\vec{x} - \vec{x}') \quad (43)$$

## Field Theory

- In field theory, the coordinates are no longer the coordinates of some ion or jelly component, but rather coordinates in an abstract sense, i.e. really “fields” that appear in a Lagrangian density.

But, we will quantize in exactly the same way as sketched above.

This approach generalizes the classical mechanics of a system of particles, and its quantization, to a continuous system, i.e. to fields ala the treatment in Goldstein's *Classical Mechanics*.

- One introduces a Lagrangian (density) from which the field equations follow via Hamilton's principle.
- One introduces momenta (density) operators conjugate to the fields and imposes canonical commutation relations between the fields and these conjugate momenta.
- Everything follows from the Lagrangian density,  $\mathcal{L}$ . In particular, all conservation laws follow from symmetries of  $\mathcal{L}$ .

- This approach will give exactly the results we have just described in the E&M field case.
- We will use manifestly relativistic covariant notation.
- We will restrict to theories which can be derived by means of a variational principle from the action

$$S(\Omega) = \int_{\Omega} d^4x \mathcal{L}(\phi_r, \phi_{r,\alpha}), \quad (44)$$

where  $r$  is some index (like components of  $\vec{A}$ ) often related to spin and  $\phi_{r,\alpha}$  is the derivative of the field  $\phi_r$  as defined by our Lorentz/Metric conventions given earlier.

**Note:** A Lagrangian which only depends on the fields and their derivatives is not the most general choice, but all cases of interest in this course fall into this category.

- The fields  $\phi_r$  may be real or complex.

In the case that they are complex, the real and imaginary components (or equivalently  $\phi$  and  $\phi^*$ ) are to be treated as independent objects.

- Postulate that the equations of motion (eom's) follow from variational principles as follows. Let

$$\phi_r(x) \rightarrow \phi_r(x) + \delta\phi_r(x) \quad (45)$$

with  $\delta\phi$  vanishing on the surface  $\Gamma(\Omega)$  bounding the region  $\Omega$ .

- Require  $\delta S(\Omega) = 0$  under this variation.
- This leads to (dropping  $r$  index for the moment)

$$\begin{aligned} \delta S(\Omega) &= \int d^4x \left[ \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \delta\phi_{,\alpha} \right] \\ &= \int d^4x \left[ \frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \right) \right] \delta\phi + \int d^4x \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \delta\phi \right) \end{aligned} \quad (46)$$

where the last line follows from partial integration and  $\delta\phi_{,\alpha} = \frac{\partial}{\partial x^\alpha} \delta\phi$ .

- The last term in the last line is zero after using Gauss's theorem in 4d to convert the volume integration to a surface integration and using the fact that  $\delta\phi = 0$  on the surface.



- Then, setting  $\delta S = 0$  and using the fact the  $\delta\phi$  at each point within the volume is independent of other points leads to

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \right) = 0 \quad (47)$$

- In order to get the conjugate momenta right, we temporarily discretize space and time by dividing space up into small cells of equal volume  $\delta\vec{x}_i$ , labeled by index  $i = 1, 2, \dots$

We approximate the values of the fields within each cell by the value at the center of the cell.

Think  $q_i(t) \equiv \phi(i, t) \equiv \phi(\vec{x}_i, t)$ .

- Write the Lagrangian of the system of cells as

$$L(t) = \sum_i \delta\vec{x}_i \mathcal{L}_i(\phi(i, t), \dot{\phi}(i, t), \phi(i', t)) \quad (48)$$

where  $i'$  denotes the label of neighboring cells.

- The conjugate momenta to the  $q_i$  are

$$p_i(t) = \frac{\partial L}{\partial \dot{q}_i} \equiv \frac{\partial L}{\partial \dot{\phi}(i, t)} \equiv \pi(i, t) \delta \vec{x}_i. \quad (49)$$

- The Hamiltonian of the discrete system is then given by

$$H = \sum_i p_i \dot{q}_i - L = \sum_i \delta \vec{x}_i \left[ \pi(i, t) \dot{\phi}(i, t) - \mathcal{L}_i \right]. \quad (50)$$

- In the continuum limit (using notation  $x = (\vec{x}, t)$ ), we define  $\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}$ , and  $L(t) \rightarrow \int d^3 \vec{x} \mathcal{L}(\phi, \phi, \alpha)$  and  $\mathcal{H}(x) = \pi(x) \dot{\phi}(x) - \mathcal{L}(\phi, \phi, \alpha)$ .
- Since  $\mathcal{L}$  does not depend explicitly on time,  $H$  will be constant in time (as we shall later prove).

### The path integral motivation

- Treating a field as a coordinate probably seems quite adhoc to you. However, the path integral approach provides a very obvious motivation.

- I hope you are all at least vaguely familiar with the fact that normal quantum mechanics is equivalent to a path integral formulation in which one writes down an action (based on a Lagrangian) involving the coordinates and momenta of the particles involved, and then integrates  $e^{iS}$  over all possible “paths” in the coordinate and momentum configuration space:

$$\int [\mathcal{D}q(t)][\mathcal{D}p(t)] e^{iS(q,p)}. \quad (51)$$

- The classical path is that for which the action is at an extremum,  $\delta S = 0$ , but in QM one allows for all possible paths with coherent (amplitude level) weighting determined by  $e^{iS}$ .
- One can show that this formulation of QM is completely equivalent to the standard  $[q, p] = i$  formulation. This equivalence is simply a mathematical identity.
- So, if we now go to field theory, where “God” gives you a Lagrangian in terms of some fields, the simple analogy would be to quantize the theory by going to the path integral of  $e^{iS}$ , where  $S$  is computed from  $\mathcal{L}$  as described above.

- The path integral would now be an integral over all possible values of the field and its conjugate momentum which are now continuous functions of the parameters  $\vec{x}, t$ :

$$\int [\mathcal{D}\phi(\vec{x}, t)][\mathcal{D}\pi(\vec{x}, t)] e^{iS(\phi, \pi)}. \quad (52)$$

- The same mathematical equivalence would then apply. This path integral approach to treating this field  $\mathcal{L}$  quantum mechanically can be shown to be equivalent to our field quantization conditions where the field at a given space time location is treated as a coordinate with non-zero commutator with its conjugate momentum

### The Klein-Gordon example

- Real scalar field  $\phi(x)$ .
- To derive the Klein-Gordon equation that you may have seen from relativistic quantum mechanics you proceed just as you did for non-relativistic QM:

1. Write down the relationship between energy and momentum, in the relativistic case:

$$E^2 - \vec{p}^2 - \mu^2 = 0. \quad (53)$$

Here  $\mu$  is a constant with  $mass = (length)^{-1}$  dimension that we identify with the particle's mass.

2. Replace  $E \rightarrow i\frac{\partial}{\partial t}$  and  $\vec{p} \rightarrow -i\vec{\nabla}$  (we are using  $\hbar = c = 1$  notation) and operate the resulting differential equation on the one-particle wave function, which we call  $\phi$ . The result is:

$$\left(-\frac{\partial^2}{\partial t^2} + \vec{\nabla}^2 - \mu^2\right)\phi(\vec{x}, t) = -(\square + \mu^2)\phi(\vec{x}, t) = 0. \quad (54)$$

This is the KG equation.

- The Lagrangian that gives this equation of motion is

$$\mathcal{L} = \frac{1}{2}(\phi_{,\alpha}\phi^{,\alpha} - \mu^2\phi^2). \quad (55)$$

**Proof:** We first must compute, using dummy indices  $\beta, \gamma$  and metric tensor

for the first term in  $\mathcal{L}$ ,

$$\begin{aligned}
 \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} &= \frac{\partial}{\partial \phi_{,\alpha}} \left( \frac{1}{2} g^{\beta\gamma} \phi_{,\beta} \phi_{,\gamma} \right) \\
 &= \frac{1}{2} g^{\beta\gamma} [\delta_{\beta\alpha} \phi_{,\gamma} + \phi_{,\beta} \delta_{\gamma\alpha}] \\
 &= \frac{1}{2} [g^{\alpha\gamma} \phi_{,\gamma} + g^{\beta\alpha} \phi_{,\beta}] \\
 &= \phi_{,\alpha}.
 \end{aligned}$$

Thus, we have

$$\begin{aligned}
 \frac{\partial \mathcal{L}}{\partial \phi} &= -\mu^2 \phi, \\
 \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} &= \phi_{,\alpha} = \frac{\partial \phi}{\partial x_\alpha}, \\
 \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \right) &= \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \phi}{\partial x_\alpha} \right) = \square \phi \\
 \Rightarrow \\
 \frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \right) &= 0 \rightarrow -(\square + \mu^2) \phi = 0.
 \end{aligned}$$

- The EOM,  $(\square + \mu^2)\phi(x) = 0$ , being the Klein-Gordon equation  $\Rightarrow$  the quantized degrees of freedom that will be associated with it will turn out to be spinless neutral bosons.

Note that the  $\mathcal{L}$  form is quite abstract until we make this connection.

- **Note:** The Lagrangian density is not completely determined by the equations of motion. The standard Euler-Lagrange form of the eom will be unchanged if we add to the Lagrangian density a total four derivative  $\Delta\mathcal{L} = \partial_\beta \Lambda^\beta$  provided  $\Lambda^\beta$  has the form  $\Lambda^\beta = f(\phi)\phi^{,\beta}$ , so that the  $\beta$  index is supplied by the field derivative, but there are no further field derivatives. In this case, we have

$$\frac{\partial \Delta\mathcal{L}}{\partial \phi} = \partial_\alpha \frac{\partial \Lambda^\alpha}{\partial \phi} = \partial_\alpha \left( \frac{\partial f}{\partial \phi} \phi^{,\alpha} \right) \quad (56)$$

and

$$\frac{\partial}{\partial x^\alpha} \left( \frac{\partial \Delta\mathcal{L}}{\partial \phi_{,\alpha}} \right) = \frac{\partial}{\partial x^\alpha} \left( \partial_\beta \frac{\partial \Lambda^\beta}{\partial \phi_{,\alpha}} \right) = \frac{\partial}{\partial x^\alpha} \left( \partial_\beta f g^{\alpha\beta} \right) = \frac{\partial}{\partial x^\alpha} \left( \partial^\alpha f \right) = \frac{\partial}{\partial x^\alpha} \left( \frac{\partial f}{\partial \phi} \phi^{,\alpha} \right) \quad (57)$$

and we see that  $\frac{\partial \Delta\mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \Delta\mathcal{L}}{\partial \phi_{,\alpha}} \right) = 0$ . If we wanted to consider a more complicated form of  $\Lambda^\beta$  such that  $f = f(\phi, A)$ , with  $A = \phi_{,\gamma}\phi^{,\gamma}$ , we would have to go back and rederive the equations of motion from  $\delta S = 0$ . They would be more complicated than the usual due to the fact the  $\Delta\mathcal{L}$  depends upon 2nd derivatives of  $\phi$  (which we did not allow for in our derivation) in a rather complicated way. The theorem would still apply in that the new equations of motion obtained from  $\delta S = 0$  allowing for 2nd derivatives in the full  $\mathcal{L}$  would be left unaltered by the addition of  $\Delta\mathcal{L}$ .

- **Conjugate momentum:**  $\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} = \dot{\phi}(x)$  by direct computation.

- The Hamiltonian density computation.

$$\begin{aligned}\mathcal{H} &= \pi\dot{\phi} - \mathcal{L} \\ &= \pi^2 - \frac{1}{2} \left[ \pi^2(x) - (\vec{\nabla}\phi)^2 - \mu^2\phi^2 \right] \\ &= \frac{1}{2} \left[ \pi^2(x) + (\vec{\nabla}\phi)^2 + \mu^2\phi^2 \right].\end{aligned}$$

## (2nd) Quantization

- Impose “usual” canonical commutation relations:

$$[\phi(j, t), p(j', t)] = [\phi(j, t), \pi(j', t)\delta\vec{x}_{j'}] = i\delta_{jj'}\delta\vec{x}_j, \quad (58)$$

with other commutators zero.

In the continuum limit, you can think of dividing by  $\delta\vec{x}_j$  to obtain

$$[\phi(\vec{x}, t), \pi(\vec{x}', t)] = i\delta^3(\vec{x} - \vec{x}') \quad (59)$$



along with other commutators = 0.

Note that the commutation relations are at equal time, as always is the case for standard quantization.

- In the Klein-Gordon case, the important non-zero commutator reduces to

$$[\phi(\vec{x}, t), \dot{\phi}(\vec{x}', t)] = i\delta^3(\vec{x} - \vec{x}') \quad (60)$$

## Symmetries and Conservation Laws

- Heisenberg eom  $\Rightarrow i\frac{dO(t)}{dt} = [O(t), H] = 0$  if  $[O, H] = 0$ .
- $[O, H] = 0$  generally derives from invariance properties under a group of transformations.

e.g. translational and rotational invariance lead to conservation of linear and angular momentum, respectively.

Such transformations lead to equivalent descriptions of the system; in the above case the descriptions are the same in two Lorentz frames related by a translation or a rotation

- Quantum mechanically, two such descriptions must be related by a unitary transformation  $U$

$$|\Psi\rangle \rightarrow |\Psi'\rangle = U|\Psi\rangle, \quad O \rightarrow O' = UOU^\dagger. \quad (61)$$

The above relations imply:

1. Operator equations are covariant (i.e. take the same form in terms of original or transformed operators).

In particular, this is true of the commutation relations of the fields and of the equations of motion. For example, consider the commutation relation. Let us compute the commutator in the prime system to see if we get the same result as in the unprimed system. We have (using  $x = (t, \vec{x}), x' = (t, \vec{x}')$ )

$$\begin{aligned} [\phi'(x), \pi'(x')] &= [\phi'(x)\pi'(x') - \pi'(x')\phi'(x)] \\ &= [U\phi(x)U^\dagger U\pi(x')U^\dagger - U\pi(x')U^\dagger U\phi(x)U^\dagger] \\ &= U[\phi(x), \pi(x')]U^\dagger \\ &= Ui\delta^3(\vec{x} - \vec{x}')U^\dagger = i\delta^3(\vec{x} - \vec{x}'), \end{aligned} \quad (62)$$

using  $U^\dagger U = 1$  (unitarity). Here, it was important to note that  $U$

is some complicated operator that contains creation and annihilation operators but should (except in special cases, such as rotation and translation that also change the coordinates) commute with  $\delta^3(\vec{x} - \vec{x}')$ . In the special cases, the coordinate arguments are also shifted to new coordinate arguments.

Another example is that Maxwell's equations will take the same form in two different frames.

2. Amplitudes and, hence, observable predictions are invariant under the transformation. An example here is to consider a typical expectation value for which we have:

$$\langle \Psi' | \mathcal{O}' | \Psi' \rangle = (\langle \Psi | U^\dagger) (U \mathcal{O} U^\dagger) (U | \Psi \rangle) = \langle \Psi | \mathcal{O} | \Psi \rangle \quad (63)$$

using the unitarity of  $U$ .

- For continuous transformations, we can write

$$U = e^{i\alpha T} \stackrel{\alpha \rightarrow 0}{\sim} 1 + i\alpha T, \quad \text{with } T = T^\dagger \quad (64)$$

( $T$  is called the generator of the transformation  $U$ .) for which

$$O' = O + \delta O = (1 + i\alpha T) O (1 - i\alpha T), \quad \Rightarrow \delta O \simeq i\alpha [T, O] \quad (65)$$

If the theory is invariant under  $U$ , then  $H$  will be invariant, i.e.  $\delta H = 0$ . Plugging  $O = H$  into the above equation implies  $[T, H] = 0$ , i.e.  $T$  is a constant of motion.

- For a field theory derived from a  $\mathcal{L}$ , conserved quantities can be constructed from the invariance of  $\mathcal{L}$  under symmetry transformations.

The procedure is known under the name “Noether’s Theorem”.

### The Noether approach to obtaining conserved quantities

- Will show that an invariance of  $\mathcal{L}$  under a symmetry transformation will always lead to an equation of the form

$$\frac{\partial f^\alpha}{\partial x^\alpha} = 0 \tag{66}$$

where the  $f^\alpha$  are functions of the field operators and their derivatives.

- Then, as you well know, we can define the spatial integrals

$$F^\alpha(t) = \int d^3\vec{x} f^\alpha(\vec{x}, t) \tag{67}$$

and use Eq. (66) to derive

$$\frac{dF^0(t)}{dt} = - \int d^3\vec{x} \sum_{j=1}^3 \frac{\partial}{\partial x^j} f^j(\vec{x}, t) = 0 \quad (68)$$

where the last equality comes from converting to a surface integral using Gauss's theorem and from assuming the fields (and hence the  $f^j$ ) vanish sufficiently fast at infinity.

In short,  $F^0$  is a conserved quantity and one can construct the corresponding unitary operator for the transformation by setting  $T = F^0$ .

So, now let us prove this theorem.

- Suppose that  $\mathcal{L}$  is invariant under  $\phi(x) \rightarrow \phi'(x) = \phi(x) + \delta\phi(x)$ .
- However, we may always use the chain rule to compute  $\delta\mathcal{L}$  under any transformation:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\phi_{,\alpha}}\delta\phi_{,\alpha} = \frac{\partial}{\partial x^\alpha} \left( \frac{\partial\mathcal{L}}{\partial\phi_{,\alpha}}\delta\phi \right) \quad (69)$$

where the last equality follows from the eom for  $\phi$ :

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\alpha} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \right) = 0 \quad (70)$$

given earlier. Since, by assumption for the particular transformation we have  $\delta \mathcal{L} = 0$ , we see that Eq. (69) implies that

$$f^\alpha = \frac{\partial \mathcal{L}}{\partial \phi_{,\alpha}} \delta \phi \quad (71)$$

is a conserved current and the constant of motion is

$$F^0 = \int d^3 \vec{x} \pi(x) \delta \phi(x). \quad (72)$$

- Thus, for a given  $\mathcal{L}$ , all that remains is to determine the field transformations for which  $\delta \mathcal{L} = 0$ . For each such transformation, there will be a conserved current and a constant of motion.

The simplest example: complex scalar field

- A particularly important and yet simple example is provided by

$$\mathcal{L} = \phi^\dagger_{,\alpha} \phi^{,\alpha} - \mu^2 \phi^\dagger \phi \quad (73)$$

where  $\phi$  is a complex field (vs. the real field case considered earlier). (The rationale for the different normalization compared to the earlier real-field case will eventually be explained.)

- The above  $\mathcal{L}$  is invariant under

$$\phi' = e^{i\epsilon} \phi \simeq (1 + i\epsilon) \phi \quad \phi'^\dagger = e^{-i\epsilon} \phi^\dagger \simeq (1 - i\epsilon) \phi^\dagger, \quad (74)$$

where  $\epsilon$  is a real parameter and we are taking  $\epsilon$  to be infinitesimal in size.

Thus, we have an invariance with  $\delta\phi = i\epsilon\phi$  and  $\delta\phi^\dagger = -i\epsilon\phi^\dagger$ .

- The corresponding conserved quantity is

$$F^0 = i\epsilon \int d^3\vec{x} [\pi(x)\phi(x) - \pi^\dagger(x)\phi^\dagger(x)] \quad (75)$$

where  $\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}(x)} = \dot{\phi}^\dagger(x)$  and  $\pi^\dagger(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}^\dagger(x)} = \dot{\phi}(x)$ .

- If  $F^0$  is conserved then so is

$$Q = -iq \int d^3\vec{x} [\pi(x)\phi(x) - \pi^\dagger(x)\phi^\dagger(x)] \quad (76)$$

as obtained using  $\epsilon = -q$ . Here,  $\pm q$  will turn out to be the electric charges of the particle and antiparticle associated with the field  $\phi$ .

- We can easily show that  $Q$  generates the symmetry transformation by computing  $\phi'(x) = e^{iQ}\phi e^{-iQ}$  which, in the infinitesimal limit (here the limit of small  $q$ ), reduces to

$$\begin{aligned} \delta\phi = i[Q, \phi(x)] &= i \left[ -iq \int d^3\vec{x}' [\pi(x'), \phi(x)] \phi(x') \right] \\ &= q \int d^3\vec{x}' [-i\delta^3(\vec{x} - \vec{x}')] \phi(x') \\ &= -iq\phi(x) \end{aligned} \quad (77)$$

which is correct for the  $\epsilon = -q$  identification made earlier. In the above, we used the fact that  $\phi$  only has non-zero commutator with its corresponding  $\pi(x)$  (remembering that  $\phi$  and  $\phi^\dagger$  are independent fields in the complex field case).



- $[Q, \phi] = -q\phi$  tells us something else as well. Namely, suppose that  $|Q'\rangle$  is an eigenstate of the operator  $Q$  with eigenvalue  $Q'$ . Then, we find

$$Q\phi|Q'\rangle = [Q, \phi]|Q'\rangle + \phi Q|Q'\rangle = (Q' - q)\phi|Q'\rangle \quad (78)$$

which in words says that operating with  $\phi$  on  $|Q'\rangle$  has reduced the charge of the state by one unit of  $q$ . This is because  $\phi$  contains an annihilation operator for the particle with charge  $q$  and a creation operator for the antiparticle with charge  $-q$ , as we shall later verify when we return to 2nd quantizing the field  $\phi$ .

- **Note:** If the field  $\phi$  is real, it will not be possible to define a charge for the field.
- The above type of symmetry is called a global (i.e.  $x$ -independent) phase symmetry or a gauge invariance of the first kind.
- Possible ambiguities associated with the ordering of the operators in the expression for  $Q$  will be resolved, once we have carried through the 2nd quantization procedure, by requiring  $Q|0\rangle = 0$  for the vacuum state in which no particles (associated with the field  $\phi$ ) are present.

## Energy, Momentum and Angular Momentum

- Conservation of energy and momentum and of angular momentum follows from the invariance of  $\mathcal{L}$  under translations and rotations.

Since these transformations form a continuous group, we need only consider infinitesimal transformations.

- However, there is an extra piece in the analysis that did not appear for the phase rotations just considered.
- The transformations are defined by:

$$x'_\alpha \equiv x_\alpha + \delta x_\alpha = x_\alpha + \epsilon_{\alpha\beta} x^\beta + \delta_\alpha, \quad (79)$$

where  $\delta_\alpha$  is the infinitesimal displacement and  $\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha}$  is an infinitesimal antisymmetric tensor (as required to ensure that  $x_\alpha x^\alpha$  is invariant when  $\delta_\alpha = 0$ ) specifying an infinitesimal rotation.

- The above transformation will induce (and here we need to bring back the possible spin indices on the field)

$$\phi'_r(x') = \phi_r(x) + \frac{1}{2}\epsilon_{\alpha\beta}S_{rs}^{\alpha\beta}\phi_s(x). \quad (80)$$

- It is important to keep in mind that  $x'$  and  $x$  label the same physical point; it is just that this physical point is specified by a different set of coordinate values in the two different frames.
- The  $S_{rs}^{\alpha\beta}$  depend upon the “representation” or nature of the spin carried by the field. If the field in question is the vector field  $A_\alpha(x)$ , then the  $S$ 's would simply correspond to the usual rules for transformation of a vector field. (Do you know for sure what these look like in the infinitesimal limit?)
- Invariance under the transformations above means that  $\mathcal{L}$ , when expressed in terms of the  $x'$  and  $\phi'(x')$  **must have the same functional form** as when expressed in terms of  $x$  and  $\phi(x)$ . If this is the case, then the field equations will be covariant, i.e. they will have the same form when expressed in terms of either the original or the transformed coordinates and fields.

An example in the real scalar field case is that

$$\mathcal{L}(\phi'(x'), \phi'_{,\alpha}(x')) = \frac{1}{2} \left[ \phi'_{,\alpha}(x') \phi'^{\alpha}(x') - \mu^2 \phi'(x') \phi'(x') \right], \quad (81)$$

which would obviously lead to the same form for the equations of motion as found in the original frame where

$$\mathcal{L}(\phi(x), \phi_{,\alpha}(x)) = \frac{1}{2} \left[ \phi_{,\alpha}(x) \phi^{\alpha}(x) - \mu^2 \phi(x) \phi(x) \right]; \quad (82)$$

it would just be that everything had a prime in the primed frame case.

- Further, since  $\mathcal{L}$  should be a scalar, it should have the same value at a given physical point, regardless of what frame is used to describe the physical point: i.e.

$$\mathcal{L}(\phi_r(x), \phi_{r,\alpha}(x)) = \mathcal{L}(\phi'_r(x'), \phi'_{r,\alpha}(x')). \quad (83)$$

For the scalar field Lagrangian, this is obviously the case since  $\phi$  is a scalar field for which  $\phi'(x') = \phi(x)$  (the field should take the same value at the same physical point). Thus, indeed  $\mathcal{L}(\phi'(x'), \dots) = \mathcal{L}(\phi(x), \dots)$ .

**Note:** We are implicitly employing the *passive* point of view (i.e. we change to a different coordinate system and that is all) in this treatment. Other treatments employ an *active* point of view and the derivations look somewhat different.

- To return to Noether's theorem, let us examine the consequences of Eq. (83) obtained by expanding the right-hand side in terms of the unprimed coordinates and fields by means of using the transformations (79) and (80).
  - We again define  $\delta\phi_r(x) \equiv \phi'_r(x) - \phi_r(x)$  as the variation of  $\phi_r$  with the argument unchanged.
  - In addition we will need  $\delta_T\phi_r(x) \equiv \phi'_r(x') - \phi_r(x)$ , i.e. the variation including the change in the argument.
  - We then have:

$$\begin{aligned}
 \delta_T\phi_r(x) &= [\phi'_r(x') - \phi_r(x')] + [\phi_r(x') - \phi_r(x)] \\
 &= \delta\phi_r(x') + \frac{\partial\phi_r}{\partial x_\beta}\delta x_\beta \\
 &\simeq \delta\phi_r(x) + \frac{\partial\phi_r}{\partial x_\beta}\delta x_\beta
 \end{aligned} \tag{84}$$

where the last approximation neglects doubly-small terms.

– Similarly, we write

$$\begin{aligned}
 0 &= \mathcal{L}(\phi'_r(x'), \phi'_{r,\alpha}(x')) - \mathcal{L}(\phi_r(x), \phi_{r,\alpha}(x)) \\
 &= \delta\mathcal{L} + \frac{\partial\mathcal{L}}{\partial x^\alpha} \delta x^\alpha,
 \end{aligned} \tag{85}$$

where  $\delta\mathcal{L} = \mathcal{L}(\phi'(x), \dots) - \mathcal{L}(\phi(x), \dots)$ , i.e. only  $x$  appears when we neglect doubly-small changes.

– For  $\delta\mathcal{L}$  we proceed as we did in the phase case and write

$$\begin{aligned}
 \delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_r} \delta\phi_r + \frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}} \delta\phi_{r,\alpha} \\
 &= \frac{\partial}{\partial x^\alpha} \left[ \frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}} \delta\phi_r \right] \\
 &= \frac{\partial}{\partial x^\alpha} \left[ \frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}} \left( \delta_T\phi_r - \frac{\partial\phi_r}{\partial x_\beta} \delta x_\beta \right) \right].
 \end{aligned} \tag{86}$$

where the first equality just follows from the good old eom's and for the 2nd equality we use Eq. (84).

– From this result, it is clear that a suitable choice of the conserved current

is

$$f^\alpha \equiv \frac{\partial \mathcal{L}}{\partial \phi_{r,\alpha}} \delta_T \phi_r - \mathcal{T}^{\alpha\beta} \delta x_\beta \quad (87)$$

with

$$\mathcal{T}^{\alpha\beta} \equiv \frac{\partial \mathcal{L}}{\partial \phi_{r,\alpha}} \frac{\partial \phi_r}{\partial x_\beta} - \mathcal{L} g^{\alpha\beta}, \quad (88)$$

which is the standard energy-momentum tensor.

In more detail, using Eq. (88) in Eq. (87) we have

$$\begin{aligned} \frac{\partial f^\alpha}{\partial x^\alpha} &= \frac{\partial}{\partial x^\alpha} \left[ \frac{\partial \mathcal{L}}{\partial \phi_{r,\alpha}} \left( \delta_T \phi_r - \frac{\partial \phi_r}{\partial x_\beta} \delta x_\beta \right) \right] - \frac{\partial}{\partial x^\alpha} (-\mathcal{L} g^{\alpha\beta} \delta x_\beta) \\ &= \delta \mathcal{L} + \frac{\partial \mathcal{L}}{\partial x^\alpha} \delta x^\alpha \\ &= 0 \end{aligned} \quad (89)$$

where the last 0 comes from Eq. (85). Note that  $\delta x_\beta$  is some fixed shift in  $x$  defined by our original symmetry statement, implying that  $\frac{\partial}{\partial x^\alpha}$  does not operate on it.

- Let us now apply this to a **translation**:

$$\epsilon_{\alpha\beta} = 0 \Rightarrow \delta x_\beta = \delta_\beta \quad \text{and} \quad \phi'_r(x') = \phi_r(x),$$

$$\text{which in turn } \Rightarrow \delta_T \phi_r(x) = 0, \quad (90)$$

which in turn implies (see Eq. (87)) that

$$f^\alpha = -\mathcal{T}^{\alpha\beta} \delta_\beta. \quad (91)$$

Since each of the 4  $\delta_\beta$ 's are independent of one another, this really implies that there are four conserved currents and four conserved charges, the latter being the spatial integrals of  $\mathcal{T}^{00}$ ,  $\mathcal{T}^{01}$ ,  $\mathcal{T}^{02}$ , and  $\mathcal{T}^{03}$ , i.e. it is useful to define

$$P^\beta = \int d^3\vec{x} \mathcal{T}^{0\beta} = \int d^3\vec{x} \left[ \pi_r(x) \frac{\partial \phi_r(x)}{\partial x_\beta} - \mathcal{L} g^{0\beta} \right], \quad (92)$$

which is just the (operator corresponding to) the energy-momentum four-vector, with explicit components

$$P^0 = \int d^3\vec{x} [\pi_r(x) \dot{\phi}_r(x) - \mathcal{L}] = \int d^3\vec{x} \mathcal{H} = H \quad (93)$$



and

$$P^j = \int d^3\vec{x} \pi_r(x) \frac{\partial \phi_r(x)}{\partial x_j}. \quad (94)$$

These are the momentum operators expressed in terms of the field operators. This will be confirmed when we express these operators in terms of the number representation that follows from the 2nd quantization procedure.

**Homework 2:** At this point, you are ready to do Problems 2.2, 2.3 and 2.4 from Mandl-Shaw.

- Now consider the Noether current for a  $\mathcal{L}$  symmetric under **rotation**:

In this case, we have  $\delta x_\beta = \epsilon_{\beta\gamma} x^\gamma$  and  $\delta_T \phi_r(x) = \frac{1}{2} \epsilon_{\beta\gamma} S_{rs}^{\beta\gamma} \phi_s(x)$ , where we have used appropriate indices and dummy summation indices for what follows.

Plugging this into Eq. (87), i.e. into

$$f^\alpha \equiv \frac{\partial \mathcal{L}}{\partial \phi_{r,\alpha}} \delta_T \phi_r - \mathcal{T}^{\alpha\beta} \delta x_\beta \quad (95)$$

gives

$$\begin{aligned}
 f^\alpha &= \frac{\partial \mathcal{L}}{\partial \phi_{r,\alpha}} \frac{1}{2} \epsilon_{\beta\gamma} S_{rs}^{\beta\gamma} \phi_s(x) - \mathcal{T}^{\alpha\beta} \epsilon_{\beta\gamma} x^\gamma \\
 &= \frac{1}{2} \left[ \frac{\partial \mathcal{L}}{\partial \phi_{r,\alpha}} S_{rs}^{\beta\gamma} \phi_s(x) - \mathcal{T}^{\alpha\beta} x^\gamma + \mathcal{T}^{\alpha\gamma} x^\beta \right] \epsilon_{\beta\gamma} \quad (96)
 \end{aligned}$$

where the 2nd equality follows by using the antisymmetry of  $\epsilon_{\beta\gamma}$ .

Now, we note that the  $\epsilon_{\beta\gamma}$  are all independent, which gives us a whole set of six different conserved currents denoted as

$$\mathcal{M}^{\alpha\beta\gamma} \equiv \frac{\partial \mathcal{L}}{\partial \phi_{r,\alpha}} S_{rs}^{\beta\gamma} \phi_s(x) - \mathcal{T}^{\alpha\beta} x^\gamma + \mathcal{T}^{\alpha\gamma} x^\beta \quad (97)$$

yielding 6 conserved quantities for  $\alpha = 0$

$$\begin{aligned}
 M^{\beta\gamma} &= \int d^3 \vec{x} \mathcal{M}^{0\beta\gamma} \\
 &= \int d^3 \vec{x} \{ [x^\beta \mathcal{T}^{0\gamma} - x^\gamma \mathcal{T}^{0\beta}] + \pi_r(x) S_{rs}^{\beta\gamma} \phi_s(x) \} . \quad (98)
 \end{aligned}$$

So what are all these objects?

- For two spatial indices ( $i, j = 1, 2, 3$ ),  $M^{ij}$  is the angular momentum operator of the field ( $M^{12}$  being the  $z$ -component, etc.).
- For one spatial index  $i$  and one 0 index, we get the operator giving rise to boosts in the  $i$  direction.
- Within the expression of Eq. (98), the [...] stuff is the orbital angular momentum and the  $S$  term contains the portion of the total angular momentum coming from the intrinsic spin of the field.

### Reiteration of Basic Steps and Points

- Nature will give us a certain type of field and a certain Lagrangian density for this field.
  - We will (2nd) quantize this system by treating the field as if it were a quantum mechanical coordinate, using the Lagrangian density to define the momentum density conjugate to the field (just as we would for a jelly or dense lattice).
  - Indeed, **everything follows from  $\mathcal{L}$** .
- In particular, any conserved quantity must be associated with an invariance of  $\mathcal{L}$ .

If the theory is to be translationally invariant, then  $\mathcal{L}$  should be invariant under translations and the corresponding conserved quantity (the momentum operator) must be that obtained by applying the Noether construction to the translationally invariant  $\mathcal{L}$ .

If the theory is to have a conserved electric charge (or similarly for hypercharge and other such things),  $\mathcal{L}$  must be invariant under a global phase rotation (or some other abelian phase transformation), and the associated charge operator must come from the Noether construction using the phase transformation.

More complicated groups are also possible, where there is another (non-spin) index attached to the field that transforms in some non-trivial way under a group generated, for example, by a Lie algebra.

We will now turn to carrying out this procedure in detail for:

1. the real and complex scalar fields.
2. the Dirac field.
3. the electromagnetic field.

in that order.

# The Scalar Field

## Basics for a real scalar field

- For one particle wave equation, we began with  $E^2 = \mu^2 + \vec{p}^2$ , made replacements  $\vec{p} \rightarrow -i\vec{\nabla}$ ,  $E \rightarrow i\frac{\partial}{\partial t}$  and obtained the Klein-Gordon equation:

$$(\square + \mu^2)\phi(x) = 0. \quad (99)$$

- We recall that one of the “difficulties” of this equation was the presence of negative energy solution.

This difficulty is characteristic of single-particle wave equations. We will see that such difficulties do not arise in the 2nd quantization (i.e. field quantization) approach.

- We saw earlier that the KG equation is the equation of motion coming from

$$\mathcal{L} = \frac{1}{2}(\phi_{,\alpha}\phi^{,\alpha} - \mu^2\phi^2) \quad (100)$$

and that the field (momentum density) conjugate to  $\phi$  is

$$\pi(\mathbf{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}. \quad (101)$$

We quantize the system by turning  $\phi$  into a hermitian operator and requiring

$$[\phi(\vec{x}, t), \dot{\phi}(\vec{x}', t)] = i\delta^3(\vec{x} - \vec{x}') \quad [\phi(\vec{x}, t), \phi(\vec{x}', t)] = [\dot{\phi}(\vec{x}, t), \dot{\phi}(\vec{x}', t)] = 0. \quad (102)$$

To implement these quantization conditions we expand  $\phi$  with operator coefficients, just as already done for the  $\vec{A}$  field earlier, assuming for the moment periodic conditions in a box:

$$\phi(\mathbf{x}) = \phi^+(\mathbf{x}) + \phi^-(\mathbf{x}) \quad (103)$$

with

$$\phi^+(\mathbf{x}) = \sum_{\vec{k}} \left( \frac{1}{2\omega_{\vec{k}}V} \right)^{1/2} a(\vec{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}, \quad \phi^-(\mathbf{x}) = \sum_{\vec{k}} \left( \frac{1}{2\omega_{\vec{k}}V} \right)^{1/2} a^\dagger(\vec{k}) e^{+i\mathbf{k}\cdot\mathbf{x}} \quad (104)$$

Here,  $k^0 = \omega_{\vec{k}} = \sqrt{\mu^2 + \vec{k}^2}$ : i.e.  $k = (k^0, \vec{k})$  is the on-mass-shell (on-shell, for short) four momentum of a relativistic particle of mass  $\mu$ .

You should be asking yourself why the appropriate expansion functions are  $e^{\pm ik \cdot x}$ , where  $k \cdot x = \omega_{\vec{k}} t - \vec{k} \cdot \vec{x}$ . The answer is that we should expand in terms of the “plane wave” solutions of the equation of motion obeyed by  $\phi$ , i.e. the solutions of the Klein-Gordon equation that is determined by the Lagrangian. We can check that these plane waves are indeed solutions:

$$\begin{aligned} (\square + \mu^2)e^{-ik \cdot x} &= \left( \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + \mu^2 \right) e^{-i\omega_{\vec{k}} t + i\vec{k} \cdot \vec{x}} \\ &= \left( -\omega_{\vec{k}}^2 + \vec{k}^2 + \mu^2 \right) e^{-i\omega_{\vec{k}} t + i\vec{k} \cdot \vec{x}} \\ &= 0 \quad \text{provided} \quad \omega_{\vec{k}}^2 = \vec{k}^2 + \mu^2 \end{aligned}$$

i.e. provided  $\omega_{\vec{k}}$  is given by the appropriate relativistic form for a particle of mass  $\mu$ .

You should also ask yourself why it is that we have chosen  $\phi^-$  to be the Hermitian conjugate of  $\phi^+$ . This is because we have chosen to discuss the case of a *real* (before quantization) scalar field, the quantized version of which should be a Hermitian field.

- So, now that we have an appropriate expansion of the field  $\phi$  in terms of operator coefficients, we must determine the commutation relations of

the  $a$  and  $a^\dagger$  such that the 2nd quantization condition of Eq. (102) are satisfied.

The quantization conditions of Eq. (102) are achieved by requiring

$$[a(\vec{k}), a^\dagger(\vec{k}')] = \delta_{\vec{k}\vec{k}'}, \quad [a(\vec{k}), a(\vec{k}')] = [a^\dagger(\vec{k}), a^\dagger(\vec{k}')] = 0. \quad (105)$$

We should prove this as an *if and only if* statement. I will prove it in the easy direction below. You will prove it in the harder direction as homework.

**Homework 3:** MS problem 3.1 is to show that if  $[\phi(\vec{x}, t), \dot{\phi}(\vec{y}, t)] = i\delta^3(\vec{x} - \vec{y})$  (and others are 0) then the above commutators are as stated.

Here we show the reverse, which is actually easier. We check that if the  $a, a^\dagger$  commutators are as stated, then  $[\phi(\vec{x}, t), \dot{\phi}(\vec{y}, t)] = i\delta^3(\vec{x} - \vec{y})$ .

**Proof:**

$$[\phi(\vec{x}, t), \dot{\phi}(\vec{y}, t)] = \left[ \sum_{\vec{k}} \left( \frac{1}{2V\omega_{\vec{k}}} \right)^{1/2} (a(\vec{k})e^{-i\vec{k}\cdot\vec{x}} + a^\dagger(\vec{k})e^{+i\vec{k}\cdot\vec{x}}), \right. \\ \left. \sum_{\vec{p}} \left( \frac{1}{2V\omega_{\vec{p}}} \right)^{1/2} (-i\omega_{\vec{p}}a(\vec{p})e^{-i\vec{p}\cdot\vec{y}} + i\omega_{\vec{p}}a^\dagger(\vec{p})e^{+i\vec{p}\cdot\vec{y}}) \right]$$



$$\begin{aligned}
&= \sum_{\vec{k}, \vec{p}} \left( \frac{1}{2V\omega_{\vec{k}}} \right)^{1/2} \left( \frac{1}{2V\omega_{\vec{p}}} \right)^{1/2} \left[ i\omega_{\vec{p}} e^{i\vec{p}\cdot\vec{y} - i\vec{k}\cdot\vec{x}} [a(\vec{k}), a^\dagger(\vec{p})] \right. \\
&\quad \left. - i\omega_{\vec{p}} e^{-i\vec{p}\cdot\vec{y} + i\vec{k}\cdot\vec{x}} [a^\dagger(\vec{k}), a(\vec{p})] \right] \\
&= \frac{i}{V} \sum_{\vec{p}} e^{-i\vec{p}\cdot(\vec{y}-\vec{x})} \rightarrow \frac{i}{(2\pi)^3} \int d^3\vec{p} e^{-i\vec{p}\cdot(\vec{y}-\vec{x})} = i\delta^3(\vec{x} - \vec{y}) \quad (106)
\end{aligned}$$

where we used  $x^0 = y^0 = t$  for equal time commutators and for the 2nd term in the next to last line did a change of summation variables from  $\vec{k}, \vec{p} \rightarrow -\vec{k}, -\vec{p}$ , and of course used the fact that  $[a(\vec{k}), a^\dagger(\vec{p})] = -[a^\dagger(-\vec{k}), a(-\vec{p})] = \delta_{\vec{k}\vec{p}}$ , which also means that we can set  $\omega_{\vec{p}} = \omega_{\vec{k}}$  everywhere. When this is done in the exponential, the time dependence disappears, as it must. Of course, it is important to keep in mind that  $\omega_{\vec{k}}$  is an even function of  $\vec{k}$ . We have noted that the evaluation of the momentum sum is most easily done by going to the large volume limit in the standard way and using the usual integral representation of the Dirac delta function.

- **A note on normalization conventions**

In writing the operator expansion of  $\phi$  using the  $\frac{1}{\sqrt{2V\omega_{\vec{k}}}}$  normalization and in writing the  $a, a^\dagger$  commutator as  $[a(\vec{k}), a^\dagger(\vec{p})] = \delta_{\vec{k}\vec{p}}$ , we made two connected normalization choices. Obviously, we could make a change in the  $\phi$  expansion normalization convention provided we made a compensating change in the  $[a(\vec{k}), a^\dagger(\vec{p})]$  normalization, adjusting one relative to the other so that we maintain  $[\phi(\vec{x}, t), \dot{\phi}(\vec{y}, t)] = i\delta^3(\vec{x} - \vec{y})$ , which is our fundamental physics input assumption.

- In any case, the  $a, a^\dagger$  commutators are precisely the Harmonic oscillator commutators as already discussed for the  $\vec{A}$  field, and we know how to proceed. In particular, we establish a occupation number space with number operator

$$N(\vec{k}) = a^\dagger(\vec{k})a(\vec{k}) \quad (107)$$

with eigenvalues  $n(\vec{k}) = 0, 1, 2, \dots$ , with  $a(\vec{k})$  and  $a^\dagger(\vec{k})$  being the annihilation and creation operators of particles with four-momentum  $k$ .

- We define the vacuum state by  $a(\vec{k})|0\rangle = 0$ , which also implies  $\phi^+(x)|0\rangle = 0$ .
- We define a single particle state via  $|\vec{k}\rangle = a^\dagger(\vec{k})|0\rangle$ .

– We define two-particle states (correctly normalized) by

$$\begin{aligned} |\vec{k}, \vec{k}'\rangle &= a^\dagger(\vec{k})a^\dagger(\vec{k}')|0\rangle, \quad \vec{k} \neq \vec{k}' \\ |\vec{k}, \vec{k}\rangle &= \frac{1}{\sqrt{2}}[a^\dagger(\vec{k})]^2|0\rangle \end{aligned} \quad (108)$$

It is perhaps useful to perform the manipulations that demonstrate why there is a  $1/\sqrt{2}$  in the last equation. The point is that all states should have the same normalization. So, let us assume that  $\langle 0|0\rangle = 1$  for convenience. Then, we wish to check the normalization of  $|\vec{k}, \vec{k}\rangle$ :

$$\begin{aligned} \langle \vec{k}, \vec{k} | \vec{k}, \vec{k} \rangle &= \frac{1}{2} \langle 0 | a a a^\dagger a^\dagger | 0 \rangle \\ &= \frac{1}{2} \langle 0 | a (1 + a^\dagger a) a^\dagger | 0 \rangle \\ &= \frac{1}{2} \langle 0 | a a^\dagger | 0 \rangle + \frac{1}{2} \langle 0 | a a^\dagger a a^\dagger | 0 \rangle \\ &= \frac{1}{2} \langle 0 | (1 + a^\dagger a) | 0 \rangle + \frac{1}{2} \langle 0 | a a^\dagger (1 + a^\dagger a) | 0 \rangle \\ &= \frac{1}{2} \langle 0 | 0 \rangle + \frac{1}{2} \langle 0 | 0 \rangle = \langle 0 | 0 \rangle \end{aligned}$$

where we stopped writing  $\vec{k}$ , used  $[a, a^\dagger] = 1$  (for the same value of  $\vec{k}$ ) repeatedly, and used  $a|0\rangle = 0$  several times.

- We clearly have Bose statistics since the order of the  $a^\dagger$  operators for a multiparticle state does not matter ( $a^\dagger$  operators commute with one another) and since we can put as many particles into the same  $\vec{k}$  state as we like.

- We already know what the Hamiltonian and momentum operators are from Noether's theorem applied to the scalar-field  $\mathcal{L}$ :

$$\begin{aligned} H &= \int d^3\vec{x} \frac{1}{2} \left[ \dot{\phi}^2 + (\vec{\nabla}\phi)^2 + \mu^2\phi^2 \right] \\ \vec{P} &= - \int d^3\vec{x} \dot{\phi} \vec{\nabla}\phi. \end{aligned} \quad (109)$$

- We now just substitute the expansion form of  $\phi$  to see what we get. We find

$$H = \sum_{\vec{k}} \omega_{\vec{k}} \left[ N(\vec{k}) + \frac{1}{2} \right], \quad \vec{P} = \sum_{\vec{k}} \vec{k} N(\vec{k}). \quad (110)$$

### Prove this as part of homework 3.

These expressions are the “proof” that the  $a(\vec{k})$  and  $a^\dagger(\vec{k})$  operators are indeed the annihilation and creation operators for a relativistic spinless particle, in the sense that if it smells like a ... and tastes like a ....., then it is a .....

More explicitly, let us examine the one particle state,  $|\vec{k}\rangle = a^\dagger(\vec{k})|0\rangle$ . We have

$$\begin{aligned} N(\vec{k}')|\vec{k}\rangle &= a^\dagger(\vec{k}')a(\vec{k}')a^\dagger(\vec{k})|0\rangle \\ &= a^\dagger(\vec{k}')[\delta_{\vec{k}'\vec{k}} + a^\dagger(\vec{k})a(\vec{k}')]|0\rangle \\ &= \delta_{\vec{k}'\vec{k}}a^\dagger(\vec{k})|0\rangle + 0 \\ &= \delta_{\vec{k}'\vec{k}}|\vec{k}\rangle \end{aligned}$$

from which we find that

$$H|\vec{k}\rangle = \left[ \omega_{\vec{k}} + \sum_{\vec{k}'} \frac{1}{2} \omega_{\vec{k}'} \right] |\vec{k}\rangle \quad (111)$$

as compared to

$$H|0\rangle = \sum_{\vec{k}'} \frac{1}{2} \omega_{\vec{k}'} |0\rangle \quad (112)$$

which is to say that the 1-particle state has an energy that is  $\omega_{\vec{k}}$  larger than the energy of the vacuum state.

Note that the  $H$  operator has only positive energy excitations, relative to the vacuum state.

Similarly, we have

$$\vec{P}|\vec{k}\rangle = \vec{k}|\vec{k}\rangle \quad (113)$$

as compared to

$$\vec{P}|0\rangle = 0 \quad (114)$$

which is to say that the 1-particle state has a momentum that is  $\vec{k}$  as compared to the vacuum state having no momentum.

Thus, it really looks like the 1-particle state that we have defined does indeed have the energy and momentum of a relativistic 1-particle state. Again, it should be emphasized that once we had specified  $\mathcal{L}$  and the 2nd quantization condition, we had no freedom in how to compute the  $H$  and  $\vec{P}$  operators.

- Well, there is another normalization convention here.

When I wrote the Lagrangian for the real scalar field, I made a certain choice of the normalization. The Klein Gordon equation would result from the eom no matter what the overall normalization of  $\mathcal{L}$  is. The normalization convention was chosen so that the energy that comes out of Noether's theorem is the normal energy and not say  $917 \times E_{\vec{k}}$  or something of the sort.

Of course, we made long ago a historical choice for how to define the units of energy and momentum, and that is built into what we have done. A phrase that is sometimes used is to say that the normalization employed for  $\mathcal{L}$  is "canonical normalization". Whenever we write a Lagrangian, it is important to write it with canonical normalization for each degree of freedom (each particle after 2nd quantization).

Thus, when we come to the charged scalar field (complex scalar field) in a short while,  $\mathcal{L}$  will be chosen without a factor of  $1/2$  in front. This is because *each* of the two independent degrees of freedom (particle and antiparticle) should carry canonically normalized energy. The form that will be chosen (without the  $1/2$ ) gives exactly this result.

- We see above that the vacuum  $|0\rangle$  will have  $H|0\rangle = \sum_{\vec{k}'} \frac{1}{2} \omega_{\vec{k}'} |0\rangle$ , which is

infinite because we are summing an infinite number of zero-point harmonic oscillator energies.

But, this infinity is unmeasurable. We only see excitations relative to it. This is just like a big lattice. The lattice would have an enormous ground state energy from all the zero-point ion energies, but all we would care about are the energies of the phonon excitations relative to this big ground state energy.

- It is convenient to always write things in such a way that this infinity is thrown away. The formal name for doing so is “normal ordering” or always using “normal products”, denoted by  $N$ . By definition,  $N$  instructs us to put all the  $a^\dagger$  operators to the left of the  $a$  operators. For example:

$$N(a(\vec{k}_1)a(\vec{k}_2)a^\dagger(\vec{k}_3)) = a^\dagger(\vec{k}_3)a(\vec{k}_1)a(\vec{k}_2) \quad (115)$$

and

$$\begin{aligned} N[\phi(x)\phi(y)] &= N[(\phi^+(x) + \phi^-(x))(\phi^+(y) + \phi^-(y))] \\ &= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) \\ &\quad + \phi^-(y)\phi^+(x) + \phi^-(x)\phi^-(y), \end{aligned} \quad (116)$$



where the order of the factors was changed only in the 3rd term, so that all  $\phi^+$  operators (which contain the annihilation operator) are to the right of all  $\phi^-$  operators.

**Note**, some minus signs will creep into the definition of  $N$  when fermions are involved.

Sometimes  $N[\dots]$  is denoted by  $:\dots:$ , to avoid confusion with the number operator.

- Clearly, the vacuum expectation value of any normal product vanishes. Thus, it is convenient to define all observables as being normal ordered. In particular, starting from our original form of  $H$ ,

$$N[H] = \sum_{\vec{k}} \omega_{\vec{k}} \frac{1}{2} N[a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})] = \sum_{\vec{k}} \omega_{\vec{k}} a^\dagger(\vec{k})a(\vec{k}). \quad (117)$$

From now on, when we write  $H$ , we will be assuming that it is the normal-ordered version of  $H$  that is being written. Thus, for example, we will

have

$$H|\vec{k}\rangle = \left[ \sum_{\vec{k}'} \omega_{\vec{k}'} N(\vec{k}') \right] |\vec{k}\rangle = \omega_{\vec{k}} |\vec{k}\rangle. \quad (118)$$

using the same result,  $N(\vec{k}')|\vec{k}\rangle = \delta_{\vec{k}'\vec{k}}|\vec{k}\rangle$ , as before.

### Basics for a complex scalar field

- Write

$$\mathcal{L} =: (\phi^\dagger, {}_\alpha \phi, {}^\alpha - \mu^2 \phi^\dagger \phi) : \quad (119)$$

Treating the field and its adjoint as independent fields, leads to the KG equations

$$(\square + \mu^2)\phi(x) = 0, \quad (\square + \mu^2)\phi^\dagger(x) = 0. \quad (120)$$

- The fields conjugate to  $\phi$  and  $\phi^\dagger$  are  $\pi = \dot{\phi}^\dagger$  and  $\pi^\dagger = \dot{\phi}$ .

- The equal-time quantization conditions are then

$$\begin{aligned} [\phi(\vec{x}, t), \dot{\phi}^\dagger(\vec{x}', t)] &= i\delta^3(\vec{x} - \vec{x}') \\ [\phi(\vec{x}, t), \phi(\vec{x}', t)] &= [\phi(\vec{x}, t), \phi^\dagger(\vec{x}', t)] = [\dot{\phi}(\vec{x}, t), \dot{\phi}(\vec{x}', t)] \end{aligned}$$

$$= [\dot{\phi}(\vec{x}, t), \dot{\phi}^\dagger(\vec{x}', t)] = [\phi(\vec{x}, t), \dot{\phi}(\vec{x}', t)] = 0 \quad (121)$$

Note that the first, i.e. the non-zero, commutator is equivalent to the other commutator that must be non-zero, namely

$$[\phi^\dagger(\vec{x}, t), \dot{\phi}(\vec{x}', t)] = i\delta^3(\vec{x} - \vec{x}'). \quad (122)$$

- These commutation relations are solved in the operator basis by:

$$\phi(x) = \phi^+(x) + \phi^-(x) = \sum_{\vec{k}} \left( \frac{1}{2\omega_{\vec{k}}V} \right)^{1/2} \left[ a(\vec{k})e^{-ik \cdot x} + b^\dagger(\vec{k})e^{+ik \cdot x} \right] \quad (123)$$

with  $\phi^\dagger(x)$  being the hermitian conjugate of the above, provided

$$[a(\vec{k}), a^\dagger(\vec{k}')] = [b(\vec{k}), b^\dagger(\vec{k}')] = \delta_{\vec{k}\vec{k}'} \quad (124)$$

with all other commutators being zero.

- Once again, we can interpret (as we shall see),  $a(\vec{k})$  and  $b(\vec{k})$  as annihilation

operators and  $a^\dagger(\vec{k})$  and  $b^\dagger(\vec{k})$  as creation operators, but this time for two different types of particles.

The corresponding vacuum state is defined by

$$a(\vec{k})|0\rangle = b(\vec{k})|0\rangle = 0, \quad \text{all } \vec{k}. \quad (125)$$

The number operators for the two different types of particles are

$$N_a(\vec{k}) = a^\dagger(\vec{k})a(\vec{k}), \quad N_b(\vec{k}) = b^\dagger(\vec{k})b(\vec{k}). \quad (126)$$

- To understand how to interpret these two different types of particle operators we must evaluate the energy and momentum operators for this system.

One finds that the four-momentum operator takes the form (recall that  $H$  is now defined as including the normal-ordering prescription)

$$P^\alpha = (H, \vec{P}) = \sum_{\vec{k}} k^\alpha (N_a(\vec{k}) + N_b(\vec{k})) \quad (127)$$

where we should keep in mind that  $k^0 = \omega_{\vec{k}}$ . This result means that both  $a^\dagger$  and  $b^\dagger$  create particles from the vacuum state with the four momentum of a relativistic particle with mass  $\mu$ .

- So what is it that distinguishes the  $a$  from the  $b$  particles?

Let us look at charge. We obtained an expression for the charge operator from  $\mathcal{L}$  from its global phase invariance. Including normal ordering (which means that we throw away an infinite charge for the vacuum state and measure charges of other states relative to the vacuum charge) we have:

$$Q = -iq \int d^3\vec{x} : [\dot{\phi}^\dagger(x)\phi(x) - \dot{\phi}(x)\phi^\dagger(x)] : \quad (128)$$

which reduces to

$$Q = q \sum_{\vec{k}} [N_a(\vec{k}) - N_b(\vec{k})], \quad (129)$$

demonstrating that the  $b$  type particle is the antiparticle of the  $a$  type particle with exactly the same mass but opposite charge.

The corresponding charge-current density is given by

$$j^\alpha(x) = (\rho(x), \vec{j}(x)) = -iq : \left[ \frac{\partial \phi^\dagger}{\partial x_\alpha} \phi - \frac{\partial \phi}{\partial x_\alpha} \phi^\dagger \right] : \quad (130)$$

which obviously (use the fact that  $\phi$  and  $\phi^\dagger$  both obey the KG equation) satisfies

$$\frac{\partial j^\alpha(x)}{\partial x^\alpha} = 0. \quad (131)$$

### Notes:

- The energy operator is positive definite, unlike what happens in trying to interpret the single particle Klein-Gordon plane wave solutions.
- If one describes a charged particle using a relativistic quantized field theory, one inevitably finds that there is an antiparticle with opposite charge (and other such quantum numbers) but with exactly the same mass.
- If the field is real, the charge operator is zero, but a chargeless particle can still have an antiparticle.

This occurs if there are other charge-like properties of the particle. An example is what is called hypercharge. The neutral meson  $K^0$  with hypercharge  $Y = 1$  has an antiparticle  $\bar{K}^0$  with  $Y = -1$ .

At this point, problem 3.2 of Mandl-Shaw is assigned, as well as the extra problem in which you are asked to derive the expression for  $Q$  of Eq. (129).

## Propagators for the scalar field

- The first example of a propagator is the covariant commutator of two fields:  $[\phi(x), \phi(y)]$ , where  $x$  and  $y$  are two different coordinate locations.

It should be a covariant scalar since the theory is supposed to be covariant and since the fields are scalars. However, the equal time commutators might have broken this in some way, but they do not.

- We first note that

$$[\phi(x), \phi(y)] = [\phi^+(x), \phi^-(y)] + [\phi^-(x), \phi^+(y)] \quad (132)$$

since the  $+$  components commute with one another and the  $-$  components also commute with one another.

- For the 1st term we get

$$\begin{aligned}
[\phi^+(x), \phi^-(y)] &= \frac{1}{2V} \sum_{\vec{k}\vec{k}'} \frac{1}{\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}}} [a(\vec{k}), a^\dagger(\vec{k}')] e^{-ik \cdot x + ik' \cdot y} \\
&\xrightarrow[V \rightarrow \infty]{} \frac{1}{2(2\pi)^3} \int \frac{d^3\vec{k}}{\omega_{\vec{k}}} e^{-ik \cdot (x-y)} \\
&\equiv i\Delta^+(x-y), \tag{133}
\end{aligned}$$

where we have employed the  $\delta_{\vec{k}\vec{k}'}$  commutator and we must keep in mind that  $k^0$  and  $k'^0$  are the on-shell energies. We have also used  $\frac{1}{V} \sum_{\vec{k}} \rightarrow \frac{1}{(2\pi)^3} \int d^3\vec{k}$  for the continuum limit [see Eq. (10)].

- Similarly, one finds

$$[\phi^-(x), \phi^+(y)] = -[\phi^+(y), \phi^-(x)] = -i\Delta^+(y-x) \equiv i\Delta^-(x-y). \tag{134}$$

- Altogether, we have

$$[\phi(x), \phi(y)] = i\Delta(x-y) = i\Delta^+(x-y) + i\Delta^-(x-y)$$



$$= -\frac{i}{(2\pi)^3} \int \frac{d^3 \vec{k}}{\omega_{\vec{k}}} \sin[k \cdot (x - y)], \quad (135)$$

and, removing the  $i$ 's, we have

$$\Delta(x) = -\frac{1}{(2\pi)^3} \int \frac{d^3 \vec{k}}{\omega_{\vec{k}}} \sin k \cdot x. \quad (136)$$

- Note that  $(\square_x + \mu^2)\Delta(x - y) = 0$  by virtue of the fact that  $\phi(x)$  satisfies this equation. This means that  $\Delta(x - y)$  is not a propagator. “Propagator” is just another word for Green’s function, and a Green’s function should give a Dirac  $\delta$  function when acted on by the equation of motion.
- We can also write

$$\Delta(x - y) = \frac{-i}{(2\pi)^3} \int d^4 k \delta(k^2 - \mu^2) \epsilon(k^0) e^{-ik \cdot (x - y)}, \quad (137)$$

where  $d^4 k = dk^0 d^3 \vec{k}$  now lets  $k^0$  run from  $-\infty$  to  $+\infty$ , but then if we

write the  $\delta$  function in the form

$$\delta(k^2 - \mu^2) = \frac{1}{2\omega_{\vec{k}}} [\delta(k^0 + \omega_{\vec{k}}) + \delta(k^0 - \omega_{\vec{k}})] \quad (138)$$

we see that it picks out  $k^0 = \pm\sqrt{\vec{k}^2 + \mu^2}$  and these terms get combined with the correct sign by virtue of the definition

$$\epsilon(k^0) \equiv k^0/|k^0| = \begin{cases} +1, & \text{if } k^0 > 0 \\ -1, & \text{if } k^0 < 0 \end{cases} . \quad (139)$$

This form makes it clear that  $\Delta(x - y)$  is an invariant under (proper, i.e. no parity or time reversal) Lorentz transformations since each factor in the integrand is Lorentz invariant. (Here,  $\epsilon(k^0)$  is Lorentz invariant since proper Lorentz transformations do not interchange past and future and  $k^0$  only appears unsquared as a multiplier of  $x^0 - y^0$ .)

- Returning to the issue of causality, we have

$$[\phi(x), \phi(y)] = i\Delta(x - y) = 0 \quad \text{for } (x - y)^2 < 0 \quad (140)$$

by virtue of

1. the fact that

$$[\phi(\vec{x}, t), \phi(\vec{y}, t)] = 0 = i\Delta(\vec{x} - \vec{y}, 0); \quad (141)$$

2. the fact that  $\Delta(x - y)$  is a Lorentz invariant; and

3. the fact that we can go from the equal time frame where we know the answer from the equal time commutation relation to any other frame with  $(x - y)^2 < 0$  via a proper Lorentz transformation.

We can also think about this more directly. As we know,

$$\begin{aligned} \Delta(x) &= \Delta^+(x) + \Delta^-(x) = \Delta^+(x) - \Delta^+(-x) \\ &= \frac{-i}{(2\pi)^3} \int \frac{d^3\vec{k}}{2\omega_{\vec{k}}} [e^{-ik \cdot x} - e^{+ik \cdot x}]. \end{aligned} \quad (142)$$

Now, if  $x^2 < 0$ , we can perform a continuous Lorentz transformation on the 2nd term taking  $x \rightarrow -x$ . (If  $x^2 > 0$ , this is not possible as it would move us from inside the forward light cone to inside the backward light cone.) After this transformation, the two terms cancel.

- **The contour integration representation of  $\Delta$**

We can write (using the complex variable residue theorem)

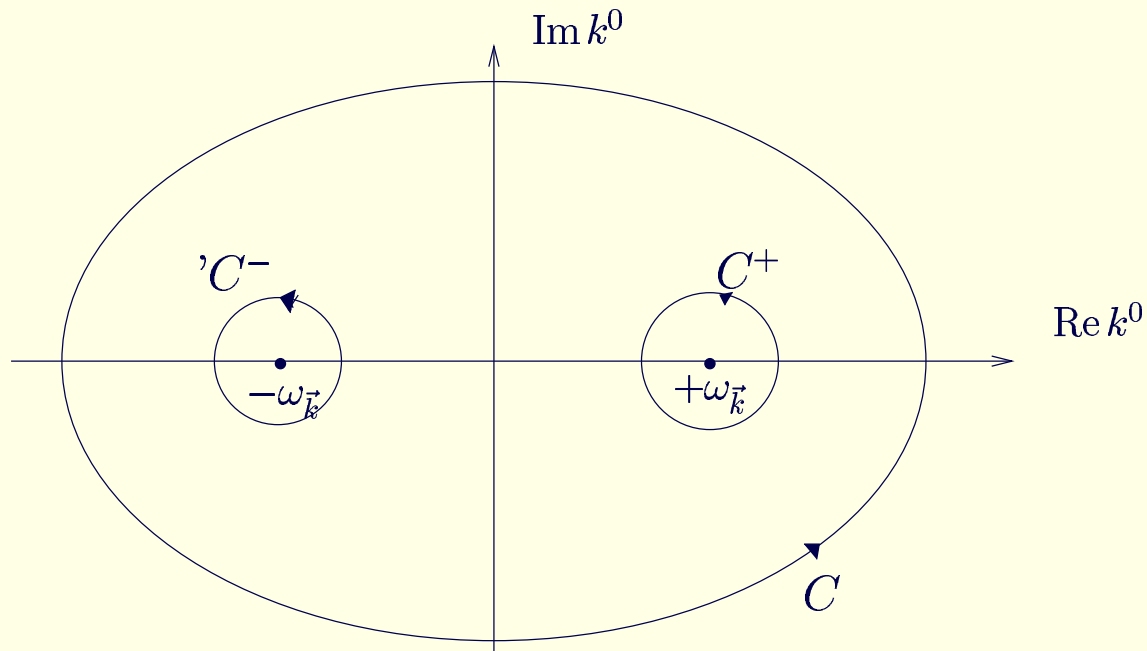
$$\Delta^\pm(x) = -\frac{1}{(2\pi)^4} \int_{C^\pm} d^4k \frac{e^{-ik \cdot x}}{k^2 - \mu^2} \quad (143)$$

where the contours are viewed as contours in the  $k^0$  complex plane,  $C^+$  is a small circular counter-clockwise contour around the pole of the denominator at  $k^0 = +\omega_{\vec{k}}$  and  $C^-$  is a small counter-clockwise contour around the pole at  $k^0 = -\omega_{\vec{k}}$ .

For example, for  $\Delta^-$ , we write

$$\begin{aligned} k^2 - \mu^2 &= (k^0)^2 - \vec{k}^2 - \mu^2 = (k^0 - \omega_{\vec{k}})(k^0 + \omega_{\vec{k}}) \\ &\sim -2\omega_{\vec{k}}(k^0 - (-\omega_{\vec{k}})), \end{aligned} \quad (144)$$

and via the residue theorem find that



$$\begin{aligned}
 \Delta^-(x) &= 2\pi i \frac{-1}{(2\pi)^4} \int d^3\vec{k} \frac{1}{-2\omega_{\vec{k}}} e^{-i(-\omega_{\vec{k}})x^0 + i\vec{k}\cdot\vec{x}} \\
 &= \frac{i}{(2\pi)^3} \int \frac{d^3\vec{k}}{2\omega_{\vec{k}}} e^{+i\omega_{\vec{k}}x^0 - i\vec{k}\cdot\vec{x}} \\
 &= \frac{i}{2(2\pi)^3} \int \frac{d^3\vec{k}}{\omega_{\vec{k}}} e^{-i(-k\cdot x)} \\
 &= -\Delta^+(-x) = \Delta^-(x).
 \end{aligned} \tag{145}$$

$\Delta(x) = \Delta^+(x) + \Delta^-(x)$  is then given by the same integral form for a large counter-clockwise contour that circles around both poles.

### The Feynman propagator

- This will be the propagator that actually enters into our calculations.

However, the simplified treatment of MS, will not allow a formal derivation of this fact.

- To begin, first note that

$$\begin{aligned} i\Delta^+(x-y) &= \langle 0 | [\phi^+(x), \phi^-(y)] | 0 \rangle = \langle 0 | \phi^+(x) \phi^-(y) | 0 \rangle \\ &= \langle 0 | \phi(x) \phi(y) | 0 \rangle \end{aligned} \quad (146)$$

where the first equality simply follows from the fact that the vacuum expectation value of a non-operator quantity (a “c-number”) is simply equal to the c-number. The 2nd equality follows from the fact that  $\langle 0 | \phi^-(y) \phi^+(x) | 0 \rangle = 0$  by virtue of the fact that  $\phi^+$  contains the annihilation operator that annihilates the vacuum. The final equality is due to the fact that all terms in  $\phi(x)\phi(y)$  other than  $\phi^+(x)\phi^-(y)$  give zero (because of  $a|0\rangle = 0$  and/or  $\langle 0|a^\dagger = 0$ ).

Further,

$$-i\Delta^-(x-y) = i\Delta^+(y-x) = \langle 0|\phi(y)\phi(x)|0\rangle. \quad (147)$$

- We now define the “time-ordered”, or  $T$ , product by

$$\begin{aligned} T\{\phi(x)\phi(y)\} &= \begin{cases} \phi(x)\phi(y), & \text{if } x^0 > y^0 \\ \phi(y)\phi(x), & \text{if } y^0 > x^0 \end{cases} \\ &= \theta(x^0 - y^0)\phi(x)\phi(y) + \theta(y^0 - x^0)\phi(y)\phi(x) \end{aligned} \quad (148)$$

That is, the operators are written in chronological order.

- The Feynman propagator is defined as the vacuum expectation value of the time-ordered product:

$$i\Delta_F(x-y) = \langle 0|T\{\phi(x)\phi(y)\}|0\rangle. \quad (149)$$

so that, using the earlier vacuum expectation value results of Eqs. (146) and (147), we have

$$\Delta_F(x) = \theta(x^0)\Delta^+(x) - \theta(-x^0)\Delta^-(x). \quad (150)$$

- Referring back to Eq. (149), we can think of  $\Delta_F(x - y)$  as representing a meson created at  $y$  and then traveling to  $x$  when  $x^0 > y^0$  and vice versa.

However, one should note that this time ordering picture is frame-dependent. What really enters a calculation in which a virtual meson is exchanged between two scattering particles is the sum of the two time orderings as incorporated in  $\Delta_F$ , which is a covariant functional form.

To be more precise, consider the scattering of two nucleons,  $n_1$  and  $n_2$ , via the exchange of a spin-0 “meson” described by our KG field. If we keep track of time-orderings, there is one diagram in which  $n_1$  first emits the  $\phi$  at time  $t_1$  which is then later absorbed by  $n_2$  at time  $t_2$  ( $t_2 > t_1$  case) and there is a 2nd time ordering in which  $n_2$  first emits the  $\phi$  at time  $t_2$  and then the  $\phi$  is later absorbed by  $n_1$  at time  $t_1$  ( $t_1 > t_2$  case). The full scattering amplitude must account for both cases, since we should integrate over all possible values of  $t_1$  and  $t_2$ . The Feynman propagator automatically accounts for the two possibilities.

- One should not be concerned that  $\Delta_F(x - y) \neq 0$  if  $(x - y)^2 < 0$  (space-like separation). A physical meson is not actually traveling faster than the speed of light between  $x$  and  $y$ ; it is a virtual meson. As a result, a real process at  $x$  (e.g. a particle emitting an on-shell meson which then travels



to  $y$  and is reabsorbed there) is not influencing a real process at  $y$ .

**Do new problem assigned at this point involving a “different” propagator.**

- Let us check explicitly that  $\Delta_F$  is a Green's function. We do this by first computing

$$\begin{aligned}
 & \frac{\partial^2}{\partial x^0{}^2} i\Delta_F(x - y) \\
 = & \frac{\partial^2}{\partial x^0{}^2} \langle 0 | T \{ \phi(x) \phi(y) \} | 0 \rangle \\
 = & \langle 0 | \frac{\partial^2}{\partial x^0{}^2} \left[ \theta(x^0 - y^0) \phi(x) \phi(y) + \theta(y^0 - x^0) \phi(y) \phi(x) \right] | 0 \rangle \\
 = & \langle 0 | \left[ \frac{\partial}{\partial x^0} \left( \left[ \frac{\partial}{\partial x^0} \theta(x^0 - y^0) \right] \phi(x) \phi(y) + \theta(x^0 - y^0) \frac{\partial}{\partial x^0} \phi(x) \phi(y) \right) \right. \right. \\
 & \left. \left. + \frac{\partial}{\partial x^0} \left( \left[ \frac{\partial}{\partial x^0} \theta(y^0 - x^0) \right] \phi(y) \phi(x) + \theta(y^0 - x^0) \phi(y) \frac{\partial}{\partial x^0} \phi(x) \right) \right] | 0 \rangle \\
 = & \langle 0 | \left[ \frac{\partial}{\partial x^0} \left( \delta(x^0 - y^0) \phi(x) \phi(y) + \theta(x^0 - y^0) \frac{\partial}{\partial x^0} \phi(x) \phi(y) \right) \right. \right. \\
 & \left. \left. + \frac{\partial}{\partial x^0} \left( -\delta(x^0 - y^0) \phi(y) \phi(x) + \theta(y^0 - x^0) \phi(y) \frac{\partial}{\partial x^0} \phi(x) \right) \right] | 0 \rangle
 \end{aligned}$$

$$\begin{aligned}
&= \langle 0 | \left[ \left( \frac{\partial}{\partial x^0} \delta(x^0 - y^0) \right) \phi(x) \phi(y) + 2\delta(x^0 - y^0) \frac{\partial}{\partial x^0} \phi(x) \phi(y) + \theta(x^0 - y^0) \frac{\partial^2}{\partial x^{02}} \phi(x) \phi(y) \right. \\
&\quad \left. + \left( -\frac{\partial}{\partial x^0} \delta(x^0 - y^0) \right) \phi(y) \phi(x) - 2\delta(x^0 - y^0) \phi(y) \frac{\partial}{\partial x^0} \phi(x) + \theta(y^0 - x^0) \phi(y) \frac{\partial^2}{\partial x^{02}} \phi(x) \right] | 0 \rangle \\
&= \langle 0 | \left[ +\delta(x^0 - y^0) \frac{\partial}{\partial x^0} \phi(x) \phi(y) + \theta(x^0 - y^0) \frac{\partial^2}{\partial x^{02}} \phi(x) \phi(y) \right. \\
&\quad \left. -\delta(x^0 - y^0) \phi(y) \frac{\partial}{\partial x^0} \phi(x) + \theta(y^0 - x^0) \phi(y) \frac{\partial^2}{\partial x^{02}} \phi(x) \right] | 0 \rangle \\
&= \langle 0 | \left[ +\delta(x^0 - y^0) [\pi(x), \phi(y)] + \theta(x^0 - y^0) \frac{\partial^2}{\partial x^{02}} \phi(x) \phi(y) + \theta(y^0 - x^0) \phi(y) \frac{\partial^2}{\partial x^{02}} \phi(x) \right] | 0 \rangle \\
&= \langle 0 | \left[ -\delta(x^0 - y^0) i\delta^3(\vec{x} - \vec{y}) + \theta(x^0 - y^0) \frac{\partial^2}{\partial x^{02}} \phi(x) \phi(y) + \theta(y^0 - x^0) \phi(y) \frac{\partial^2}{\partial x^{02}} \phi(x) \right] | 0 \rangle \\
&= \langle 0 | \left[ -i\delta^4(x - y) + \theta(x^0 - y^0) \frac{\partial^2}{\partial x^{02}} \phi(x) \phi(y) + \theta(y^0 - x^0) \phi(y) \frac{\partial^2}{\partial x^{02}} \phi(x) \right] | 0 \rangle \tag{15}
\end{aligned}$$

where we used some partial integrations to remove derivatives on the  $\delta$  functions, and used the equal time commutator for  $[\phi(y), \pi(x)]$ . The final step is to bring in the  $-\vec{\nabla}_x^2 + \mu^2$  part of the  $\square_x$  operator, which does not operate on the  $\theta$  functions (since they only contain the time variables) which gives

$$\begin{aligned}
& (\square_x + \mu^2)i\Delta_F(x - y) \\
= & \langle 0 | \left[ -i\delta^4(x - y) + \theta(x^0 - y^0)(\square_x + \mu^2)\phi(x)\phi(y) + \theta(y^0 - x^0)\phi(y)(\square_x + \mu^2)\phi(x) \right] | 0 \rangle \\
= & -i\delta^4(x - y), \tag{152}
\end{aligned}$$

where the last step just uses the fact that  $(\square_x + \mu^2)\phi(x) = 0$ .

- **A reminder about Green's functions:**

In general, the Green's function for a differential equation, defined by operator  $\mathcal{O}_x$  (here we use notation appropriate for a one-dimensional problem with coordinate  $x$ ) is defined by the fact that it allows a solution to the equation  $\mathcal{O}_x f(x) = j(x)$ , where  $j(x)$  is some source term. As you know,  $f(x)$  will consist of a sum of homogeneous solutions and an inhomogeneous component. The Green's function solution to this equation is to write

$$f(x) = \int dy G(x - y) j(y), \tag{153}$$

which will satisfy  $\mathcal{O}_x f(x) = j(x)$  if  $\mathcal{O}_x G(x - y) = \delta(x - y)$ . Another phraseology is that if  $j(y)$  is a localized impulse,  $j(y) = \delta(y - y_0)$ , then  $f(x) = G(x - y_0)$  is the response to that localized impulse. In

general,  $f(x)$  is obtained by integrating over the full set of impulses as specified by the source function  $j(y)$ . In general, there is more than one Green's function satisfying the inhomogeneous equation. The differences between different Green's functions are always homogeneous solutions of the differential equation. Usually, boundary conditions are imposed that determine precisely which (or what combination of) homogeneous solutions is to be added to the inhomogeneous solution.

In mathematical physics, you probably computed the Green's function for a theory using a complete basis of eigenfunctions as follows. Let us use the notation  $e_{\lambda_i}(x)$  for the (normalized) eigenstates of  $\mathcal{O}_x$  ( $\mathcal{O}_x e_{\lambda_i}(x) = \lambda_i e_{\lambda_i}(x)$ ). Then,

$$G(x, y) = \sum_i \lambda_i^{-1} e_{\lambda_i}(x) e_{\lambda_i}^*(y), \quad (154)$$

by virtue of the fact that  $\mathcal{O}_x G(x, y) = \sum_i e_{\lambda_i}(x) e_{\lambda_i}^*(y) = \delta(x - y)$ , where the 2nd equality is a statement of the completeness of the eigenfunction basis.

In the present case, we extend to 4 dimensions and use the eigenfunctions

$e_k(x) = \frac{1}{(2\pi)^2} e^{-ik \cdot x}$  which have

$$(\square_x + \mu^2)e_k(x) = (-k^2 + \mu^2)e_k(x) \equiv \lambda_k e_k(x) \quad (155)$$

(note that we don't require on-shell  $k^2 = \mu^2$  — we want all possible eigenvalues and not just the homogenous modes) and are normalized according to

$$\frac{1}{(2\pi)^4} \int d^4x e^{ip \cdot x} e^{-ik \cdot x} = \delta^4(k - p). \quad (156)$$

Thus,

$$\begin{aligned} G(x - y) &= \int d^4k \frac{1}{(2\pi)^2} e^{ik \cdot x} \frac{1}{-k^2 + \mu^2} \frac{1}{(2\pi)^2} e^{-ik \cdot y} \\ &= \frac{1}{(2\pi)^4} \int d^4k \frac{e^{ik \cdot (x-y)}}{-k^2 + \mu^2}. \end{aligned} \quad (157)$$

Here, we allow an extra minus sign, which is sort of a convention as to how to define a Green's function. Different Green's functions are defined by different ways of evading the denominator singularity (which is obviously closely related to what homogenous solutions, i.e. those with  $k^2 = \mu^2$  are included).

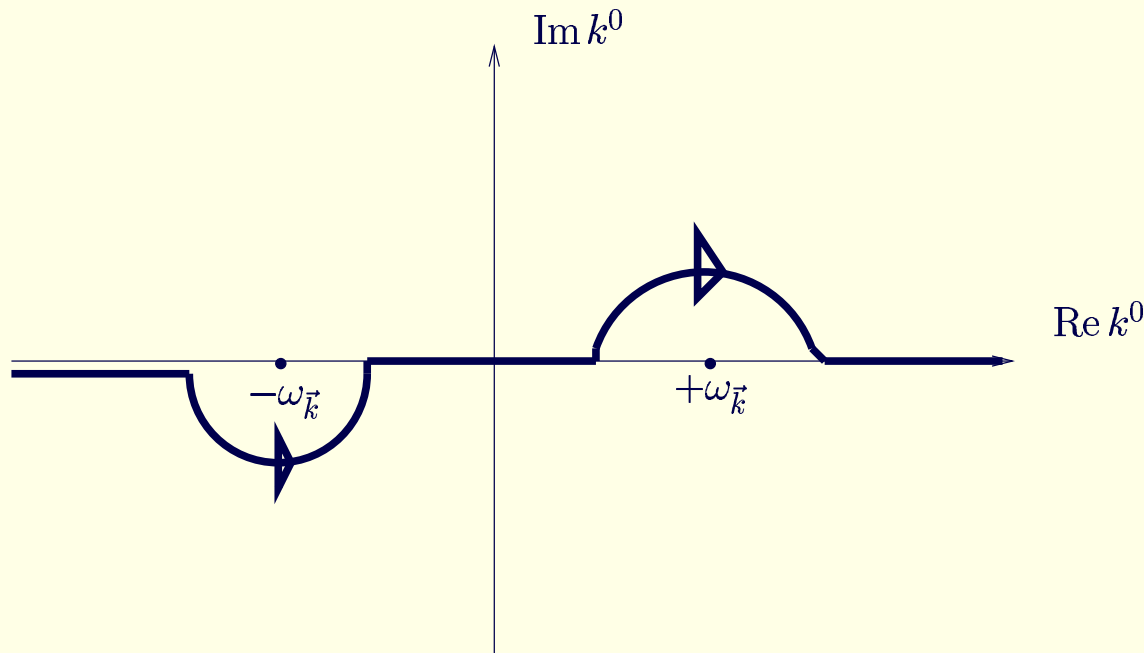
- The contour representation of  $\Delta_F$

We will show that

$$\Delta_F(x) = \frac{1}{(2\pi)^4} \int_{C_F} d^4k \frac{e^{-ik \cdot x}}{k^2 - \mu^2} \quad (158)$$

where the contour  $C_F$  runs along the real  $k^0$  axis, passing above the pole at  $k^0 = +\omega_{\vec{k}}$  and below the pole at  $k^0 = -\omega_{\vec{k}}$ .

Note: if this expression is correct, then  $(\square_x + \mu^2)\Delta_F(x - y) = -\delta^4(x - y)$  (using integral rep. of  $\delta^4$  function) in agreement with above result.



- First, we note that for  $x^0 > 0$ , we can close the contour in the lower 1/2 plane without picking up any contribution from the big half circle in the lower 1/2 plane (since  $e^{-ik^0 x^0} \propto e^{-i(i\text{Im}k^0 x^0)} \sim e^{(\text{Im}k^0)x^0}$  is damped when  $\text{Im}k^0 < 0$  and  $x^0 > 0$ ). The resulting contour is equivalent to  $C^+$ , except for the clockwise direction, and we see that we get  $\Delta_F(x) = \Delta^+(x)$  when  $x^0 > 0$ , where we recall the result of Eq. (143):

$$\Delta^\pm(x) = -\frac{1}{(2\pi)^4} \int_{C^\pm} d^4k \frac{e^{-ik \cdot x}}{k^2 - \mu^2} \quad (159)$$

- If  $x^0 < 0$ , we close in the upper 1/2 plane. The contour is equivalent to  $C^-$  (including contour direction), and we get  $\Delta_F(x) = -\Delta^-(x)$ .
- These results are in agreement with the earlier result

$$\Delta_F(x) = \theta(x^0)\Delta^+(x) - \theta(-x^0)\Delta^-(x). \quad (160)$$

- It is very useful to give another form for the expression of Eq. (158) using a different contour. The idea is that we can move the  $k^0$  contour to lie

exactly along the real axis provided we displace the poles by an infinitesimal amount in the correct direction. Defining  $\eta$  to be  $> 0$  and very small,

- the  $k^0 = -\omega_{\vec{k}}$  pole should be displaced by  $+i\eta$  into the upper 1/2 plane;
- the  $k^0 = +\omega_{\vec{k}}$  pole should be displaced by  $-i\eta$  into the lower 1/2 plane.

This is equivalent to replacing the denominator  $k^2 - \mu^2 = (k^0)^2 - \omega_{\vec{k}}^2$  by

$$\begin{aligned}
 & [k^0 - (\omega_{\vec{k}} - i\eta)][k^0 + (\omega_{\vec{k}} - i\eta)] \\
 = & [k^0]^2 - (\omega_{\vec{k}} - i\eta)^2 \\
 \sim & [k^0]^2 - \omega_{\vec{k}}^2 + i\epsilon \\
 = & [k^0]^2 - [\vec{k}^2 + \mu^2] + i\epsilon \\
 = & k^2 - \mu^2 + i\epsilon
 \end{aligned} \tag{161}$$

where we defined  $\epsilon = 2\eta\omega_{\vec{k}}$  which is again an infinitesimally small  $> 0$  quantity.

- Exactly equivalent results are obtained for the complex scalar field:

$$\langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle = i\Delta_F(x - y) \tag{162}$$



where  $\Delta_F$  is exactly the same function with the same contour representations as in the real scalar field case.

# The Dirac Field

## Anticommutator basics

- For Dirac fields (obeying the Dirac equation) We will convert from commutation conditions to anticommutation conditions for quantization (as we shall detail in a moment).

This also leads to a perfectly satisfactory number representation, but one implicitly based on Fermi statistics for the particles being created or annihilated.

This choice is required by causality (along with other basic requirements) in the field theory context.

- The bosonic quantization in terms of the  $a_r, a_r^\dagger$  operators ( $r = 1, 2, \dots$  denoting momentum states) required the commutation relations

$$[a_r, a_s^\dagger] = \delta_{rs}, \quad [a_r, a_s] = [a_r^\dagger, a_s^\dagger] = 0 \quad (163)$$

leading to our ability to define the vacuum (0-particle) state by

$$a_r|0\rangle = 0, r = 1, 2, \dots \quad (164)$$

and a number operator

$$N_r = a_r^\dagger a_r \quad (165)$$

obeying

$$[N_r, a_s] = -\delta_{rs}a_s, \quad [N_r, a_s^\dagger] = +\delta_{rs}a_s^\dagger \quad (166)$$

as follows from the general operator identity

$$[AB, C] = A[B, C] + [A, C]B \quad (167)$$

- The interpretation of the above commutators as implying particle annihilation or creation follows from, for example,

$$\begin{aligned} N a_s | \dots, s \rangle &= \sum_r N_r a_s | \dots, s \rangle \\ &= \sum_r ([N_r, a_s] + a_s N_r) | \dots, s \rangle \end{aligned}$$

$$\begin{aligned}
&= \sum_r (-\delta_{rs} a_s + a_s N_r) |\dots, s\rangle \\
&= a_s (-1 + N) |\dots, s\rangle \\
&= (n - 1) a_s |\dots, s\rangle
\end{aligned} \tag{168}$$

where we assumed  $N|\dots, s\rangle = n|\dots, s\rangle$ . In fact, it is precisely the  $s$  state that has been removed from  $|\dots, s\rangle$ . This can be checked once a Lagrangian has been specified and we know how to specify the physical properties (e.g.  $E$ ,  $\vec{P}$ , charge, ...) of a given state  $s$ .

- However, there is an alternative way to get a number operator obeying the creation and annihilation commutators of Eq. (166), based on using the anticommutator

$$[A, B]_+ \equiv \{A, B\} \equiv AB + BA, \tag{169}$$

for which we have

$$[AB, C] = A[B, C]_+ - [A, C]_+ B. \tag{170}$$

We need only require anticommutator quantization conditions,

$$[a_r, a_s^\dagger]_+ = \delta_{rs}, \quad [a_r, a_s]_+ = [a_r^\dagger, a_s^\dagger]_+ = 0 \quad (171)$$

to obtain exactly the same results as in Eq. (166) for  $N_r$  by using Eq. (170).

For example,

$$\begin{aligned} [N_r, a_s] &= [a_r^\dagger a_r, a_s] = a_r^\dagger [a_r, a_s]_+ - [a_r^\dagger, a_s]_+ a_r \\ &= 0 - \delta_{rs} a_r = -\delta_{rs} a_s. \end{aligned} \quad (172)$$

This leads to the interpretation of  $a_r$ ,  $a_r^\dagger$  and  $N_r = a_r^\dagger a_r$  as annihilation, creation and number operators operating on states based on the vacuum state defined by  $a_r|0\rangle = 0$ .

- For this setup, it is clear that we have Fermi-Dirac statistics.

For example, if we try to create 2 particles in the same state via  $a_r^\dagger a_r^\dagger |0\rangle$ , we can rewrite this as  $\frac{1}{2}[a_r^\dagger, a_r^\dagger]_+ |0\rangle$  which is zero because of the last anticommutator quantization condition of Eq. (171).

In fact, we have

$$N_r^2 = a_r^\dagger a_r a_r^\dagger a_r = a_r^\dagger (1 - a_r^\dagger a_r) a_r = N_r \quad (173)$$

(using  $a_r^2 = 0$ ) which can be rewritten as

$$N_r(N_r - 1) = 0 \quad (174)$$

implying that  $N_r$  can only have the eigenvalues 1 or 0.

Further, if we have two particles created from the vacuum in different states  $r \neq s$ , we have

$$|1_r 1_s\rangle = a_r^\dagger a_s^\dagger |0\rangle = -a_s^\dagger a_r^\dagger |0\rangle = -|1_s 1_r\rangle, \quad (175)$$

i.e. the state is antisymmetric under interchange of the particle labels.

- Of course, if we are forced to use these anticommutation quantization conditions in order to have proper causality (as we shall verify), we must simply accept the fact that there is no analogy to anything we have seen in non-relativistic or relativistic single particle quantum mechanics, all of which is based on the usual commutator quantization conditions.

## The Dirac Equation

- Recall that KG equation began with the relativistic energy momentum relation

$$E^2 = c^2 \vec{p}^2 + \mu^2 c^4 \quad (176)$$

followed by substitutions  $E \rightarrow i\hbar \frac{\partial}{\partial t}$ ,  $\vec{p} \rightarrow -i\hbar \vec{\nabla}$ , yielding the wave equation ( $\hbar = c = 1$ )

$$(\square + \mu^2)\phi(x) = 0 \quad (177)$$

with solutions  $\phi = e^{i(\vec{k}\cdot\vec{x} - \omega_{\vec{k}}t)}$  with  $\omega_{\vec{k}} = \pm(\vec{k}^2 + \mu^2)^{\frac{1}{2}}$ . The problem was that  $E = i\hbar \frac{\partial}{\partial t}$  gave positive energy for one choice but negative energy for the other choice.

- Dirac wanted to start with an equation that was linear in  $E$  by roughly taking the square root of Eq. (176). Beginning with

$$i\frac{\partial}{\partial t}\psi(\vec{x}, t) = H\psi(\vec{x}, t) \quad (178)$$

his first thought was to write this in terms of  $E$  and  $\vec{k}$  as

$$E = (\vec{k}^2 + m^2)^{\frac{1}{2}} \quad (179)$$

followed by  $E \rightarrow i\frac{\partial}{\partial t}$ ,  $\vec{p} \rightarrow -i\vec{\nabla}$ . (We use  $m$  instead of  $\mu$  for the Dirac equation.)

But, this obviously leads to an equation where time and space are treated asymmetrically implying that the equation is not relativistically covariant.

- At the very least, one needs an equation that is linear in  $\vec{k}$  as well as in  $E$ . A general form is

$$H = \vec{\alpha} \cdot \vec{k} + \beta m \quad (180)$$

leading to

$$(E - \vec{\alpha} \cdot \vec{k} - \beta m)\psi = 0, \quad (181)$$

and, thence, the wave equation

$$(i\frac{\partial}{\partial t} + i\vec{\alpha} \cdot \vec{\nabla} - \beta m)\psi = 0. \quad (182)$$



- Now, what about the  $\vec{\alpha}$  and  $\beta$  objects. They cannot have explicit dependence on  $\vec{x}$  or  $t$ , or derivatives with respect to the same, as such dependence would lead to space time dependent energy and therefore give rise to forces, whereas the equation we seek is one for a free particle.
- So, the only choices for  $\vec{\alpha}$  and  $\beta$  are numbers or matrices. We learn more by requiring that any solution  $\psi$  of Eq. (182) also be a solution of the KG equation (but not vice versa). This is a reasonable requirement since, in the absence of external fields, the wave packet solutions of Eq. (182) should have the classical relation between energy, momentum and mass. So, let us multiply Eq. (181) on the left by  $(E + \vec{\alpha} \cdot \vec{k} + \beta m)$  to obtain

$$\left[ E^2 - (\alpha_x^2 k_x^2 + \alpha_y^2 k_y^2 + \alpha_z^2 k_z^2) + (\alpha_x \alpha_y + \alpha_y \alpha_x) k_x k_y + (\alpha_z \alpha_y + \alpha_y \alpha_z) k_z k_y + (\alpha_x \alpha_z + \alpha_z \alpha_x) k_x k_z - m^2 \beta^2 - m \{ (\alpha_x \beta + \beta \alpha_x) k_x + (\alpha_y \beta + \beta \alpha_y) k_y + (\alpha_z \beta + \beta \alpha_z) k_z \} \right] \psi = 0. \quad (183)$$

To agree with the relativistic energy momentum relationships we require the following identities

$$\begin{aligned}
\alpha_x^2 &= \alpha_y^2 = \alpha_z^2 = \beta^2 = 1 \\
\alpha_x\alpha_y + \alpha_y\alpha_x &= \alpha_y\alpha_z + \alpha_z\alpha_y = \alpha_z\alpha_x + \alpha_x\alpha_z = 0 \\
\alpha_x\beta + \beta\alpha_x &= \alpha_y\beta + \beta\alpha_y = \alpha_z\beta + \beta\alpha_z = 0
\end{aligned} \tag{184}$$

i.e. the four quantities anticommute in pairs and their squares are unity. The smallest representation of this algebra is for  $\alpha_i$  and  $\beta$  to be  $4 \times 4$  matrices (see Schiff QM Sec. 52 for more detail on this).

The choice of Dirac was

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \tag{185}$$

where the  $\sigma_i$  are the  $2 \times 2$  Pauli matrices and the  $1$ 's are  $2 \times 2$  unit matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{186}$$

which we know anticommute and have unit square (implying the same for the  $\alpha_i$ ). That  $\beta$  and the  $\alpha_i$  anticommute is apparent from the explicit expressions above.

- So, now let us return to the basic Dirac equation, which you have seen can be written in the form

$$i\hbar \frac{\partial \psi(x)}{\partial t} = [c\vec{\alpha} \cdot (-i\hbar \vec{\nabla}) + \beta mc^2] \psi(x), \quad (187)$$

where I have temporarily reintroduced  $\hbar$  and  $c$ . In this equation  $\vec{\alpha}$  and  $\beta$  are  $4 \times 4$  matrices, and  $\psi$  is a 4-component object:  $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$ , where the individual components are often denoted  $\psi_\alpha$ ,  $\alpha = 1, 2, 3, 4$ , where  $\alpha$  is referred to as the Dirac index of the spinor.

Very often, as above, we will write equations in such a way that this Dirac index and the matrix etc. multiplications are implicit rather than explicit. Thus, for example, in the above equation there is an implicit Dirac index, call it  $\rho$ , and the last term could be written more explicitly as  $\beta_{\rho\lambda} \psi_\lambda(x)$ .

The Dirac equation can be written (after multiplying the equation by  $\beta$  and using  $\beta^2 = 1_{4 \times 4}$ , the  $4 \times 4$  identity matrix) in a more covariant looking

form (setting  $\hbar = c = 1$ ),

$$i\gamma^\mu \frac{\partial \psi(x)}{\partial x^\mu} - m 1_{4 \times 4} \psi(x) = 0 \quad (188)$$

by defining

$$\gamma^0 \equiv \beta, \quad \gamma^i = \beta \alpha^i, \quad i = 1, 2, 3, \quad (189)$$

where these “Dirac matrices”  $\gamma^\mu$  obey the anticommutation relations

$$[\gamma^\mu, \gamma^\nu]_+ = 2g^{\mu\nu} 1_{4 \times 4}, \quad (190)$$

and the hermiticity conditions

$$\gamma^{0\dagger} = \gamma^0, \quad \gamma^{j\dagger} = -\gamma^j \quad j = 1, 2, 3, \quad (191)$$

so that (using the anticommutation relation also)

$$\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0. \quad (192)$$

Usually, one does not explicitly write the  $1_{4 \times 4}$  matrix. It is understood to be present in the matrix equations we will be writing.

**Note:** The above hermiticity properties guarantee hermiticity of the Hamiltonian and momentum operators constructed below.

- There is more than one specific numerical representation of these  $\gamma$  matrices.

The so-called Dirac representation is that you may have already seen, and is most useful for discussing the low-energy, non-relativistic limit of the Dirac equation.

A better representation for high-energy applications is the so-called *Weyl* or *chiral* representation that is discussed in detail in Peskin and Schroeder.

We will avoid as long as possible writing down any specific representation. All representations of the matrices are related by unitary transformations.

- In a more general context, one could arrive at this collection of  $\gamma^\mu$  matrices by showing that they form the simplest general representation of the Lorentz group structure beyond the usual vector representation.

Later, we will give enough information about this approach in order that we can specify what the form of the rotation and Lorentz transformation generators is.

- We will define an adjoint field

$$\bar{\psi}(x) = \psi^\dagger(x)\gamma^0, \quad (193)$$

which, by starting with the Dirac equation for  $\psi(x)$ , is easily shown to satisfy

$$i\frac{\partial\bar{\psi}(x)}{\partial x^\mu}\gamma^\mu + m\bar{\psi}(x) = 0. \quad (194)$$

Note that any given component of the adjoint field  $\bar{\psi}_\alpha$  is independent of the  $\psi_\alpha$  component of the non-adjoint field since  $\psi_\alpha$  is a complex quantity and has an independent real and imaginary part. This is the analogue of the statement that  $\phi$  and  $\phi^\dagger$  are independent objects when dealing with a complex scalar field.

### The Dirac Lagrangian

- In order to proceed with the 2nd quantization game, we must first find a Lagrangian that yields the Dirac equation as its equation of motion.

An appropriate choice is

$$\mathcal{L} = \bar{\psi}(x) \left[ i\gamma^\mu \frac{\partial}{\partial x^\mu} - m \right] \psi(x). \quad (195)$$

(We will see that we must use  $\bar{\psi}$  instead of  $\psi^\dagger$  on the left in order that  $\mathcal{L}$  is a Lorentz invariant scalar.)

One should keep in mind that one can make this  $\mathcal{L}$  look different by adding the total derivative  $i \frac{\partial}{\partial x^\mu} [\bar{\psi} \gamma^\mu \psi]$  times any coefficient you like. Such a total derivative will not affect the action  $\propto \int d^4x \mathcal{L}$ .

For the equations of motion, we must keep in mind that  $\psi$  is a complex field and so we obtain two independent equations of motion by varying with respect to  $\psi_\alpha$  and  $\bar{\psi}_\alpha$ . First, recall the general form of the Euler Lagrange equations coming from  $\delta S = 0$ :

$$\frac{\partial \mathcal{L}}{\partial \phi_r} - \frac{\partial}{\partial x^\nu} \left( \frac{\partial \mathcal{L}}{\partial \left( \frac{\partial \phi_r}{\partial x^\nu} \right)} \right) = 0. \quad (196)$$

We now apply this to the Dirac  $\mathcal{L}$  above, treating  $\psi_\alpha$  and  $\bar{\psi}_\alpha$  as independent

fields. The Dirac equation for  $\psi_\alpha$  is very easily obtained by applying the Euler-Lagrange equation to the Dirac  $\mathcal{L}$  for the case of  $\phi_r = \bar{\psi}_\alpha$ :

$$\frac{\partial \mathcal{L}}{\partial \bar{\psi}_\alpha} = 0 = \left[ i \frac{\partial}{\partial x^\mu} \gamma_{\alpha\beta}^\mu - m 1_{\alpha\beta} \right] \psi_\beta(x) \quad (197)$$

where there is no term coming from  $\frac{\partial}{\partial x^\nu} \left( \frac{\partial \mathcal{L}}{\partial \left( \frac{\partial \bar{\psi}_\alpha}{\partial x^\nu} \right)} \right)$  simply because (for the particular form of  $\mathcal{L}$  employed) there is no term in  $\mathcal{L}$  that depends on  $\frac{\partial \bar{\psi}_\alpha}{\partial x^\nu}$ . In the above, I have made the Dirac indices completely explicit. Making them implicit, the above is identical to the covariant form of the Dirac equation given earlier.

- For 2nd quantization, we need the conjugate momentum densities of  $\psi_\alpha$  and  $\bar{\psi}_\alpha$ . From  $\mathcal{L}$  we have

$$\begin{aligned} \pi_\alpha(x) &= \frac{\partial \mathcal{L}}{\partial \dot{\psi}_\alpha} = i[\bar{\psi}(x)\gamma^0]_\alpha = i[\psi^\dagger \gamma^0 \gamma^0]_\alpha = i\psi^\dagger_\alpha \\ \bar{\pi}_\alpha(x) &= \frac{\partial \mathcal{L}}{\partial \dot{\bar{\psi}}_\alpha} = 0. \end{aligned} \quad (198)$$



## The Hamiltonian and Momentum operators

- The **Hamiltonian** and **momentum** operators are obtained from the Noether procedure and take the form:

$$\begin{aligned}
 H &= \int d^3\vec{x} \mathcal{H} = \int d^3\vec{x} [\pi_\alpha(x) \dot{\psi}_\alpha(x) + \bar{\pi}_\alpha(x) \dot{\bar{\psi}}_\alpha(x) - \mathcal{L}] \\
 &= \int d^3\vec{x} [i\psi^\dagger(x) \dot{\psi}(x) - \mathcal{L}] \\
 &= \int d^3\vec{x} [i\bar{\psi}(x) \gamma^0 \frac{\partial}{\partial x^0} \psi(x) - i\bar{\psi}(x) \gamma^\mu \frac{\partial}{\partial x^\mu} \psi(x) + m\bar{\psi}(x) \psi(x)] \\
 &= \int d^3\vec{x} \bar{\psi}(x) [-i\gamma^j \frac{\partial}{\partial x^j} + m] \psi(x) \\
 P^j &= \int d^3\vec{x} [\pi_\alpha(x) \frac{\partial \psi_\alpha(x)}{\partial x_j} + 0] \\
 &= \int d^3\vec{x} [i\psi^\dagger(x) (-\nabla^j) \psi(x) \quad \text{i.e.} \\
 \vec{P} &= -i \int d^3\vec{x} \psi^\dagger(x) \vec{\nabla} \psi(x) .
 \end{aligned}$$

## The Angular Momentum operator

- The **angular momentum** operator requires knowing what the generator of rotations is for the Dirac field. To this end, we must go to Peskin and Schroeder, chapter 3.

I outline the approach.

1. First, you must understand that rotations are in general generated by rotation operators of the form  $R = \exp[-i\theta^i J^i]$ , where the  $J^i$ ,  $i = 1, 2, 3$ , are called the generators of the rotation group.

The exact form of these generators depends upon the representation of the rotation group. Some examples that you know about already are:

- the non-relativistic spin-1/2 generators which are the Pauli matrices:

$$\vec{J} = \vec{\sigma}/2.$$

- the angular momentum operator in non-relativistic quantum mechanics:

$$\vec{J} = \vec{x} \times \vec{p} = \vec{x} \times (-i\vec{\nabla}).$$

These operators operate on different spaces. The Pauli matrices operate in the space of two-component spinors. The NRQM angular momentum operator operates on NRQM wave functions.

However, they have a common property; namely the commutation relations:

$$[J^i, J^j] = i\epsilon^{ijk} J^k . \quad (199)$$

These commutation relations for the group generators are what actually specifies the rotation group.

2. We can generalize the NRQM rotation generators to include Lorentz boosts by using an antisymmetric tensor notation:

$$J^{ij} = -i(x^i \nabla^j - x^j \nabla^i) \quad (200)$$

is identified with the above  $J^k$  according to  $J^3 = J^{12}$ , and so forth. This we generalize to

$$J^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu). \quad (201)$$

These 6 operators generate the three boosts and three rotations of the Lorentz group.

3. Using simple algebra, one finds the generalized Lorentz group commutators:

$$[J^{\mu\nu}, J^{\rho\sigma}] = i (g^{\nu\rho} J^{\mu\sigma} - g^{\mu\rho} J^{\nu\sigma} - g^{\nu\sigma} J^{\mu\rho} + g^{\mu\sigma} J^{\nu\rho}) . \quad (202)$$

4. Another simple representation of these  $J^{\mu\nu}$  is that for Lorentz transformations of the usual 4-vectors:

$$(\mathcal{J}^{\mu\nu})_{\alpha\beta} = i(\delta_{\alpha}^{\mu}\delta_{\beta}^{\nu} - \delta_{\beta}^{\mu}\delta_{\alpha}^{\nu}), \quad (203)$$

where the  $\alpha$  and  $\beta$  are the Lorentz indices that are connected with the vector indices on the coordinates.

You can easily verify that this representation of the  $J^{\mu\nu}$  obeys the general commutator relations above that specify the group.

Lorentz 4-vectors transform under infinitesimal boosts or rotations according to

$$V^{\alpha} \equiv \Lambda_{\beta}^{\alpha} V^{\beta} \rightarrow (\delta_{\beta}^{\alpha} - \frac{i}{2}\omega_{\mu\nu}(\mathcal{J}^{\mu\nu})_{\beta}^{\alpha}) V^{\beta}, \quad (204)$$

or equivalently

$$V_{\alpha} \rightarrow V_{\alpha} - \frac{i}{2}\omega_{\mu\nu}(\mathcal{J}^{\mu\nu})_{\alpha\beta} V^{\beta} \quad (205)$$

Plugging in Eq. (203) gives

$$V_{\alpha} \rightarrow V_{\alpha} + \omega_{\alpha\beta} V^{\beta}. \quad (206)$$

If we look back at the transform used to define the Noether current

Eq. (79) one finds

$$x_\alpha \rightarrow x_\alpha + \epsilon_{\alpha\beta} x^\beta \quad (207)$$

implying that  $\omega_{\alpha\beta} = \epsilon_{\alpha\beta}$ . This will be important shortly. For now let us look at two examples.

For example, a  $z$ -axis rotation would be specified by  $\omega_{12} = -\omega_{21} = -\theta_R$ . An infinitesimal boost in the  $x$ -direction would correspond to  $\omega_{01} = -\omega_{10} = -\beta_B$  (the boost velocity). In both cases, I have chosen the signs appropriate for the passive point of view. We can verify these claims as follows. Let us consider the boost case, for which the generator of interest is  $(\mathcal{J}^{01})^\alpha_\beta = i(g^{0\alpha}\delta_\beta^1 - \delta_\beta^0 g^{1\alpha})$ . The antisymmetry of  $\omega_{01} = -\omega_{10} = -\beta_B$  means that

$$-\frac{i}{2}\omega_{\mu\nu}(\mathcal{J}^{\mu\nu})^\alpha_\beta = -\beta_B(g^{0\alpha}\delta_\beta^1 - \delta_\beta^0 g^{1\alpha}). \quad (208)$$

Looking at the  $\alpha, \beta = 0, 1$  components of the general boost equation, we find  $\Lambda_0^0 = 1$ ,  $\Lambda_1^0 = -\beta_B$ ,  $\Lambda_0^1 = -\beta_B$  and  $\Lambda_1^1 = 1$ , where we had to use  $g^{00} = 1$  and  $g^{11} = -1$ . The result is

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & -\beta_B \\ -\beta_B & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}, \quad (209)$$

which is of course correct for the visualization (passive) where you look at some fixed physical location using the coordinates of a prime frame that is boosted with positive velocity along the  $x$  axis in the unprimed frame. Returning to the rotation example, we have

$$-\frac{i}{2}\omega_{\mu\nu}(\mathcal{J}^{\mu\nu})_{\beta}^{\alpha} = -\theta_R(g^{1\alpha}\delta_{\beta}^2 - \delta_{\beta}^1g^{2\alpha}) \quad (210)$$

and focusing on the  $\alpha, \beta = 1, 2$  components we have  $\Lambda_1^1 = 1$ ,  $\Lambda_2^2 = 1$ ,  $\Lambda_2^1 = \theta_R$  and  $\Lambda_1^2 = -\theta_R$ , yielding

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & \theta_R \\ -\theta_R & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (211)$$

which describes the coordinates of a given physical point using the coordinates defined in a new frame obtained from the old frame by rotating by an amount  $\theta_R$  about the positive  $\hat{z}$  axis (again the passive point of view).

Of course, these infinitesimal forms can be exponentiated to give the

usual finite results. For example, in the case of the boost we would have

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \exp \left[ \begin{pmatrix} 0 & -\beta_B \\ -\beta_B & 0 \end{pmatrix} \right] \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} \cosh \beta_B & -\sinh \beta_B \\ -\sinh \beta_B & \cosh \beta_B \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \quad (212)$$

5. An interesting exercise is to find all the finite dimensional representations of the Lorentz group. This we will not do in general. However, if we define  $n \times n$  matrices obeying  $[\gamma^\mu, \gamma^\nu]_+ = 2g^{\mu\nu} \times 1_{n \times n}$ , then one can show that the matrices defined by

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu] \equiv \frac{1}{2}\sigma^{\mu\nu} \quad (213)$$

satisfy the Lorentz generator commutator algebra.

The demonstration that the  $S^{\mu\nu}$  defined in this way satisfy the Lorentz generator commutator algebra (for any  $n$ ) is assigned as a problem.

Obviously the case of  $n = 4$  is what we shall employ.

6. The form of the  $\gamma^\mu$  in the Weyl representation of the Dirac matrices (this use of the word “representation” is not the same as the group theory

use) is

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \quad (214)$$

For this choice, we can work out the  $S^{\mu\nu}$  explicitly:

$$\begin{aligned} S^{0i} &= \frac{i}{4}[\gamma^0, \gamma^i] = -\frac{i}{2} \begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix} \\ S^{ij} &= \frac{i}{4}[\gamma^i, \gamma^j] = \frac{1}{2}\epsilon^{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \end{aligned} \quad (215)$$

We will need this explicit representation later.

7. A finite Lorentz transformation when working with the 4-component  $\psi_\alpha$  basis has the form

$$\Lambda_{1/2} = \exp\left(-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}\right). \quad (216)$$

In a moment, it will be important to note that  $\Lambda_{1/2}$  is not a unitary matrix for boosts, since the  $S^{0i}$  are not Hermitian. In fact, it is easily seen from the above explicit form of  $S^{0i}$  that  $S^{0i}$  is antihermitian; in contrast, the  $S^{ij}$  are hermitian.



8. There are several very important properties of these  $S^{\rho\sigma}$ .

(a)  $[\gamma^\mu, S^{\rho\sigma}] = (\mathcal{J}^{\rho\sigma})^\mu{}_\nu \gamma^\nu$ , which implies

$$\left(1 + \frac{i}{2}\omega_{\rho\sigma}S^{\rho\sigma}\right)\gamma^\mu\left(1 - \frac{i}{2}\omega_{\rho\sigma}S^{\rho\sigma}\right) = \left(1 - \frac{i}{2}\omega_{\rho\sigma}\mathcal{J}^{\rho\sigma}\right)^\mu{}_\nu\gamma^\nu \quad (217)$$

which is just the infinitesimal form of

$$\Lambda_{1/2}^{-1}\gamma^\mu\Lambda_{1/2} = \Lambda^\mu{}_\nu\gamma^\nu. \quad (218)$$

(b) Using these results, one can show that the Dirac equation is Lorentz covariant (takes the same form in different frames).

(c) It also allows us to see why we must use  $\bar{\psi}$  rather than  $\psi^\dagger$  in writing the Dirac field Lagrangian if the Lagrangian is to be a Lorentz invariant.

In particular, consider the form  $\psi^\dagger\psi$  (which, multiplies  $m$  in  $\mathcal{L}$ ). Under a Lorentz transformation,  $\psi^\dagger\psi \rightarrow \psi^\dagger\Lambda_{1/2}^\dagger\Lambda_{1/2}\psi$ . However,  $\Lambda_{1/2}$  is not a unitary matrix in the case of boosts for which the generators are not Hermitian. The trick is to note that the  $S^{ij}$  are hermitian whereas the  $S^{0i}$  are antihermitian ( $(S^{0i})^\dagger = -S^{0i}$ ) and that  $\gamma^0 S^{ij} = +S^{ij}\gamma^0$  whereas  $\gamma^0 S^{0i} = -S^{0i}\gamma^0$ .

As a result,

$$\Lambda_{1/2}^\dagger \gamma^0 = e^{+\frac{i}{2}\omega_{\mu\nu} S^{\mu\nu\dagger}} \gamma^0 = \gamma^0 e^{+\frac{i}{2}\omega_{\mu\nu} S^{\mu\nu}} = \gamma^0 \Lambda_{1/2}^{-1} \quad (219)$$

keeping in mind that the exponential is defined by power series expansion and that in each term of the power series we can use the fact that  $\gamma^0$  commutes with the  $S^{ij}$  but anticommutes with the  $S^{0i}$ .

Thus, we can compensate for the antihermiticity of the  $S^{0i}$  by introducing a  $\gamma^0$ , i.e. by using the form  $\psi^\dagger \gamma^0 \psi = \bar{\psi} \psi$ , which is a Lorentz invariant by virtue of the fact that

$$\bar{\psi} = \psi^\dagger \gamma^0 \rightarrow \psi^\dagger \Lambda_{1/2}^\dagger \gamma^0 = \psi^\dagger \gamma^0 \Lambda_{1/2}^{-1} = \bar{\psi} \Lambda_{1/2}^{-1} \quad (220)$$

under a Lorentz transformation.

9. A slightly more complicated derivation is required to show that the  $\bar{\psi}(i\gamma^\mu \partial_\mu)\psi$  part of  $\mathcal{L}$  is also a Lorentz invariant. For this, we must use Eq. (218).

As a problem, you are asked to show this as well as to show that  $\mathcal{L}$  is hermitian.

So, finally we are able to derive what the angular momentum operator is using the infinitesimal form of  $\Lambda_{1/2}$ . We have:

$$\begin{aligned}\psi_\alpha(x) \rightarrow \psi'_\alpha(x') &= \psi_\alpha(x) - \frac{i}{2} \omega_{\mu\nu} S_{\alpha\beta}^{\mu\nu} \psi_\beta(x) \\ &= \psi_\alpha(x) - \frac{i}{4} \epsilon_{\mu\nu} \sigma_{\alpha\beta}^{\mu\nu} \psi_\beta(x),\end{aligned}\tag{221}$$

where the last equation is written in the notation of Mandl and Shaw, i.e. we used  $\omega_{\mu\nu} = \epsilon_{\mu\nu}$ , and we used Eq. (213). This is now in the form required for obtaining the conserved current from the Noether theorem.

- **The Noether current for the angular momentum operator.**

Recall that for  $x_\mu \rightarrow x'_\mu = x_\mu + \epsilon_{\mu\nu} x^\nu$ , the Noether current depended upon the matrix appearing in

$$\phi_r(x) \rightarrow \phi'_r(x') = \phi_r(x) + \frac{1}{2} \epsilon_{\mu\nu} S_{rs}^{\mu\nu} \phi_s(x)\tag{222}$$

(where of course we think of the  $\phi_r$  as being the  $\psi_\alpha$  in the present Dirac

situation). Comparing Eq. (221) and Eq. (222), we see that

$$S_{rs}^{\mu\nu} \rightarrow -\frac{i}{2}\sigma^{\mu\nu}. \quad (223)$$

This appeared in the angular momentum operator in the form

$$M^{\mu\nu} = \int d^3\vec{x} \{ [x^\mu \mathcal{T}^{0\nu} - x^\nu \mathcal{T}^{0\mu}] + \pi_r(x) S_{rs}^{\mu\nu} \phi_s(x) \}. \quad (224)$$

Now, for discussing angular momentum and spin, we are only interested in the spatial coordinates,  $\mu, \nu = i, j$ . In this case, we write

$$M^{ij} = \int d^3\vec{x} \{ x^i \mathcal{T}^{0j} - x^j \mathcal{T}^{0i} + \pi_r(x) S_{rs}^{ij} \phi_s(x) \}, \quad (225)$$

where  $\mathcal{T}^{0j}(x) = \mathcal{P}^j(x) = \pi_r(x) \frac{\partial \phi_r(x)}{\partial x_j}$  is the general result, where we sum over all the fields in the  $r, s$  sums. (In the above  $\mathcal{P}^j(x)$  is the momentum density, the integral of which gives the 3-momentum given earlier in Eq. (199).) Now, in the Dirac case, this sum over fields and their conjugate momenta includes both  $\psi_\alpha$  (recall that the Dirac field

components are independent of one another) and  $\overline{\psi}_\alpha$  (since the real and imaginary components are independent of one another). However,  $\overline{\pi}_\alpha = 0$  for the specific form of  $\mathcal{L}$  we are employing. Thus, we are left with only the  $\psi_\alpha$  and  $\pi_\alpha = i\psi_\alpha^\dagger$  stuff. Thus, we find  $\mathcal{P}^j = i\psi_\alpha^\dagger \frac{\partial \psi_\alpha}{\partial x_j} = i\psi_\alpha^\dagger (-\vec{\nabla}^j \psi_\alpha)$ , and, using  $S_{\alpha\beta}^{ij} = -\frac{i}{2}\sigma_{\alpha\beta}^{ij}$ , we obtain

$$\pi_r S_{rs}^{ij} \phi_s \rightarrow i\psi_\alpha^\dagger \left[ -\frac{i}{2}\sigma_{\alpha\beta}^{ij} \right] \psi_\beta = \psi^\dagger \left[ \frac{\sigma^{ij}}{2} \right] \psi \quad (226)$$

(making the Dirac index sums implicit) and therefore

$$M^{ij} = \int d^3\vec{x} \psi^\dagger(x) \left\{ [x^i(-i\vec{\nabla}^j) - x^j(-i\vec{\nabla}^i)] + \frac{\sigma^{ij}}{2} \right\} \psi(x). \quad (227)$$

We now use the usual cyclic identification  $M^1 = M^{23}$ , etc. and rewrite the above equation as

$$\vec{M} = \int d^3\vec{x} \psi^\dagger(x) \left\{ [\vec{x} \times (-i\vec{\nabla})] + \frac{\vec{\Sigma}}{2} \right\} \psi(x), \quad (228)$$

where the  $4 \times 4$  matrices  $\vec{\Sigma}$  are defined as

$$\vec{\Sigma} = (\sigma^{23}, \sigma^{31}, \sigma^{12}). \quad (229)$$

(I don't like the Mandl-Shaw convention of using  $\vec{\sigma}$  here.)

In the Weyl representation of the Dirac matrices we have simply,

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}. \quad (230)$$

- From the above, it should be clear that the field appearing in the Dirac Lagrangian must necessarily correspond to a spin-1/2 particle.

To verify this explicitly, you would keep only the spin part of Eq. (228), perform the spatial integral and get a structure involving the  $b, b^\dagger, d, d^\dagger$  operators and then operate the operator form of  $\vec{M}$  on a single particle or antiparticle state,  $b^\dagger|0\rangle$  or  $d^\dagger|0\rangle$  and verify that these single particle states have eigenvalues of  $\pm\frac{1}{2}$  using, say, the helicity basis for the spin.

The Charge operator

- Since  $\mathcal{L}$  is invariant under a phase rotation of the complex field,  $\psi \rightarrow e^{i\phi}\psi$  (which implies  $\bar{\psi} \rightarrow e^{-i\phi}\bar{\psi}$ ), the Noether procedure leads to a conserved charge current. The general equation

$$Q = -iq \int d^3\vec{x} [\pi_r(x)\phi_r(x) - \pi_r^\dagger(x)\phi_r^\dagger(x)] \quad (231)$$

reduces in the present Dirac case to

$$\begin{aligned} Q &= -iq \int d^3\vec{x} [\pi_\alpha(x)\psi_\alpha(x) - 0] \\ &= -iq \int d^3\vec{x} [i\psi_\alpha^\dagger(x)\psi_\alpha(x)] = q \int d^3\vec{x} \psi^\dagger(x)\psi(x), \end{aligned}$$

and the corresponding charge-current density is

$$j^\mu = (\rho(x), \vec{j}(x)) = q\bar{\psi}(x)\gamma^\mu\psi(x). \quad (232)$$

One can explicitly check that  $\frac{\partial j^\mu}{\partial x^\mu} = 0$  follows by use of the Dirac equation for the  $\psi$  and  $\bar{\psi}$  fields.

- You may recall that Dirac actually wanted to have  $\psi^\dagger\psi = \bar{\psi}\gamma^0\psi$  be positive definite since it looked to him like a probability density. Thus, we have come 180°; our Lagrangian + Noether theorem implies that this *must* be the charge density, with opposite sign for particles vs. antiparticles. This we will see explicitly when we 2nd quantize.

The plane wave solutions of the Dirac equation

- To 2nd quantize the Dirac field, we need a complete set of solutions to the Dirac equation. This is closely analogous to what we did for  $\vec{A}$  and for  $\phi$ , where we expanded in terms of solutions of the KG equation or of  $\square\vec{A} = 0$ .

We then 2nd quantize by turning the c-number expansion coefficients into operators. In the Dirac case, we will demand that these operator coefficients obey **anticommutation** quantization conditions.

- Since the Dirac wave function has four components, and since we saw in the KG case that plane waves come in the form  $e^{ip\cdot x}$  and  $e^{-ip\cdot x}$ , we can anticipate that the complete set will look like

$$u_r(\vec{p})\frac{e^{-ip\cdot x}}{\sqrt{V}}, \quad v_r(\vec{p})\frac{e^{+ip\cdot x}}{\sqrt{V}}, \quad r = 1, 2, \quad (233)$$



where  $p^0 = E_{\vec{p}} \equiv \sqrt{m^2 + \vec{p}^2}$  is the 0th component of the 4-vector  $p$ .

- If these are plane wave solutions, they must obey

$$\begin{aligned} \left( i \frac{\partial}{\partial x^\mu} \gamma^\mu - m \right) u_r(\vec{p}) e^{-ip \cdot x} &= (\not{p} - m) u_r(\vec{p}) e^{-ip \cdot x} = 0 \\ \left( i \frac{\partial}{\partial x^\mu} \gamma^\mu - m \right) v_r(\vec{p}) e^{+ip \cdot x} &= (-\not{p} - m) v_r(\vec{p}) e^{+ip \cdot x} = 0 \end{aligned}$$

where we employ the notation  $\not{p} = v_\mu \gamma^\mu$ .

In fact,  $u_r$  and  $v_r$  must obey the equations

$$(\not{p} - m) u_r(\vec{p}) = 0, \quad (\not{p} + m) v_r(\vec{p}) = 0 \quad (234)$$

without the exponents in order that the Dirac equation plane waves obey the Dirac equation for any  $x$ .

- $u_r$  and  $v_r$  are called the positive and negative energy solutions by Dirac, and we will often use this language, but really they are the particle and antiparticle solutions, as we shall see.

- The  $r = 1, 2$  possibilities for each correspond to the two choices of spin along some direction.

The most convenient choice for this direction is the momentum itself, since the spin along the direction of motion is a constant of motion.

- To formalize this we define the helicity operator

$$\Sigma_{\vec{p}} \equiv \frac{\vec{\Sigma} \cdot \vec{p}}{|\vec{p}|} \quad (235)$$

for which we can choose our plane wave solutions so that

$$\begin{aligned} \Sigma_{\vec{p}} u_1(\vec{p}) &= +u_1(\vec{p}), \\ \Sigma_{\vec{p}} u_2(\vec{p}) &= -u_2(\vec{p}), \\ \Sigma_{\vec{p}} v_1(\vec{p}) &= -v_1(\vec{p}), \\ \Sigma_{\vec{p}} v_2(\vec{p}) &= +v_2(\vec{p}), \end{aligned}$$

where we can motivate the sign reversal for the  $v$ 's for the moment by referring to the “hole” way of thinking about the negative energy states; an antiparticle with  $+$  helicity is the absence of the  $-$  helicity plane wave state.

Note that the helicity operator has eigenvalues  $\pm 1$ ; we will see that this means that the Dirac particle has  $1/2$  unit of spin ( $\frac{\vec{\Sigma}}{2}$ ) either along the  $+\vec{p}$  direction or the  $-\vec{p}$  direction.

- We could write down explicit forms for the  $u_r$  and  $v_r$  that obey the following normalization and orthogonality relations:

$$\begin{aligned} u_r^\dagger(\vec{p})u_s(\vec{p}) &= v_r^\dagger(\vec{p})v_s(\vec{p}) = 2E_{\vec{p}}\delta_{rs} \\ u_r^\dagger(\vec{p})v_s(-\vec{p}) &= v_r^\dagger(\vec{p})u_s(-\vec{p}) = 0. \end{aligned}$$

**Note:** the  $\frac{E_{\vec{p}}}{m}$  normalization of Mandl-Shaw is archaic and extremely inconvenient in the high-energy or small mass limits. I hope I can consistently avoid using it. The normalization above puts the formulas for spin-1/2 and spin-0 fields on a much more parallel basis.

## 2nd Quantization

- To implement 2nd Quantization, we expand the Dirac field in terms of the above plane wave states:

$$\psi(x) = \psi^+(x) + \psi^-(x)$$

$$= \sum_{r, \vec{p}} \left( \frac{1}{2V E_{\vec{p}}} \right)^{1/2} \left[ c_r(\vec{p}) u_r(\vec{p}) e^{-ip \cdot x} + d_r^\dagger(\vec{p}) v_r(\vec{p}) e^{+ip \cdot x} \right] .$$

The conjugate field  $\bar{\psi} = \psi^\dagger \gamma^0$  will then have the expansion

$$\begin{aligned} \bar{\psi}(x) &= \bar{\psi}^+(x) + \bar{\psi}^-(x) \\ &= \sum_{r, \vec{p}} \left( \frac{1}{2V E_{\vec{p}}} \right)^{1/2} \left[ d_r(\vec{p}) \bar{v}_r(\vec{p}) e^{-ip \cdot x} + c_r^\dagger(\vec{p}) \bar{u}_r(\vec{p}) e^{+ip \cdot x} \right] , \end{aligned}$$

where we have defined  $\bar{u}_r \equiv u_r^\dagger \gamma^0, \dots$

In all of this, we are temporarily using a finite volume for which the momenta will take on the usual discrete values.

We have written  $c_r^\dagger$  instead of  $c_r^*$ , and  $d_r^\dagger$  instead of  $d_r^*$  in anticipation of these becoming operators as part of 2nd quantization.

- Quantize using *anti*commutators. We will see later that this is required by causality.

$$[c_r(\vec{p}), c_s^\dagger(\vec{k})]_+ = [d_r(\vec{p}), d_s^\dagger(\vec{k})]_+ = \delta_{rs} \delta_{\vec{p}\vec{k}} \quad (236)$$

with all other anticommutators equal to zero.

- Define the vacuum by

$$c_r(\vec{p})|0\rangle = d_r(\vec{p})|0\rangle = 0, \quad \text{all } \vec{p}, \text{ and } r=1,2. \quad (237)$$

- Define the number operators

$$N_r(\vec{p}) = c_r^\dagger(\vec{p})c_r(\vec{p}), \quad \bar{N}_r(\vec{p}) = d_r^\dagger(\vec{p})d_r(\vec{p}). \quad (238)$$

- From the above, we can show that it is consistent to think of  $c_r$  and  $d_r$  as being particle and antiparticle annihilation operators,  $c_r^\dagger$  and  $d_r^\dagger$  as being the corresponding creation operators, where  $N_r$  and  $\bar{N}_r$  are the number operators for particles and antiparticles, respectively
- But, we must check the other physical properties of the particles by looking for the constants of motion (conserved quantities) and seeing that they are sensible quantities related to what we expect for properties of particles.

- In computing these constants of motion, we will naturally measure things like energy, momentum and charge relative to the vacuum state if we, from the beginning, just employ the normal ordering prescription.

But, for fermions, we need to modify our definition just slightly to account for the anticommutation of the operators.

In arranging operators in the Bose case, we reordered assuming that all commutators vanished. In the Dirac case, we will reorder assuming all anticommutators vanish.

$$\begin{aligned}
 : \psi_\alpha \psi_\beta : &= : (\psi_\alpha^+ + \psi_\alpha^-)(\psi_\beta^+ + \psi_\beta^-) : \\
 &= \psi_\alpha^+ \psi_\beta^+ - \psi_\beta^- \psi_\alpha^+ + \psi_\alpha^- \psi_\beta^+ + \psi_\alpha^- \psi_\beta^- .
 \end{aligned}$$

- So now we compute the constants of motion obtained earlier from the symmetries of the Dirac  $\mathcal{L}$ , incorporating the normal ordering prescription, e.g.

$$H = \int d^3\vec{x} : \bar{\psi}(x) \left[ -i\gamma^j \frac{\partial}{\partial x^j} + m \right] \psi(x) : . \quad (239)$$

- We must repeatedly use the orthonormality properties for the plane waves. We obtain:

$$H = \sum_{r\vec{p}} E_{\vec{p}} [N_r(\vec{p}) + \bar{N}_r(\vec{p})] \quad (240)$$

$$\vec{P} = \sum_{r\vec{p}} \vec{p} [N_r(\vec{p}) + \bar{N}_r(\vec{p})] \quad (241)$$

$$Q = -e \sum_{r\vec{p}} [N_r(\vec{p}) - \bar{N}_r(\vec{p})] \quad (242)$$

where in the last equation we have adopted the usual convention in which the particle of the system is identified as the electron with charge  $-e$ , where  $e > 0$ .

Because we will need the pre-normal-ordered version of  $H$  later (to discuss the spin-statistics connection), let us derive the form of  $H$ . The steps are (without normal ordering):

$$\begin{aligned} H &= \int d^3\vec{x} \bar{\psi}(\vec{x}) [-i\gamma^j \partial_j + m] \psi(\vec{x}) \\ &= \int d^3\vec{x} \sum_{\vec{p},r} \frac{1}{\sqrt{2VE_{\vec{p}}}} \sum_{\vec{k},s} \frac{1}{\sqrt{2VE_{\vec{k}}}} \left( d_r(\vec{p}) \bar{v}_r(\vec{p}) e^{-ip\cdot x} + c_r^\dagger(\vec{p}) \bar{u}_r(\vec{p}) e^{ip\cdot x} \right) \times \end{aligned}$$

$$\begin{aligned}
& [-i\gamma^j \partial_j + m] \left( c_s(\vec{k}) u_s(\vec{k}) e^{-ik \cdot x} + d_s^\dagger(\vec{k}) v_s(\vec{k}) e^{ik \cdot x} \right) \\
= & \int d^3 \vec{x} \sum_{\vec{p}, r} \frac{1}{\sqrt{2V E_{\vec{p}}}} \sum_{\vec{k}, s} \frac{1}{\sqrt{2V E_{\vec{k}}}} \left( d_r(\vec{p}) \bar{v}_r(\vec{p}) e^{-ip \cdot x} + c_r^\dagger(\vec{p}) \bar{u}_r(\vec{p}) e^{ip \cdot x} \right) \times \\
& \left( c_s(\vec{k}) ((-i)(-i)\gamma^j k_j + m) u_s(\vec{k}) e^{-ik \cdot x} + d_s^\dagger(\vec{k}) ((-i)(+i)\gamma^j k_j + m) v_s(\vec{k}) e^{ik \cdot x} \right) \\
= & \int d^3 \vec{x} \sum_{\vec{p}, r} \frac{1}{\sqrt{2V E_{\vec{p}}}} \sum_{\vec{k}, s} \frac{1}{\sqrt{2V E_{\vec{k}}}} \left( d_r(\vec{p}) \bar{v}_r(\vec{p}) e^{-ip \cdot x} + c_r^\dagger(\vec{p}) \bar{u}_r(\vec{p}) e^{ip \cdot x} \right) \times \\
& \left( c_s(\vec{k}) (+k^0 \gamma^0) u_s(\vec{k}) e^{-ik \cdot x} + d_s^\dagger(\vec{k}) (-k^0 \gamma^0) v_s(\vec{k}) e^{ik \cdot x} \right) \\
= & \sum_{\vec{p}, r} \frac{1}{\sqrt{2V E_{\vec{p}}}} \sum_{\vec{k}, s} \frac{1}{\sqrt{2V E_{\vec{k}}}} V \left[ d_r(\vec{p}) c_s(\vec{k}) \bar{v}_r(\vec{p}) E_{\vec{k}} \gamma^0 u_s(\vec{k}) \delta_{\vec{p}, -\vec{k}} e^{-2iE_{\vec{k}} x^0} \right. \\
& + c_r^\dagger(\vec{p}) c_s(\vec{k}) \bar{u}_r(\vec{p}) (E_{\vec{k}} \gamma^0) u_s(\vec{k}) \delta_{\vec{p}, \vec{k}} \\
& + d_r(\vec{p}) d_s^\dagger(\vec{k}) \bar{v}_r(\vec{p}) (-E_{\vec{k}} \gamma^0) v_s(\vec{k}) \delta_{\vec{p}, \vec{k}} \\
& \left. + c_r^\dagger(\vec{p}) d_s^\dagger(\vec{k}) \bar{u}_r(\vec{p}) (-E_{\vec{k}} \gamma^0) v_s(\vec{k}) \delta_{\vec{p}, -\vec{k}} e^{+2iE_{\vec{k}} x^0} \right] \\
= & \frac{1}{2} \sum_{\vec{p}, r, s} \left[ d_r(\vec{p}) c_s(-\vec{p}) \bar{v}_r(\vec{p}) \gamma^0 u_s(-\vec{p}) e^{-2iE_{\vec{p}} x^0} + c_r^\dagger(\vec{p}) c_s(\vec{p}) \bar{u}_r(\vec{p}) \gamma^0 u_s(\vec{p}) \right. \\
& \left. - d_r(\vec{p}) d_s^\dagger(\vec{p}) \bar{v}_r(\vec{p}) \gamma^0 v_s(\vec{p}) - c_r^\dagger(\vec{p}) d_s^\dagger(-\vec{p}) \bar{u}_r(\vec{p}) \gamma^0 v_s(-\vec{p}) e^{+2iE_{\vec{p}} x^0} \right]
\end{aligned}$$



$$\begin{aligned}
&= \frac{1}{2} \sum_{\vec{p}, r, s} \left[ d_r(\vec{p}) c_s(-\vec{p})(0) e^{-2iE_{\vec{p}}x^0} + c_r^\dagger(\vec{p}) c_s(\vec{p}) 2E_{\vec{p}} \delta_{rs} \right. \\
&\quad \left. - d_r(\vec{p}) d_s^\dagger(\vec{p}) 2E_{\vec{p}} \delta_{rs} - c_r^\dagger(\vec{p}) d_s^\dagger(-\vec{p})(0) e^{+2iE_{\vec{p}}x^0} \right] \\
&= \sum_{\vec{p}, r} E_{\vec{p}} \left[ c_r^\dagger(\vec{p}) c_r(\vec{p}) - d_r(\vec{p}) d_r^\dagger(\vec{p}) \right] \tag{243}
\end{aligned}$$

where we used

$$(\not{k} - m)u(\vec{k}) = (\gamma^0 k_0 + \gamma^j k_j - m)u(\vec{k}) = 0, \quad \Rightarrow (-\gamma^j k_j + m)u(\vec{k}) = \gamma^0 k^0 u(\vec{k}) \tag{244}$$

$$(\not{k} + m)v(\vec{k}) = (\gamma^0 k_0 + \gamma^j k_j + m)v(\vec{k}) = 0, \quad \Rightarrow (\gamma^j k_j + m)v(\vec{k}) = -\gamma^0 k^0 v(\vec{k}). \tag{245}$$

and the orthonormal properties outlined earlier of the spinors outlined earlier:

$$\bar{u}_r(\vec{p}) \gamma^0 u_s(\vec{p}) = u_r^\dagger(\vec{p}) u_s(\vec{p}) = 2E_{\vec{p}} \delta_{rs} \tag{246}$$

$$\bar{v}_r(\vec{p}) \gamma^0 v_s(\vec{p}) = v_r^\dagger(\vec{p}) v_s(\vec{p}) = 2E_{\vec{p}} \delta_{rs} \tag{247}$$

$$\bar{v}_r(\vec{p}) \gamma^0 u_s(-\vec{p}) = v_r^\dagger(\vec{p}) u_s(-\vec{p}) = 0 \tag{248}$$

$$\bar{u}_r(\vec{p}) \gamma^0 v_s(-\vec{p}) = u_r^\dagger(\vec{p}) v_s(-\vec{p}) = 0. \tag{249}$$

You are assigned the problem of showing that  $Q$  is given by the equation Eq. (242), following similar procedures to the  $H$  derivation above.

- The last item on the list is the spin operator. We employ the helicity

(longitudinal spin in the words of Mandl-Shaw) operator which is the non-angular-momentum part of the total angular momentum operator we wrote down from the rotational symmetry of  $\mathcal{L}$ :

$$S_{\vec{p}} = \frac{1}{2} \int d^3\vec{x} : \psi^\dagger(x) \Sigma_{\vec{p}} \psi(x) :, \quad (250)$$

from which one finds by direct computation

$$S_{\vec{p}} c_r^\dagger(\vec{p}) |0\rangle = (-1)^{r+1} \frac{1}{2} c_r^\dagger(\vec{p}) |0\rangle, \quad (251)$$

$$S_{\vec{p}} d_r^\dagger(\vec{p}) |0\rangle = (-1)^{r+1} \frac{1}{2} d_r^\dagger(\vec{p}) |0\rangle, \quad (252)$$

which is to say that the  $r = 1$  states for both the particle and antiparticle have  $+1/2$  helicity, while the  $r = 2$  states have  $-1/2$  helicity.

The “flip” of the antiparticle projection helicity relative to the simple result of  $\Sigma_{\vec{p}}$  operating on the  $v_r$  states comes about as a result of the anticommutations that emerge in the above computations. Let us check this.

$$\begin{aligned}
S_{\vec{p}} d_r^\dagger(\vec{p}) |0\rangle &= \frac{1}{2} \int d^3 \vec{x} : \psi^\dagger(\vec{x}) \vec{\Sigma} \cdot \hat{p} \psi(\vec{x}) : d_r^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} \int d^3 \vec{x} \sum_{\vec{k},s} \frac{1}{\sqrt{2V E_{\vec{k}}}} e^{-i\vec{k}\cdot\vec{x}} \sum_{\vec{l},t} \frac{1}{\sqrt{2V E_{\vec{l}}}} e^{+i\vec{l}\cdot\vec{x}} \\
&\quad v_s^\dagger(\vec{k}) \vec{\Sigma} \cdot \hat{p} v_t(\vec{l}) : d_s(\vec{k}) d_t^\dagger(\vec{l}) : d_r^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} \sum_{\vec{k},s} \frac{1}{\sqrt{2V E_{\vec{k}}}} \sum_{\vec{l},t} \frac{1}{\sqrt{2V E_{\vec{l}}}} V \delta_{\vec{k}\vec{l}} \\
&\quad v_s^\dagger(\vec{k}) \vec{\Sigma} \cdot \hat{p} v_t(\vec{l}) : d_s(\vec{k}) d_t^\dagger(\vec{l}) : d_r^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} \sum_{\vec{k},s,t} \frac{1}{2E_{\vec{k}}} v_s^\dagger(\vec{k}) \vec{\Sigma} \cdot \hat{p} v_t(\vec{k}) : d_s(\vec{k}) d_t^\dagger(\vec{k}) : d_r^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} \sum_{\vec{k},s,t} \frac{1}{2E_{\vec{k}}} v_s^\dagger(\vec{k}) \vec{\Sigma} \cdot \hat{p} v_t(\vec{k}) (-) d_t^\dagger(\vec{k}) d_s(\vec{k}) d_r^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} \sum_{\vec{k},s,t} \frac{1}{2E_{\vec{k}}} v_s^\dagger(\vec{k}) \vec{\Sigma} \cdot \hat{p} v_t(\vec{k}) (-) d_t^\dagger(\vec{k}) \\
&\quad (-d_r^\dagger(\vec{p}) d_s(\vec{k}) + \delta_{\vec{p}\vec{k}} \delta_{rs}) |0\rangle
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \sum_t \frac{1}{2E_{\vec{p}}} v_r^\dagger(\vec{p}) \vec{\Sigma} \cdot \hat{p} v_t(\vec{p}) (-) d_t^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} \sum_t \frac{1}{2E_{\vec{p}}} v_r^\dagger(\vec{p}) (-1)^t v_t(\vec{p}) (-) d_t^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} \sum_t \frac{1}{2E_{\vec{p}}} (-1)^t 2E_{\vec{p}} \delta_{rt} (-) d_t^\dagger(\vec{p}) |0\rangle \\
&= \frac{1}{2} (-1)^{r+1} d_r^\dagger(\vec{p}) |0\rangle \tag{253}
\end{aligned}$$

Can you justify all the steps? There is one little detail that I sort of slipped in. In principle, what you should actually do is compute  $\vec{J}_{op} \cdot \vec{P}_{op} d_r^\dagger(\vec{p}) |0\rangle$ . But, of course  $\vec{p}$  pops out, giving you  $\vec{J}_{op} \cdot \vec{p} = \dots$ . There is another way that should be equivalent. That is to employ the helicity density  $\vec{J}(x) \cdot \vec{P}(x)$  (each being the appropriate bilinear form), integrated over  $\int d^3\vec{x}$  and operate this on  $d_r^\dagger(\vec{p})$ . This should give the same answer but is a far harder computation.

- Thus the particles and antiparticles of the theory have all the same properties except for the sign of the charge, as required for the use of the words particle and antiparticle.

## The Majorana representation of the Dirac matrices

- To make the symmetry between antiparticles and particles more apparent, one would use a particular representation of the  $\gamma$  matrices, the “Majorana” representation which has the property:

$$\gamma_M^{\mu*} = -\gamma_M^{\mu}, \quad \mu = 0, 1, 2, 3 \quad (254)$$

where the  $*$  denotes simple complex conjugation. In other words, in the Majorana rep. all four  $\gamma$  matrices are pure imaginary. Explicit forms of the  $\gamma_M$  that also obey the anticommutation relations for the  $\gamma$  matrices are given in the Appendix of Mandl-Shaw, for example.

- In the Majorana representation, the Dirac operator

$$\left( i\gamma_M^{\mu} \frac{\partial}{\partial x^{\mu}} - m \right) \quad (255)$$

is real. Hence, if  $\psi_M$  is a solution of the Dirac equation in a Majorana representation, so is its complex conjugate  $\psi_M^*$ .

It follows that if the positive energy solutions are denoted by

$$u_{Mr}(\vec{p}) \frac{e^{-ip \cdot x}}{\sqrt{V}}, \quad (256)$$

in the Majorana representation, then the corresponding negative energy solutions are

$$u_{Mr}^*(\vec{p}) \frac{e^{ip \cdot x}}{\sqrt{V}}. \quad (257)$$

**Proof:**

$$\begin{aligned} 0 &= \left[ \left( i\gamma_M^\mu \frac{\partial}{\partial x^\mu} - m \right) u_{Mr}(\vec{p}) e^{-ix \cdot p} \right]^* \\ &= \left[ (\not{p} - m) u_{Mr}(\vec{p}) e^{-ix \cdot p} \right]^* \\ &= (-\not{p} - m) u_{Mr}^*(\vec{p}) e^{+ix \cdot p} \\ &= \left( i\gamma_M^\mu \frac{\partial}{\partial x^\mu} - m \right) u_{Mr}^*(\vec{p}) e^{+ix \cdot p}, \end{aligned}$$

where we, of course, used  $\not{p}^* = -\not{p}$  in the Majorana representation.

- Thus, we have  $v_r(\vec{p}) = u_{Mr}^*(\vec{p})$  which implies the expansions:

$$\psi_M(x) = \sum_{r\vec{p}} \left( \frac{1}{2VE_{\vec{p}}} \right)^{1/2} [c_r(\vec{p})u_{Mr}(\vec{p})e^{-ip \cdot x} + d_r^\dagger(\vec{p})u_{Mr}^*(\vec{p})e^{+ip \cdot x}]$$

$$\psi_M^\dagger(x) = \sum_{r\vec{p}} \left( \frac{1}{2VE_{\vec{p}}} \right)^{1/2} [d_r(\vec{p})u_{Mr}(\vec{p})e^{-ip \cdot x} + c_r^\dagger(\vec{p})u_{Mr}^*(\vec{p})e^{+ip \cdot x}].$$

In the last equation, we avoided using  $\bar{\psi}$  so as to show the complete symmetry of  $\psi$  and  $\psi^\dagger$ : i.e.  $c_r(\vec{p})$  and  $d_r(\vec{p})$  are multiplied by exactly the same plane wave function.

- The Majorana representation is useful for particles that are their own antiparticles (such particles must be neutral), but is not particularly useful in other cases.

Thus, for the moment, we will stick to the general representation notation.

The field anticommutation relations

- Starting from the anticommutator algebra for the  $c, c^\dagger, d, d^\dagger$ , one can “easily” show that

$$[\psi_\alpha(x), \psi_\beta(y)]_+ = [\bar{\psi}_\alpha(x), \bar{\psi}_\beta(y)]_+ = 0 \quad (258)$$

$$[\psi_\alpha^\pm(x), \bar{\psi}_\beta^\mp(y)]_+ = i \left( i\gamma^\mu \frac{\partial}{\partial x^\mu} + m \right)_{\alpha\beta} \Delta^\pm(x - y) = iS_{\alpha\beta}^\pm(x - y), \quad (259)$$

where the  $\Delta^\pm(x)$  are the invariant  $\Delta$ -functions encountered in our study of the KG equation, except that they are defined using mass  $m$  of the Dirac particle rather than some other unrelated mass  $\mu$  of a Klein Gordon particle. **You are assigned the problem of verifying Eq. (259).**

- Using the above two equations we obtain easily

$$[\psi(x), \bar{\psi}(y)]_+ = iS(x - y) \quad (260)$$

where we define, using  $\Delta(x) = \Delta^+(x) + \Delta^-(x)$  from KG results,

$$S(x) \equiv S^+(x) + S^-(x) = \left( i\gamma^\mu \frac{\partial}{\partial x^\mu} + m \right) \Delta(x) \equiv (i\partial\!\!\!/ + m)\Delta(x) \quad (261)$$



where Dirac matrix indices are implicit.

- It is thought-provoking to use this result to compute the equal time anticommutator

$$[\psi(\vec{x}, t), \bar{\psi}(\vec{y}, t)]_+ = \left[ \left( i\gamma^\mu \frac{\partial}{\partial x^\mu} + m \right) \Delta(x - y) \right]_{x^0=y^0=t} . \quad (262)$$

One way of proceeding is to recall that  $\Delta(x - y) = [\phi(x), \phi(y)]$  ( $\phi$  being a real scalar field) and that  $[\dot{\phi}(x), \phi(y)]_{x^0=y^0} = -i\delta^3(\vec{x} - \vec{y})$  was the defining equal time commutator for the scalar field 2nd quantization.

Now, since the  $i\gamma^i \frac{\partial}{\partial x^i} + m$  part of the operator above does not involve the time derivative we have

$$\begin{aligned} \left[ \left( i\gamma^i \frac{\partial}{\partial x^i} + m \right) [\phi(x), \phi(y)] \right]_{x^0=y^0} &= \left( i\gamma^i \frac{\partial}{\partial x^i} + m \right) [\phi(x), \phi(y)]_{x^0=y^0} \\ &= \left( i\gamma^i \frac{\partial}{\partial x^i} + m \right) 0 \\ &= 0 \end{aligned} \quad (263)$$

where  $[\phi(x), \phi(y)]_{x^0=y^0} = 0$  comes from the fact that the fields commute for space like separations.

Thus, we are left only with the time derivative piece:

$$\begin{aligned}
 \left[ i\gamma^0 \frac{\partial}{\partial x^0} [\phi(x), \phi(y)] \right]_{x^0=y^0} &= i\gamma^0 [\dot{\phi}(x), \phi(y)]_{x^0=y^0} \\
 &= i\gamma^0 [-i\delta^3(\vec{x} - \vec{y})] \\
 &= \gamma^0 \delta^3(\vec{x} - \vec{y}). \tag{264}
 \end{aligned}$$

Of course, remembering that  $\bar{\psi}(y) = \psi^\dagger(y)\gamma^0$ , our result can be rewritten (displaying the previously implicit Dirac indices explicitly) as:

$$[\psi_\alpha(\vec{x}, t), \psi_\beta^\dagger(\vec{y}, t)]_+ = \delta_{\alpha\beta} \delta^3(\vec{x} - \vec{y}). \tag{265}$$

If we recall that  $\pi_\alpha(y) = i\psi_\alpha^\dagger$ , the above gives

$$[\psi_\alpha(\vec{x}, t), \pi_\beta(\vec{y}, t)]_+ = i\delta_{\alpha\beta} \delta^3(\vec{x} - \vec{y}). \tag{266}$$

This is the exact analogue of the 2nd quantization we used for the scalar field case. The only difference is that for the Dirac field we have employed

the anticommutator instead of the commutator. In short, an equivalent way to get at the anticommutator quantization rules for the  $c, d, c^\dagger, d^\dagger$  operators would have been to require Eq. (266).

- This brings up the question of **normalizations**. Just as discussed in the scalar field case, there is a certain arbitrariness in the normalizations employed when writing

$$\psi(x) = \sum_{r, \vec{p}} \left( \frac{1}{2V E_{\vec{p}}} \right)^{1/2} [c_r(\vec{p}) u_r(\vec{p}) e^{-ip \cdot x} + d_r^\dagger(\vec{p}) v_r(\vec{p}) e^{+ip \cdot x}] \quad (267)$$

and in the  $[c_r(\vec{p}), c_r^\dagger(\vec{k})]_+ = \delta_{\vec{p}\vec{k}}, \dots$  anticommutators. Keeping the same normalization convention for the spinors  $u_r$  and  $v_r$ , we could still change the  $\left( \frac{1}{2V E_{\vec{p}}} \right)^{1/2}$  factor to  $f^{1/2}$  (say) times this while retaining Eq. (266) (which has the **standard** 2nd quantization normalization) provided we make a compensating change in the normalization of the Fock space  $c, d, c^\dagger, d^\dagger$  anticommutators to have an extra factor of  $1/f$ .

At the same time, you must consider what would happen to the form of

***H***. You would find that

$$H = \sum_{\vec{p}, r} f E_{\vec{p}} [N_r(\vec{p}) + \bar{N}_r(\vec{p})] . \quad (268)$$

(Here, we retain the definitions

$$N_r(\vec{p}) = c_r^\dagger(\vec{p})c_r(\vec{p}) , \quad \bar{N}_r(\vec{p}) = d_r^\dagger(\vec{p})d_r(\vec{p}) \quad (269)$$

employed earlier.) One would then compute

$$H|\vec{k}, s\rangle = Hc_s^\dagger(\vec{k})|0\rangle = E_{\vec{k}}|\vec{k}, s\rangle , \quad (270)$$

where the last equality comes from using the new rescaled anticommutators (including the extra  $1/f$  factor, which cancels the extra  $f$  in the expression for  $H$ ). So, such rescalings do not affect the fact that the Hamiltonian operator gives the correct result for the energy of a single particle state.

At this point, you should ask about the normalization of our starting  $\mathcal{L}$ , since it was this normalization that determined the normalization of  $H$ . (Note that  $\mathcal{L}$  could be multiplied by an arbitrary constant without changing

the equation of motion for  $\psi$ .) In fact, we chose the normalization of  $\mathcal{L}$  so that the normalization of the  $H$  and  $\vec{P}$  operators corresponded to our usual conventions for defining what we mean by energy and momentum. This is referred to as “canonical normalization”. It will always be important to refer to canonical normalization when defining what we mean by the canonically normalized fields that define the particles (the physical eigenstates) of our theory. This seemingly trivial remark can have some profound implications once interactions and mixings between particles are incorporated into our theory.

- Starting from the contour integral representations of the  $\Delta^\pm$ , we easily obtain from Eq. (259) that

$$S^\pm(x) = -\frac{1}{(2\pi)^4} \int_{C^\pm} d^4p e^{-ip \cdot x} \frac{\not{p} + m}{p^2 - m^2}, \quad (271)$$

where the  $C^\pm$  contours are the anticlockwise paths enclosing the poles at  $p_0 = \pm E_{\vec{p}}$ .

Since  $(\not{p} \pm m)(\not{p} \mp m) = p^2 - m^2$ , the last equation can be abbreviated

in the form

$$S^\pm(x) = -\frac{1}{(2\pi)^4} \int_{C^\pm} d^4p e^{-ip \cdot x} \frac{1}{\not{p} - m}, \quad (272)$$

Note:

$$\begin{aligned} \not{a} \not{a} &= a_\mu a_\nu \gamma^\mu \gamma^\nu \\ &= a_\mu a_\nu \frac{1}{2} [\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu] = a_\mu a_\nu \frac{1}{2} [2g^{\mu\nu}] = a \cdot a. \end{aligned} \quad (273)$$

### Discussion on spin-statistics connection

- We can ask what would have happened had we quantized the Dirac system using commutators instead of anticommutators.

At the first level, i.e. directly from the expression for  $H$  and the field expansion forms but before assuming commutation or anticommutation, one obtains

$$H = \sum_{r\vec{p}} E_{\vec{p}} [c_r^\dagger(\vec{p}) c_r(\vec{p}) - d_r(\vec{p}) d_r^\dagger(\vec{p})]. \quad (274)$$

If we employ anticommutation, then we can rewrite the 2nd term as (roughly)  $dd^\dagger = -d^\dagger d + 1$  to obtain

$$\begin{aligned}
 H &= \sum_{r\vec{p}} E_{\vec{p}} [c_r^\dagger(\vec{p})c_r(\vec{p}) + d_r^\dagger(\vec{p})d_r(\vec{p}) - 1] \\
 &= \sum_{r\vec{p}} E_{\vec{p}} [N_r(\vec{p}) + \bar{N}_r(\vec{p})] - \infty. \tag{275}
 \end{aligned}$$

However, had we chosen to use commutation relations, we would have to use  $dd^\dagger = d^\dagger d + 1$ , yielding

$$\begin{aligned}
 H &= \sum_{r\vec{p}} E_{\vec{p}} [c_r^\dagger(\vec{p})c_r(\vec{p}) - d_r^\dagger(\vec{p})d_r(\vec{p}) - 1] \\
 &= \sum_{r\vec{p}} E_{\vec{p}} [N_r(\vec{p}) - \bar{N}_r(\vec{p})] - \infty. \tag{276}
 \end{aligned}$$

In both cases, we throw away the  $\infty$  to define the 0 of energy and then consider excitations relative to this. From the later result, we see that we could have eigenstates with arbitrarily negative energy (since  $\bar{N}_r = 0, 1, 2, \dots$  are all possible in the commutation case) relative to

the vacuum state. If we want a state of lowest energy, we must use anticommutators.

- Similarly, we can ask what would happen if we quantized the KG case using anticommutators. Here, even more bizarre things take place.

At the first stage, i.e. before assuming commutation or anticommutation, for the real field case, one finds

$$H = \sum_{\vec{p}} \frac{1}{2} \omega_{\vec{p}} [a^\dagger(\vec{p})a(\vec{p}) + a(\vec{p})a^\dagger(\vec{p})]. \quad (277)$$

Assuming commutation, we write  $aa^\dagger = a^\dagger a + 1$ , and find

$$H = \sum_{\vec{p}} \omega_{\vec{p}} [a^\dagger(\vec{p})a(\vec{p}) + \frac{1}{2}] = \sum_{\vec{p}} \omega_{\vec{p}} [N(\vec{p}) + \frac{1}{2}]. \quad (278)$$

We would throw away the  $\infty$  from the  $\frac{1}{2}$  and obtain the usual result.



Assuming anticommutation, we would write  $aa^\dagger = -a^\dagger a + 1$  and get

$$H = \sum_{\vec{p}} \omega_{\vec{p}} \left[ \frac{1}{2} \right]; \quad (279)$$

i.e. there would be no number operator in the expression for  $H$ !. The same sort of thing happens in the charged scalar field case.

### Causality and the spin-statistics connection

- In the scalar field case, we verified that  $[\phi(x), \phi(y)] = 0$  for space-like separation  $(x - y)^2 < 0$ . Obviously any construct from  $\phi(x)$  would commute with any other construct from  $\phi(y)$  for  $(x - y)^2 < 0$ .

Had we tried to quantize using anticommutators, we would have found  $[\phi(x), \phi(y)] \neq 0$  for  $(x - y)^2 < 0$ . Mandl-Shaw says that it would be ok if just  $[\phi(x), \phi(y)]_+ = 0$  for  $(x - y)^2 < 0$ . This is incorrect. There are many observables in nature that are linear in  $\phi$  and so it is not sufficient that bilinears in  $\phi$  commute. In any case, one would find using anticommutation that  $[\phi(x), \phi(y)]_+ \neq 0$  as well.

- In the Dirac field case, it is sufficient that

$$[\psi(x), \psi(y)]_+ = [\psi(x), \bar{\psi}(y)]_+ = [\bar{\psi}(x), \bar{\psi}(y)]_+ = 0 \quad (280)$$

for  $(x - y)^2 < 0$  and any choice of Dirac indices (suppressed above) on the two fields. This is because, in the Dirac field case, for the observables to be scalar constructs in the Dirac indices they must be bilinears of the form  $\mathcal{O}_i(x) = \bar{\psi}(x)\Gamma_i\psi(x)$ , where, in fact, there are precisely 16 independent  $\Gamma_i$  (simply the number of independent  $4 \times 4$  matrices).

Causality will be obeyed if

$$[\mathcal{O}_i(x), \mathcal{O}_j(y)] = 0 \quad (281)$$

for all choices of  $i, j = 1, \dots, 16$  when  $(x - y)^2 < 0$ .

**Homework:** Show that this is the case given Eq. (280). Problem 4.3 of Mandl-Shaw is a special case of this general result.

The only possibly non-zero anticommutator above is

$$[\psi(x), \bar{\psi}(y)]_+ = iS(x - y) = (i\partial\!\!\!/ + m)\Delta(x - y) \quad (282)$$

[see Eq. (261)] which is 0 for  $(x - y)^2 < 0$  since  $\Delta(x - y)$  is 0 for  $(x - y)^2 < 0$ .

Had we quantized the Dirac theory using commutators, not only would we have ended up with energies unbounded from below, but also we would not have been able to get causality (without the bizarre manipulations of Peskin and Schroeder, which end up violating the positive norm requirement for single particle states).

## The Fermion Feynman propagator

- We define the fermion Feynman propagator as

$$\langle 0 | T \{ \psi(x) \bar{\psi}(x') \} | 0 \rangle, \quad (283)$$

where the spinor indices on the fermionic fields are not written. Like normal ordering, the  $T$  instruction for fermionic fields is defined by including an extra  $-$  sign when two fermi fields are passed by one another:

$$T \{ \psi(x) \bar{\psi}(x') \} = \theta(t - t') \psi(x) \bar{\psi}(x') - \theta(t' - t) \bar{\psi}(x') \psi(x). \quad (284)$$

To compute the propagator, we note that

$$\begin{aligned}
 \langle 0 | \psi(x) \bar{\psi}(x') | 0 \rangle &= \langle 0 | \psi^+(x) \bar{\psi}^-(x') | 0 \rangle \\
 &= \langle 0 | [\psi^+(x), \bar{\psi}^-(x')]_+ | 0 \rangle \\
 &= iS^+(x - x'), \quad \text{and} \\
 \langle 0 | \bar{\psi}(x') \psi(x) | 0 \rangle &= iS^-(x - x'), \quad (285)
 \end{aligned}$$

where  $S^+$  and  $S^-$  were defined earlier in Eq. (261):

$$S^\pm(x) = (i\partial + m)\Delta^\pm(x)$$

The result is that

$$\langle 0 | T\{\psi(x) \bar{\psi}(x')\} | 0 \rangle = iS_F(x - x'), \quad (286)$$

where

$$S_F(x) = \theta(t)S^+(x) - \theta(-t)S^-(x) = (i\partial + m)\Delta_F(x), \quad (287)$$

where we used the earlier result:

$$\Delta_F(x) = \theta(t)\Delta^+(x) - \theta(-t)\Delta^-(x), \quad (288)$$

and the fact that when the  $i\gamma_0\partial_0$  part of  $i\partial\!\!\!/$  acts on the time theta functions one obtains  $i\gamma^0$  times

$$\delta(t)\Delta^+(x) + \delta(-t)\Delta^-(x) = \delta(t)\Delta(x) = 0 \quad (289)$$

by virtue of the fact that  $\Delta(x)$  is zero for space like  $x$ .

- Given Eq. (287), and the contour representation of  $\Delta_F$  derived earlier, we have

$$S_F(x) = \frac{1}{(2\pi)^4} \int d^4p e^{-ip\cdot x} \frac{\not{p} + m}{p^2 - m^2 + i\epsilon}, \quad (290)$$

where the  $+i\epsilon$  in the denominator once again instructs us to pass below the  $-E_{\vec{p}}$  pole and above the  $+E_{\vec{p}}$  pole in the  $p^0$  complex plane.

- To show that this is, indeed, a Green's function, we need to verify that  $(i\partial\!\!\!/ - m)S_F(x) = \delta^4(x)$ . This can be done in two ways.

1. Operate on the form given in Eq. (290). Using the fact that

$$i\partial\!\!\!/ e^{-ip\cdot x} = i\gamma^\mu \frac{\partial}{\partial x^\mu} e^{-ip_\mu x^\mu} = (i\gamma^\mu) \times (-ip_\mu) e^{-ip\cdot x} = \not{p} e^{-ip\cdot x}, \quad (291)$$

and  $\not{p} \not{p} = p^2$ , we obtain

$$\begin{aligned} (i\partial\!\!\!/ - m)S_F(x) &= \frac{1}{(2\pi)^4} \int d^4p e^{-ip\cdot x} \frac{(\not{p} - m)(\not{p} + m)}{p^2 - m^2 + i\epsilon} \\ &= \frac{1}{(2\pi)^4} \int d^4p e^{-ip\cdot x} \frac{(p^2 - m^2)}{p^2 - m^2 + i\epsilon} \\ &= \frac{1}{(2\pi)^4} \int d^4p e^{-ip\cdot x} = \delta^4(x). \end{aligned} \quad (292)$$

2. Alternatively, we can note that

$$\begin{aligned} (i\partial\!\!\!/ - m)S_F(x) &= (i\partial\!\!\!/ - m)(i\partial\!\!\!/ + m)\Delta_F(x) \\ &= (-\partial\!\!\!/\partial\!\!\!/ - m^2)\Delta_F(x) \\ &= -(\square + m^2)\Delta_F(x) \end{aligned}$$

$$= (-)^2 \delta^4(x), \quad (293)$$

using the fact that  $\Delta_F(x)$  is the Green's function for the Klein Gordon operator (taking care with the sign in  $(\square + m^2)\Delta_F(x) = -\delta^4(x)$  — see the earlier material on  $\Delta_F$  — and remembering that  $\Delta_F$  used here is defined with mass  $m$  of the Dirac particle).

- Analogously to the spin-0 discussion,  $S_F(x_1 - x_2) = \langle 0|T\{\psi(x_1)\bar{\psi}(x_2)\}|0\rangle$  incorporates simultaneously two processes:
  1. For  $t_1 > t_2$ , an electron is emitted or created at location  $x_2$  and then absorbed later at  $x_1$ .
  2. For  $t_2 > t_1$ , a positron is emitted or created at location  $x_1$  and later absorbed at  $x_2$ .

(Recall that  $\bar{\psi}$  creates a  $e^-$  and annihilates a  $e^+$ , while  $\psi$  does the reverse.)

## Interactions with the electromagnetic field

- Problem 1.2 introduced you to the minimal substitution interaction, which

in NRQM is defined by the substitutions

$$i\frac{\partial}{\partial t} \rightarrow i\frac{\partial}{\partial t} - q\phi(x), \quad -i\vec{\nabla} \rightarrow -i\vec{\nabla} - q\vec{A}, \quad (294)$$

which is summarized in 4-vector notation (after factoring away an  $i$ ) as

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + iqA_\mu. \quad (295)$$

To recognize that this is what is implied by problem 1.2, note that the final result of that problem was that

$$H = \frac{(\vec{p} - q\vec{A})^2}{2m} + q\phi, \quad (296)$$

which after doing the standard QM substitutions becomes

$$i\frac{\partial}{\partial t} - q\phi = \frac{(-i\vec{\nabla} - q\vec{A})^2}{2m} \quad (297)$$

in QM operator form.



This clearly follows from the NRQM procedure and minimal substitution rules asserted above as follows:

$$\begin{aligned}
 H &= \frac{\vec{p}^2}{2m} \rightarrow \text{using QM operator substitutions} \\
 i\frac{\partial}{\partial t} &= \frac{(-i\vec{\nabla})^2}{2m} \rightarrow \text{using minimal substitution rules} \\
 i\frac{\partial}{\partial t} - q\phi &= \frac{(-i\vec{\nabla} - q\vec{A})^2}{2m}.
 \end{aligned} \tag{298}$$

- If we follow this prescription in the Dirac Lagrangian, we get

$$\mathcal{L} = \bar{\psi}(x)(i\not{D} - m)\psi(x) = \bar{\psi}(x)(i\not{\partial} - m - q\not{A})\psi(x) = \mathcal{L}_F + \mathcal{L}_I \tag{299}$$

with (using  $q = -e$  for the electron)

$$\mathcal{L}_F = \bar{\psi}(x)(i\not{\partial} - m)\psi(x), \quad \mathcal{L}_I = e\bar{\psi}(x)\not{A}\psi(x) = e\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu(x). \tag{300}$$

$\mathcal{L}_I$  is the interaction Lagrangian density.

- To obtain the complete Lagrangian density, we must add the Lagrangian density for the noninteracting  $A_\mu$  field,  $\mathcal{L}_A = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$  (plus some gauge fixing stuff).
- **The concept of an Elementary Particle**

Now, the above interaction Lagrangian is not the only interaction Lagrangian that can be written down. One could include, for example, an anomalous magnetic moment interaction

$$\propto \bar{\psi}(x) \frac{\sigma_{\mu\nu}}{2m} \psi(x) F^{\mu\nu} \quad (301)$$

which is gauge invariant (see ensuing discussion), Lorentz invariant and so forth. This type of interaction must be included when describing the way in which a proton interacts with the  $A^\mu$  field.

However, we know that the proton's anomalous magnetic moment comes from the orbital motion of quarks around one another when the quarks are treated as having only the interaction generated by the minimal substitution rule.

Further, the proton is known to have a form factor that suppresses its interaction when probed by  $\vec{A}$  at high momentum transfer. This is also

understood as being due to the composite nature of the proton, whereas the quarks would not have such a form factor.

The minimal substitution rule implies that the fermion's interactions with the  $A^\mu$  field are such that there is no anomalous magnetic moment (at “tree-level”, i.e. before we include loop “radiative” corrections) and no form factor.

This is our definition of what we mean by an *elementary particle*. It is a particle whose interactions are generated entirely by the minimal substitution rule.

We are continuing to test the hypothesis that the leptons (the electron being one of the leptons) and quarks are truly elementary in this sense. Every time we build a new accelerator capable of higher momentum transfer probes, we look to see if the electron has a form factor that is not in agreement with that computed using the interaction generated by the minimal substitution rule to compute loop corrections to the tree-level prediction of no form factor (i.e.  $F(q^2) = 1$ , where  $F$  is the form factor and  $q$  is the momentum transfer squared).

- **Gauge invariance**

1. We know that  $\mathcal{L}_A$  is invariant under  $A_\mu \rightarrow A_\mu + \partial_\mu f(x)$ .
2. Under this substitution, we have

$$\mathcal{L}_I \rightarrow \mathcal{L}_I + e\bar{\psi}(x)\gamma^\mu\psi(x)\partial_\mu f(x), \quad (302)$$

which means a lack of gauge invariance.

3. To compensate, we must introduce a transformation for  $\psi$ . The one that works is a generalization of the “global” phase transformation related to charge to a local phase transformation (i.e.  $x$ -dependent phase):

$$\psi(x) \rightarrow e^{ief(x)}\psi(x), \quad \bar{\psi}(x) \rightarrow \bar{\psi}(x)e^{-ief(x)} \quad (303)$$

This additional phase transformation does not affect  $\mathcal{L}_I$  (the phases of  $\psi$  and  $\bar{\psi}$  just cancel one another), but because of the derivative in  $\mathcal{L}_F$  we get

$$\mathcal{L}_F \rightarrow \mathcal{L}_F - e\bar{\psi}(x)\gamma^\mu\psi(x)\partial_\mu f(x) \quad (304)$$

so that the change in  $\mathcal{L}_I$  is just compensated by the change in  $\mathcal{L}_F$  and  $\mathcal{L}$  as a whole is unchanged by the local gauge transformation.

4. There is an intimate connection between local gauge invariance and a property of perturbative field theory called “renormalizability”. Basically,

a theory must be renormalizable if we are to be able to use perturbation theory to compute predictions in the context of the field theory.

As a result, a cardinal rule of constructing a field theory is that it should exhibit local gauge invariance under gauge transformations related to any vector field, including  $A_\mu$  of electromagnetism as well as similar vector fields associated with color QCD and weak forces.

**Homework:** Repeat the minimal substitution rule for the complex field KG Lagrangian and define the local gauge transformation for  $\phi$  that leaves the full interacting  $\mathcal{L}$  invariant. This is roughly the 1st 1/2 of Problem 5.4 of Mandl-Shaw.

**Homework:** Problem 4.5 of Mandl-Shaw. We already did Problems 4.1 and 4.2 in class.

# The Electromagnetic Field (again)

- The quantization procedure of Chapter 1 hides the Lorentz-invariance of the theory — the decomposition of the fields into transverse and longitudinal components is frame-dependent.
- We need an explicitly Lorentz-covariant formulation in order to prove that we can carry out calculations to arbitrary order in perturbation theory.
- We will employ the 4-vector potential  $A^\mu(x) = (\phi(x), \vec{A}(x))$ .

However, it contains more degrees of freedom than the system actually possesses and the extra dof must be removed by imposing constraints.

- The formulation presented now is actually equivalent to the formulation of Chapter 1.

The Classical Fields

- We employ

$$F^{\mu\nu} = \partial^\mu A^\nu(x) - \partial^\nu A^\mu(x), \quad (305)$$

where  $\mu$  labels the row and  $\nu$  labels the column. This is opposite the Mandl-Shaw convention (which is a terrible excursion from the norm); my convention is the more usual convention. For example, see the book by Ryder for correct conventions. I will be following Ryder's notation in what follows. Sorry, but I don't want you to get used to an uncommon convention set.

Then,  $F^{\mu\nu}$  reduces to

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix} \quad (306)$$

To get the signs right in the above, we had to be very careful. For example,

$$F^{01} = \partial^0 A^1 - \partial^1 A^0 = -\left(-\frac{\partial}{\partial x^0} A^1 - \vec{\nabla}^1 \phi\right) = -E^1 \quad (307)$$

where we used  $\partial^1 = \frac{\partial}{\partial x_1} = -\frac{\partial}{\partial x^1} = -\vec{\nabla}^1$ . More generally,  $E^i = -F^{0i}$  and

$$\epsilon^{ijk} B^k = -F^{ij}.$$

- We include a charge-current density,  $j^\mu(x) = (\rho(x), \vec{j}(x))$  in terms of which Maxwell's equations take the form

$$\partial_\mu F^{\mu\nu}(x) = j^\nu(x) \quad (308)$$

$$\partial^\lambda F^{\mu\nu}(x) + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0. \quad (309)$$

Since  $F^{\mu\nu}$  is antisymmetric, the first of these equations immediately implies that  $j^\mu$  must be conserved:  $\partial_\nu j^\nu(x) = 0$ . Note that the first of these equations differs from Mandl-Shaw, because I don't like their definition of  $F^{\mu\nu}$ .

- A convenient way of reformulating Eq. (309) is to use the dual tensor

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}, \quad (310)$$

in which case Eq. (309) is equivalent to

$$\partial_\mu \tilde{F}^{\mu\nu} = 0. \quad (311)$$



- $F^{\mu\nu}$  of Eq. (305) automatically obeys (309), while substituting the form of  $F^{\mu\nu}$  from Eq. (305) into Eq. (308) results in the eom

$$\square A^\nu - \partial^\nu(\partial_\mu A^\mu) = j^\nu. \quad (312)$$

- These equations are Lorentz-covariant and they are also invariant under the gauge transformation

$$A^\mu(x) \rightarrow A'^\mu = A^\mu(x) + \partial^\mu f(x). \quad (313)$$

- The 4 equations contained in (312) are the eom of

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j_\mu A^\mu \quad (314)$$

if the four components of  $A_\nu$  are treated as independent fields in the variational technique.

- To check the equations of motion for the above Lagrangian, we must write

out  $\mathcal{L}$  in terms of the  $A^\mu$  fields:

$$\mathcal{L} = -\frac{1}{4}(\partial_\rho A_\sigma - \partial_\sigma A_\rho)g^{\rho\alpha}g^{\sigma\beta}(\partial_\alpha A_\beta - \partial_\beta A_\alpha) - j_\alpha A^\alpha, \quad (315)$$

where we have been careful to use dummy summation indices in  $\mathcal{L}$ . Now, think of dealing with each  $A_\nu$  component as the independent field.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -\frac{1}{2}(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho})g^{\rho\alpha}g^{\sigma\beta}(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \\ &= -g^{\mu\alpha}g^{\nu\beta}(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \\ &= -(\partial^\mu A^\nu - \partial^\nu A^\mu) \end{aligned} \quad (316)$$

where, in the first equality we have used the fact that the  $\alpha, \beta$  dummy indices would work just like the  $\rho, \sigma$  indices, thereby providing a factor of 2. The 2nd equality comes from using the  $\alpha \leftrightarrow \beta$  antisymmetry of  $(\partial_\alpha A_\beta - \partial_\beta A_\alpha)$ . We next compute

$$-\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = \square A^\nu - \partial^\nu(\partial_\mu A^\mu) \quad (317)$$

which is the LH side of Eq. (312). Meanwhile,

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = j^\nu \quad (318)$$

so that the eom becomes

$$\begin{aligned} 0 &= -\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu A_\nu)} \right) + \frac{\partial \mathcal{L}}{\partial A_\nu} \\ &= \square A^\nu - \partial^\nu (\partial_\mu A^\mu) - j^\nu \end{aligned} \quad (319)$$

which is the desired result, Eq. (312), and obviously holds for each value of  $\nu$  independently of every other value of  $\nu$ .

- We note that  $\mathcal{L}$  is clearly Lorentz invariant by construction. Further, although  $\mathcal{L}$  is not gauge invariant, a gauge transformation does not affect the equations of motion derived from  $\mathcal{L}$  under certain provisos. To show this and get the provisos, we note that  $F^{\mu\nu}$  is obviously gauge invariant while, under a gauge transformation,

$$j_\mu A^\mu \rightarrow j_\mu A^\mu + j_\mu \partial^\mu f$$

$$\begin{aligned}
&= j_\mu A^\mu + \partial^\mu(j_\mu f) - (\partial^\mu j_\mu) f \\
&= j_\mu A^\mu + \partial^\mu(j_\mu f)
\end{aligned} \tag{320}$$

by virtue of current conservation  $\partial^\mu j_\mu = 0$ . The final form is equivalent to the original  $\mathcal{L}$  since this total 4-derivative can be added for free to  $\mathcal{L}$  without changing the equations of motion, provided that  $j_\mu f$  only depends upon the fields and not their derivatives. (This is problem 2.1 of Mandl-Shaw.)

**Proof:**

$f$  must depend upon the fields in order for it to have any impact upon the equations of motion, so let us assume it has some arbitrary dependence upon the fields  $A_\nu$ , **but no dependence on the derivatives of the fields.**  $j_\mu$  is assumed to depend upon other fields (so-called matter fields such as electrons, quarks, ....). So, let us write the addition to  $\mathcal{L}$  as  $\delta\mathcal{L} = \partial^\alpha \Lambda_\alpha(A_\beta)$ . Then, we first note that

$$\delta\mathcal{L} = \partial^\alpha \Lambda_\alpha = \frac{\partial \Lambda_\alpha}{\partial A_\beta} \partial^\alpha A_\beta \tag{321}$$

by the chain rule. Then, the eom contribution from  $\delta\mathcal{L}$  takes the form:

$$\begin{aligned}
 \frac{\partial\delta\mathcal{L}}{\partial A_\nu} - \partial_\mu \left( \frac{\partial\delta\mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) &= \partial^\alpha \left( \frac{\partial\Lambda_\alpha}{\partial A_\nu} \right) - \partial_\mu \left( \frac{\partial\Lambda_\alpha}{\partial A_\beta} g^{\mu\alpha} g^\nu_\beta \right) \\
 &= \partial^\alpha \left( \frac{\partial\Lambda_\alpha}{\partial A_\nu} \right) - \partial_\mu \left( \frac{\partial\Lambda^\mu}{\partial A_\nu} \right) \\
 &= 0
 \end{aligned} \tag{322}$$

In getting the form of the 2nd term on the LH side of the 1st equality, we used the explicit dependence of  $\delta\mathcal{L}$  on  $\partial_\mu A_\nu$  written in Eq. (321). The final cancellation just follows from dummy index relabeling.

Note that the cancellation between the two eom terms would fail if  $\Lambda_\alpha$  depended on the derivatives of the  $A^\mu$  fields. For example, if it depended only on the field derivatives,  $\frac{\partial\delta\mathcal{L}}{\partial A_\nu} = 0$  while  $\partial_\mu \left( \frac{\partial\delta\mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) \neq 0$ .

So, this implies that if we want to change the equation of motion (which we must in order to have an invertible kernel, i.e. a well-defined Green's function/propagator) gauge fixing must involve derivatives of the  $A^\mu(x)$  fields.

## 2nd quantization

- The conjugate fields are computed as

$$\pi^\nu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\nu} = -F^{0\nu} \quad (323)$$

where there are 4 terms contributing to the derivative, all of which contribute equally. This result implies that  $\pi^0 = 0$ , and so imposition of canonical quantization conditions in the usual way is impossible for the  $A^0$  field.

- The other thing that goes wrong for the  $\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$  Lagrangian form is the following.

One wants to be able to define the Feynman propagator as being the Greens function for the Kernel defined by the equation of motion in the absence of any interactions. Recall that the equation of motion for the above  $\mathcal{L}$  is (for the  $j = 0$ , free field, case)

$$\square A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0 \quad (324)$$

which we rewrite as

$$(\square g_\mu^\nu - \partial^\nu \partial_\mu) A^\mu \equiv K_\mu^\nu A^\mu = 0 \quad (325)$$

The  $K$  is the Kernel and the Greens function should obey

$$K_\alpha^\nu(x) G_\mu^\alpha(x - y) = g_\mu^\nu \delta^4(x - y). \quad (326)$$

Let us transform to momentum space. One then gets the equation

$$K_\alpha^\nu(k) G_\mu^\alpha(k) = -(k^2 g_\alpha^\nu - k^\nu k_\alpha) G_\mu^\alpha(k) = g_\mu^\nu. \quad (327)$$

In other words  $G(k)$  should be basically the inverse of  $K(k)$ .

I will now show that there is no solution to this equation, implying that  $K$  does not have an inverse and that a Feynman propagator cannot be defined. Let us suppose that there is a solution  $G(k)$ . Lorentz covariance implies that the most general form it can take is

$$G_\mu^\alpha(k) = a(k^2) k^\alpha k_\mu + b(k^2) g_\mu^\alpha. \quad (328)$$

Plug this into Eq. (327) to obtain the consistency requirement

$$-k^2 b(k^2) g_\mu^\nu + b(k^2) k^\nu k_\mu = g_\mu^\nu, \quad (329)$$

to which there is no solution. (Note how the  $a(k^2)$  stuff cancelled.)

This lack of an inverse is because  $K$  has many zero eigenvalues. Any  $A^\mu \propto \partial^\mu f$  (i.e. a pure gauge transformation) obeys

$$K_\mu^\nu(x) \partial^\mu f(x) = 0! \quad (330)$$

An operator with zero eigenvalues has no inverse.

Thus, we see that in order to get an invertible Kernel from our Lagrangian we must specify a Lagrangian that has no zero eigenvalues, which in the present context means that it **should not** be invariant under gauge transformations — the Lagrangian should be that obtained after fixing some particular gauge.

- An alternative form for  $\mathcal{L}$  that does allow canonical procedures and does



generate an invertible equation of motion Kernel is

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) - j_\mu A^\mu. \quad (331)$$

As we will see shortly, this  $\mathcal{L}$  corresponds to having assumed the gauge-fixing condition  $\partial^\mu A_\mu = 0$ , called the Lorentz gauge.

For this  $\mathcal{L}$  we find

$$\pi^\nu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\nu} = -\dot{A}^\nu, \quad (332)$$

which are all non-vanishing so that the canonical procedure can be applied for each  $A^\nu$  field independently.

- The eom for this alternative  $\mathcal{L}$  are

$$\square A^\nu(x) = j^\nu(x). \quad (333)$$

- Is there a justification for making this change in  $\mathcal{L}$ .

We note that (312) ( $\square A^\nu - \partial^\nu(\partial_\mu A^\mu) = j^\nu$ ) reduces to (333) if we take

$$\partial_\mu A^\mu = 0. \quad (334)$$

This gauge condition must be imposed *after the fact*, i.e. after the 2nd quantization procedure based on  $\mathcal{L}$  of (331), as a constraint or subsidiary condition in order to maintain consistency with the original Maxwell equation (312).

- In the classical theory, starting from an arbitrary  $A_\mu(x)$  we can always find a gauge transformation that gives  $\partial^\mu A'_\mu = 0$ . We simply choose the function  $f$  in  $A'_\mu(x) = A_\mu(x) + \partial_\mu f(x)$  so that

$$\partial_\mu A^\mu(x) + \square f(x) = 0 \quad (335)$$

which can be solved for  $f(x)$ , but not uniquely since additional gauge transformations obeying  $\square f = 0$  would not further change  $\partial_\mu A^\mu$ .

- The condition  $\partial^\mu A_\mu = 0$  is called the Lorentz gauge condition.
- The computational advantage of this choice over the previously discussed  $\vec{\nabla} \cdot \vec{A} = 0$  choice is that the eom are so much simpler. In particular, for  $j^\mu = 0$ , the eom reduce to

$$\square A^\nu(x) = 0, \quad (336)$$

i.e. each  $A^\nu$  obeys a massless Klein-Gordon equation and we can 2nd quantize by taking over results from the KG discussion.

- We write

$$A^\nu(x) = A^{\nu+}(x) + A^{\nu-}(x) \quad (337)$$

with (assuming real polarization vectors,  $\epsilon_r$ )

$$A^{\nu+} = \sum_{r\vec{k}} \left( \frac{1}{2V\omega_{\vec{k}}} \right)^{1/2} \epsilon_r^\nu(\vec{k}) a_r(\vec{k}) e^{-ik \cdot x}, \quad (338)$$

$$A^{\nu-} = \sum_{r\vec{k}} \left( \frac{1}{2V\omega_{\vec{k}}} \right)^{1/2} \epsilon_r^\nu(\vec{k}) a_r^\dagger(\vec{k}) e^{+ik \cdot x} \quad (339)$$

where  $k^0 = \omega_{\vec{k}} = |\vec{k}|$  and  $r = 0, 1, 2, 3$  to describe 4 fields. We choose the  $\epsilon_r^\nu$  to be real and obey

$$\epsilon_r(\vec{k}) \cdot \epsilon_s(\vec{k}) = -\zeta_r \delta_{rs}, \quad r, s = 0, 1, 2, 3, \quad \sum_r \zeta_r \epsilon_r^\mu(\vec{k}) \epsilon_r^\nu(\vec{k}) = -g^{\mu\nu}, \quad (340)$$

where  $\zeta_0 = -1$ ,  $\zeta_{1,2,3} = +1$ . The two extra polarization states compared to transverse (Coulomb) gauge provide a covariant description of the instantaneous Coulomb interaction. A particularly convenient choice that realizes the above conditions is:

$$\epsilon_0(\vec{k}) = n^\mu \equiv (1, 0, 0, 0) \quad (341)$$

$$\epsilon_r(\vec{k}) = (0, \vec{\epsilon}_r(\vec{k})), \quad r = 1, 2, 3 \quad (342)$$

$$\vec{\epsilon}_3(\vec{k}) = \frac{\vec{k}}{|\vec{k}|} \quad (343)$$

$$\vec{k} \cdot \vec{\epsilon}_r(\vec{k}) = 0, \quad r = 1, 2 \quad (344)$$

$$\vec{\epsilon}_r(\vec{k}) \cdot \vec{\epsilon}_s(\vec{k}) = \delta_{rs}, \quad r, s = 1, 2, 3. \quad (345)$$

The  $\epsilon_{1,2}$  are the transverse polarization vectors,  $\epsilon_3$  is the longitudinal polarization vector, and  $\epsilon_0$  is referred to as the scalar or time-like polarization vector.

For later use, we note that (using 4-vector notation)

$$\epsilon_3(\vec{k}) = \frac{\mathbf{k} - (\mathbf{k} \cdot \mathbf{n})\mathbf{n}}{[(\mathbf{k} \cdot \mathbf{n})^2 - k^2]^{1/2}}. \quad (346)$$

We have kept  $k^2 \neq 0$  so that we can describe off-shell photons as well as on-shell photons using this set up.

### Covariant 2nd Quantization

- We take  $j^\mu = 0$  for the free-field case.
- We require

$$[A^\mu(\vec{x}, t), \dot{A}^\nu(\vec{y}, t)] = -ig^{\mu\nu} \delta^3(\vec{x} - \vec{y}) \quad (347)$$

and other commutators = 0. Note that covariance requires the  $g^{\mu\nu}$ , so we must learn how to deal with it. For  $\mu = \nu = j = 1, 2, 3$ ,  $g^{\mu\nu} = -1$ , and the resulting  $-ig^{jj} = +i$  is exactly like the  $+i$  of the KG single field case. The implication and need for the  $-i$  for the  $\mu = \nu = 0$  case will emerge later.

- Let us implement these conditions using KG or KG-like commutators for the  $a, a^\dagger$  appearing in the “Fourier” decomposition written earlier. In particular, we will employ

$$[a_r(\vec{k}), a_s^\dagger(\vec{p})] = \zeta_r \delta_{rs} \delta_{\vec{k}\vec{p}} \quad (348)$$

with  $[a, a]$  and  $[a^\dagger, a^\dagger]$  commutators = 0. The fact that  $\zeta_0 = -1$  appears above is required to get the “wrong”, i.e. negative, sign, i.e.  $-i$ , for the  $\mu = \nu = 0$  field commutator of Eq. (347).

- Using the above, and following the KG procedures, we will obviously obtain

$$[A^\mu(x), A^\nu(y)] = iD^{\mu\nu}(x - y), \quad (349)$$

where

$$D^{\mu\nu}(x) = -g^{\mu\nu} \Delta(x, m = 0). \quad (350)$$

The Feynman propagator will similarly be given by

$$\langle 0|T\{A^\mu(x)A^\nu(y)\}|0\rangle = iD_F^{\mu\nu}(x - y) \quad (351)$$

where

$$D_F^{\mu\nu}(x) = -g^{\mu\nu} \Delta_F(x, m = 0) = \frac{-g^{\mu\nu}}{(2\pi)^4} \int d^4k \frac{e^{-ik \cdot x}}{k^2 + i\epsilon}. \quad (352)$$

- That the above is the correct result can be verified by simply checking that the Feynman propagator is indeed the Green's function for the equation of motion, which for the Lagrangian we are employing,  $\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu)$ , means it should obey

$$\square[iD_F^{\mu\nu}(x)] = -ig^{\mu\nu} \delta^4(x - y) \quad (353)$$

which is obviously the case. Note that as above for  $\mu = \nu = i$  the sign of the RHS is exactly what we had for a scalar field, whereas we have the opposite sign for  $\mu = \nu = 0$  — this sign is required for a covariant form above and elsewhere.

### Gupta-Bleuler approach

- The basic steps are:

1. Interpret  $a_r(\vec{k})$  as absorption operators for all  $r = 0, 1, 2, 3$ .
2. Interpret  $a_r^\dagger(\vec{k})$  as creation operators for all  $r = 0, 1, 2, 3$ .
3. Define the vacuum state by  $a_r(\vec{k})|0\rangle = 0$  for all  $\vec{k}$  and  $r = 0, 1, 2, 3$ .
4. Define the 1-photon states by  $|1_{\vec{k}r}\rangle = a_r^\dagger(\vec{k})|0\rangle$ .
5. Compute

$$H = \int d^3\vec{x} : [\pi^\mu \dot{A}_\mu - \mathcal{L}] : \quad (354)$$

by substituting the expansions and using the properties of the  $\epsilon_s$ 's of Eqs. (340)-(345) to obtain

$$H = \sum_{s\vec{p}} \omega_{\vec{p}} \zeta_s a_s^\dagger(\vec{p}) a_s(\vec{p}) . \quad (355)$$

6. The “bad” sign for the  $s = 0$  photons is actually ok since

$$H|1_{\vec{k}r}\rangle = \sum_{\vec{p}s} \omega_{\vec{p}} \zeta_s a_s^\dagger(\vec{p}) a_s(\vec{p}) a_r^\dagger(\vec{k})|0\rangle = +\omega_{\vec{k}} a_r^\dagger(\vec{k})|0\rangle, \quad r = 0, 1, 2, 3 \quad (356)$$

is always positive. The extra  $\zeta_r \delta_{rs}$  in the commutators compensates the extra  $\zeta_s$  in the  $H$  expression.



## 7. The number operators should be defined as

$$N_r(\vec{k}) = \zeta_r a_r^\dagger(\vec{k}) a_r(\vec{k}) \quad (357)$$

for the same reason.

- But, there is a remaining problem. One finds states of negative norm:

$$\langle 1_{\vec{k}r} | 1_{\vec{k}r} \rangle = \langle 0 | a_r(\vec{k}) a_r^\dagger(\vec{k}) | 0 \rangle = \zeta_r \langle 0 | 0 \rangle = \zeta_r \quad (358)$$

which is  $-1$  for  $r = 0$ . To eliminate this problem and to eliminate the extra  $r = 0, 3$  non-transverse photon states, we must somehow bring back in the Lorentz gauge condition. We cannot do so as an operator identity because that would be incompatible with the commutation relations since

$$[\partial_\mu A^\mu(x), A^\nu(y)] = i \partial_\mu D^{\mu\nu}(x - y) \neq 0 \quad (359)$$

for the form of  $D^{\mu\nu}$  obtained earlier as given in Eq. (350).

- The solution is to restrict the physical states by requiring

$$\partial_\mu A^{\mu+}(x) |\Psi\rangle = 0, \quad (360)$$

which involves only absorption operators. The adjoint of this equation is

$$\langle \Psi | \partial_\mu A^{\mu-}(x) = 0. \quad (361)$$

Then, for any state obeying this condition we have

$$\langle \Psi | \partial_\mu A^\mu(x) | \Psi \rangle = \langle \Psi | \partial_\mu A^{\mu+}(x) + \partial_\mu A^{\mu-}(x) | \Psi \rangle = 0 \quad (362)$$

so that the Lorentz condition holds in the Quantum Mechanical expectation value sense for the allowed physical states. If we rewrite (360) in terms of the  $a$  operators, we find

$$[a_3(\vec{k}) - a_0(\vec{k})] | \Psi \rangle = 0, \quad \text{for all } \vec{k}, \quad (363)$$

and the adjoint

$$\langle \Psi | [a_3^\dagger(\vec{k}) - a_0^\dagger(\vec{k})] = 0. \quad (364)$$

These are constraints on the linear combinations of longitudinal and scalar photons allowed in the state  $|\Psi\rangle$ . It places no constraint on the physical transverse photons present in a given state.

**Derivation: We note that**

$$\begin{aligned}
\partial_\mu A^{\mu+} &= \sum_{\vec{k}, r} \frac{1}{\sqrt{2V|\vec{k}|}} \epsilon_r^\mu(\vec{k}) a_r(\vec{k}) \partial_\mu e^{-ik \cdot x} \\
&= -i \sum_{\vec{k}, r} \frac{1}{\sqrt{2V|\vec{k}|}} k_\mu \epsilon_r^\mu(\vec{k}) a_r(\vec{k}) e^{-ik \cdot x} \\
&\stackrel{k \cdot \epsilon_{1,2} = 0}{=} -i \sum_{\vec{k}, r=0,3} \frac{1}{\sqrt{2V|\vec{k}|}} k_\mu \epsilon_r^\mu(\vec{k}) a_r(\vec{k}) e^{-ik \cdot x} \\
&= -i \sum_{\vec{k}} \frac{1}{\sqrt{2V|\vec{k}|}} e^{-ik \cdot x} \left( k_0 \epsilon_0^0(\vec{k}) a_0(\vec{k}) + k_0 \epsilon_3^0(\vec{k}) a_3(\vec{k}) \right. \\
&\quad \left. - \vec{k} \cdot \vec{\epsilon}_0(\vec{k}) a_0(\vec{k}) - \vec{k} \cdot \vec{\epsilon}_3(\vec{k}) a_3(\vec{k}) \right) \\
&= -i \sum_{\vec{k}} \frac{1}{\sqrt{2V|\vec{k}|}} e^{-ik \cdot x} \left( k_0 a_0(\vec{k}) + 0 a_3(\vec{k}) - 0 a_0(\vec{k}) - \vec{k} \cdot \hat{k} a_3(\vec{k}) \right) \\
&= -i \sum_{\vec{k}} \frac{1}{\sqrt{2V|\vec{k}|}} e^{-ik \cdot x} |\vec{k}| \left( a_0(\vec{k}) - a_3(\vec{k}) \right). \tag{365}
\end{aligned}$$

In going from the 2nd to 3rd line, I used the fact that  $k \cdot \epsilon_{1,2} = 0$  for the transverse polarization states. The 4th line is obtained by simply writing the 3rd line very explicitly. The 5th line is obtained by noting that  $\epsilon_0^0 = 1$ ,  $\epsilon_3^0 = 0$ ,  $\vec{\epsilon}_0 = 0$  and  $\vec{\epsilon}_3 = \hat{k}$ . Since  $\partial_\mu A^{\mu+} |\Psi\rangle = 0$  for all  $x$ , and since each

$\vec{k}$  gives a different spatial dependence, we must have Eq. (363).

- The result of this condition is apparent if we compute the vacuum expectation value of the energy of an allowed state. From the form of  $H$ , this calculation will include

$$\langle \Psi | a_3^\dagger(\vec{k}) a_3(\vec{k}) - a_0^\dagger(\vec{k}) a_0(\vec{k}) | \Psi \rangle = \langle \Psi | a_3^\dagger(\vec{k}) [a_3(\vec{k}) - a_0(\vec{k})] | \Psi \rangle = 0 \quad (366)$$

where the first equality is obtained using the adjoint form of the constraint. Thus, we have

$$\langle \Psi | H | \Psi \rangle = \langle \Psi | \sum_{\vec{k}} \sum_{r=1,2} \omega_{\vec{k}} a_r^\dagger(\vec{k}) a_r(\vec{k}) | \Psi \rangle \quad (367)$$

so that only the transverse photons contribute to the expectation value of the energy. The same applies to other observables.

As far as negative norm states are concerned, it is not possible to have a physical state defined by  $a_0^\dagger(\vec{k})|0\rangle$  since

$$[a_0(\vec{k}) - a_3(\vec{k})] a_0^\dagger(\vec{k}) |0\rangle = a_0(\vec{k}) a_0^\dagger(\vec{k}) |0\rangle = -|0\rangle \neq 0. \quad (368)$$

An allowed state would have to be a superposition state such as  $\frac{1}{\sqrt{2}}[a_0^\dagger(\vec{k}) - a_3^\dagger(\vec{k})]|0\rangle$  which would obey the subsidiary condition as follows:

$$[a_0(\vec{k}) - a_3(\vec{k})]\frac{1}{\sqrt{2}}[a_0^\dagger(\vec{k}) - a_3^\dagger(\vec{k})]|0\rangle = \frac{1}{\sqrt{2}}[a_0(\vec{k})a_0^\dagger(\vec{k}) + a_3(\vec{k})a_3^\dagger(\vec{k})]|0\rangle = \frac{1}{\sqrt{2}}[-1 + 1]|0\rangle = 0. \quad (369)$$

However, this kind of state is not actually a state at all. One way to see this is that the norm of the above sample state is computed as:

$$\frac{1}{2}\langle 0|[a_0(\vec{k}) - a_3(\vec{k})][a_0^\dagger(\vec{k}) - a_3^\dagger(\vec{k})]|0\rangle = \frac{1}{2}[a_0(\vec{k})a_0^\dagger(\vec{k}) + a_3(\vec{k})a_3^\dagger(\vec{k})]|0\rangle = \frac{1}{2}[-1 + 1]|0\rangle = 0. \quad (370)$$

In other words, adding in some allowed superposition of longitudinal and time-like polarization states does not actually yield anything. You must have some transverse stuff present in order to have a state with non-zero norm. A typical state would be of the form that you will consider for Problem 5.3 of Mandl and Shaw:

$$|\Psi\rangle = \left(1 + c[a_0^\dagger(\vec{k}) - a_3^\dagger(\vec{k})]\right) |\Psi_T\rangle, \quad (371)$$

where  $|\Psi_T\rangle$  is a state containing some number of transversely polarized photons. From the computation of Eq. (370), it should be apparent that

$$\langle\Psi|\Psi\rangle = \langle\Psi_T|\Psi_T\rangle. \quad (372)$$

- Although only the transverse photons matter for physical states, the presence of the time-like and longitudinal photons is needed in a virtual sense when interactions mediated by the electromagnetic field enter the picture.

Of the two extra ( $r = 0, 3$ ) dof for each  $\vec{k}$ , one is removed by the subsidiary state condition (363) and the other corresponds to arbitrariness in the choice of the Lorentz gauge (i.e. the specific way of implementing the Lorentz condition). For instance, one can define the vacuum state  $|0\rangle$  as containing no  $r = 0$  or  $r = 3$  states. This is the usual choice. So for the so-called “asymptotic states” where particles are not interacting, only the transverse photons need to be kept.

But one could also define  $|0\rangle$  as containing an appropriate mixture of  $r = 0, 3$  states so as to still satisfy Eq. (363). The many different such possibilities correspond to different Lorentz gauge choices, i.e. different choices of  $f$  obtained by shifting  $f \rightarrow f + \Delta f$  with  $\square \Delta f = 0$ .

Once a vacuum choice is made, one simply adds in the transverse photons in the obvious way.

- Once we introduce charged matter fields, we can have virtual photon

propagation between charged fields. The propagator form will depend upon the gauge choice.

However, so long as the initial and final “free-particle-like” states of a process obey the subsidiary conditions (i.e. are properly composed), at the very end of the calculation the result will be independent of the gauge chosen.

- The better way to handle all of this is through path integrals. We will return to this in the section of notes labelled QFT-III.

### The GB photon propagator interpretation

- Writing

$$D_F^{\mu\nu} = \frac{1}{(2\pi)^4} \int d^4k D_F^{\mu\nu}(k) e^{-ik \cdot x} \quad (373)$$

we have

$$D_F^{\mu\nu}(k) = \frac{-g^{\mu\nu}}{k^2 + i\epsilon} = \frac{1}{k^2 + i\epsilon} \sum_r \zeta_r \epsilon_r^\mu(\vec{k}) \epsilon_r^\nu(\vec{k}). \quad (374)$$

- If we separate out the transverse part  $r = 1, 2$  of this sum to define

$_T D_F^{\mu\nu}(k)$  then we can write

$$D_F^{\mu\nu}(k) =_T D_F^{\mu\nu}(k) +_C D_F^{\mu\nu}(k) +_R D_F^{\mu\nu}(k) \quad (375)$$

where the first of the latter two guys corresponds to an instantaneous Coulomb potential interaction and the last “remainder” component can be shown to make no contribution to any physical process by virtue of current conservation. Please read the Mandl-Shaw material on this.

In this way, we can understand that this covariant treatment is the same as the Coulomb gauge treatment of Chapter 1, which also had real propagating transverse photons and an instantaneous Coulomb potential interaction (we did not show this in lecture), and nothing more.

In any case, we shall employ the fully covariant form of  $D_F^{\mu\nu}$  in computing scattering processes and such.

**Problems 5.1 and 5.3 of Mandl-Shaw are assigned at this point.**