A minimal introduction to Quantum Field Theory

Arpit Das[1,2†],

1 School of Maths, University of Edinburgh, Edinburgh, EH9 3FD, U.K.
2 Higgs Centre for Theoretical Physics, University of Edinburgh, Edinburgh EH8 9YL, U.K.

† arpit.das@ed.ac.uk

Abstract

These lecture notes were prepared for the School on Quantum Field Theory organized by ICTP: Physics Without Frontiers – Physics for Bangladesh program. They aim to provide a concise introduction to relativistic field theory and its canonical quantization in $(1+3)$ dimensions.

We begin with a discussion of the Lorentz and Poincaré groups, together with their Lie algebras. We then turn to classical field theory, where we derive the Euler–Lagrange equations, introduce Noether's theorem, and illustrate the deep connection between symmetries and conservation laws.

Next, we quantize the complex scalar field, examine its various propagators, and explore how the notion of causality emerges in field-theoretic settings. Following this, we briefly describe how symmetries act on states in Hilbert space through unitary operators, and how these symmetry operators in turn act on operator-valued fields.

We then move on to the Dirac equation. Starting with the spinor field, we study how it transforms under spinor representations, together with its discrete symmetries such as parity. Upon quantization, we see how the spin–statistics theorem arises: the requirement of a Hamiltonian bounded from below forces the use of anticommutation relations for fermionic fields. We conclude with a short discussion of propagators in the Dirac theory.

Throughout these lectures, we have restricted ourselves to free, non-interacting theories.

Contents

1 Introduction **5**
- 1.1 Why quantum field theory? 5
- 1.2 Locality and causality 6
- 1.3 Fluctuations of particle number 6
- 1.4 Relativistic propagator and causality 7
- 1.5 From particles to fields and then to operators 9

2 Lorentz Invariance **10**
- 2.1 2d Rotational Invariance 10

2.2 Lorentz group 11
2.3 Translations and the Poincaré Group 15
2.4 Realizing transformations on fields 15
2.4.1 Scalar fields 16
2.4.2 Vector fields 17
2.5 Representation of groups and Lie algebras 18

3 Classical Field Theory 19
3.1 Euler–Lagrange equations 20
3.1.1 Action of a real scalar field 20
3.2 Noether's theorem 23
3.2.1 Complex scalar field 23
3.2.2 Lorentz Transformations and Angular Momentum 29
3.3 Hamiltonian methods 31

4 Canonical quantization 33
4.1 Canonical commutators 34
4.2 Classical theory 35
4.3 Operator-valued Fourier expansions 36
4.4 Algebra of creation and annihilation operators 37
4.5 Hamiltonian and the energy of the vacuum 39
4.6 Normal ordering 40
4.7 Single-particle states 41
4.8 The Fock space 45
4.9 Position-space particle creation 46

5 Propagators and Causality 47
5.1 Lorentz-invariant integration measures 48
5.2 Commutators and the light-cone 49
5.3 Particle propagation 50
5.4 Feynman propagator 52
5.5 The $i\epsilon$ prescription 55

6 Symmetries in QM 57

7 Dirac equation 59
7.1 The spinor representation 60
7.1.1 Spinor fields 61
7.2 Constructing the Dirac action 63
7.2.1 Equations of Motion 65
7.3 Weyl spinors 66
7.3.1 The Weyl Equation 67

7.4 γ^5 and Parity 68
7.5 Chiral interaction terms 70
7.6 Symmetries and Conserved Currents 72
7.7 Plane wave solutions 74

8 Quantization of Dirac equation 76
8.1 A Glimpse at the Spin–Statistics Theorem 76
8.2 The Hamiltonian 78
8.3 Fermionic Quantization 79
8.3.1 Fermi–Dirac Statistics 80
8.4 Propagators 81
8.4.1 The Feynman Propagator 82

A Translational Symmetry and Fourier Transform 83

B Very brief complex analysis 85
B.1 Cauchy's three theorems 85
B.2 Rules for real integrals via residues 85
B.3 Branch points and branch cuts (logs and square roots) 86

C Fun with delta distribution identities 87
C.1 A nice delta identity 87
C.2 Derivative of the Heaviside step function 88
C.3 Another delta identity 88
C.4 Delta function pole identity 89

D More on representations 89
D.1 $G = SU(2)$ 90
D.2 $G = SO(3)$ 91
D.3 $G = U(1)$ 92

Acknowledgements

These lecture notes draw heavily on a number of excellent sources: lecture notes, textbooks, and videos—and thus contain little that is truly new beyond some reorganization and restructuring. I remain deeply indebted to the following references, from which I have learned the workings of QFT and continue to learn with every passing day. Last but not the least, I am grateful to the students who participated in this school on QFT and made this an enriching

experience with their curious questionings and discussions.

Note: A word of caution – the references listed below adopt different metric conventions. Be mindful of this, as occasional sign differences will appear across sources. Also, some references discuss directly the *path integral quantisation* approach instead of the *canonical quantisation* approach.

Lecture notes

- Advanced Quantum Theory by Nabil Iqbal
- Quantum Field Theory by David Tong
- Quantum Field Theory I by Niklas Beisert
- Quantum Field Theory I by Edgar Shaghoulian
- Relativistic Quantum Field Theory I by Daniel Harlow
- Quantum Field Theory: part one by John McGreevy
- Quantum Field Theory I by Ashoke Sen

Books

- Quantum field theory for the gifted amateur by Tom Lancaster, Stephen Blundell
- Quantum Field Theory and the Standard Model by Matthew Schwartz
- Quantum Field Theory: An Integrated Approach by Eduardo Fradkin
- Quantum Field Theory: From Basics to Modern Topics by François Gelis
- Quantum Field Theory: The Why, What and How by Thanu Padmanabhan

Videos

- Relativistic Quantum Field Theory I by Hong Liu: https://www.youtube.com/watch?v=xQFgk-nEihg&list=PLUl4u3cNGP61AV6bhf4mB3tCyWQrI_uU5
- Quantum Field Theory I by Ashoke Sen: https://www.youtube.com/watch?v=Uk_YQkonLBU&list=PLt4li39O1V1WGB_mI3S7ia1GbnJLOf8t1
- Quantum Field Theory by Tobias Osborne: https://www.youtube.com/watch?v=T58H6ofIOpE&list=PLDfPUNusx1EpRs-wku83aqYSKfR5fFmfS

Units and convention

We will adopt natural units $\hbar = c = 1$ [1] and the mostly-minus Minkowski metric:

$$\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1), \tag{0.1}$$

i.e. the $(+,-,-,-)$ convention. For the Levi-Civita symbol we will have: $\epsilon^{0123} = 1$ and $\epsilon_{0123} = -1$.

1 Introduction

1.1 Why quantum field theory?

This course is about quantum field theory (QFT). It is both the conceptual framework behind our most precise descriptions of nature and the day–to–day language of modern quantum physics. However, we still don't have an understanding of it – we can't even *properly* define it. We just know that it works :). For instance, the anomalous magnetic moment of the electron, computed from QFT, matches very well to its experimental value:

$$\begin{aligned} a_e^{\text{th}} &= 0.001159652181643 \pm 0.764 \times 10^{-12}, \\ a_e^{\text{exp}} &= 0.00115965218073 \pm 0.28 \times 10^{-12}. \end{aligned} \tag{1.1}$$

QFT can be subtle (the same phenomenon often admits multiple equivalent viewpoints) and technically heavy; that diversity is part of the appeal and also the challenge. Now let us ask *why* QFT is needed – QFT is required when physics is simultaneously *small in size* and *fast in motion*. Typical situations include:

1. *Light.* Treating electromagnetic radiation quantum mechanically leads to photons—massless quanta that always travel at c. A faithful quantum description of light is a field theory. Note that, light has been a guiding principle for many foundational theories: Electromagnetism, Special Relativity, Quantum Mechanics and now QFT. Yet, we are still developing our understanding of light [2]

2. *Relativistic beams.* Particle colliders (e.g. the LHC) accelerate particles to near light speed and scatter them. Predicting cross-sections and decay distributions is a QFT problem. Practically, these collisions serve as microscopes for very short distances.

3. *Very hot systems.* At sufficiently high temperature typical velocities become relativistic (e.g. the early universe), again demanding field-theoretic tools.

[1] At times, for clarity, we shall keep $\hbar$ and c in various expressions.

[2] Roughly, about a decade ago, it was revealed that the *photon* can be understood as the *goldstone boson* for the *spontaneous symmetry breaking* of a *higher-form* symmetry.

Roughly speaking, the attempt to *nicely* combine quantum mechanics and special relativity leads to the formalism of *quantum fields*. In these introductory lectures, we shall learn what is called *perturbative QFT* – where we will be interested in local fields/operators and in their correlation functions. We will study *free QFT* where everything will be non-interacting and potentials will be quadratic.

Note: Along these basic lectures we will also come across some mathematical concepts which we will try to motivate physically as much as possible. Additionally, everything will be discussed at a physics level of rigour and hence my humble apologies to the very mathematically inclined people.

1.2 Locality and causality

Now let us unpack the word *nicely* in the above paragraph. By this we mean, that our theory of quantum fields should obey the following constraints: *causality, locality* and *unitarity* – for it to be a physically sensible theory. First, let us understand how locality and causality are related. Consider a system of N positive charges (each with charge q) with net charge being: $Q = Nq$. Now imagine that you can destroy a charge at point A and then simultaneously create another charge at point B. So, the net charge is still Q, right? Hence, the total charge is still conserved. Is this argument okay or, is it flawed?

You will quickly notice that, causality is violated with the above type of charge conservation. Why? Since, simultaneity is not a causal concept. So, above type of conservation processes are not compatible with causality. The way to remedy this is to have local conservation laws:

$$\partial_\mu j^\mu = \partial_t j^0 + \vec{\nabla} \cdot \vec{j} = 0, \tag{1.2}$$

which suggests that locality is a very crucial concept to prevent superluminal information transfer [3]. Note, that this is necessary condition for causality since QM is *acausal* yet it has a local conservation law like above.

§1: Which local conservation law in QM is hinted to above? Derive it.

One might try to retrofit single-particle quantum mechanics to be "relativistic" but new physics enters that cannot be captured by a fixed-particle-number framework. A quick heuristic makes this plausible.

1.3 Fluctuations of particle number

In nonrelativistic classical mechanics, a free particle of mass m has energy

$$E_{\text{non-rel}} = \frac{|\vec{p}|^2}{2m}. \tag{1.3}$$

[3] You must have seen the above in the *Electrodynamics* course.

Relativistically, we have:

$$E_{\text{rel}} = \sqrt{|\vec{p}|^2c^2 + m^2c^4} = mc^2 + \frac{|\vec{p}|^2}{2m} + \mathcal{O}(|\vec{p}|^4), \tag{1.4}$$

where $\vec{p}$ is the 3-vactor corresponding to momentum and m is the rest-mass with c being the speed of light. Above, we have expanded about $|\vec{p}| \to 0$. Note, rest mass corresponds to rest energy mc^2. Quantum mechanics allows energy fluctuations, and injecting enough energy can create *quanta*; in particular, producing a particle–antiparticle pair of mass m requires energy roughly equal to [4],

$$\Delta E \approx 2mc^2. \tag{1.5}$$

Now confine the system to a box of size L. The Heisenberg relation gives

$$\Delta q\, \Delta p \gtrsim \frac{\hbar}{2} \quad \Rightarrow \quad \Delta p \gtrsim \frac{\hbar}{2L}. \tag{1.6}$$

For ultra-relativistic momenta – that is expanding (1.4) about $|\vec{p}| \to \infty$, we get: $E \approx |\vec{p}|c$, so the corresponding energy uncertainty is

$$\Delta E \sim c\, \Delta p \sim \frac{\hbar c}{2L}. \tag{1.7}$$

Equating (1.7) with the pair-creation threshold (1.5) yields the characteristic length

$$L \sim \frac{\hbar}{mc}, \tag{1.8}$$

namely the (reduced) *Compton wavelength*,

$$\lambda_{\text{C}} = \frac{\hbar}{mc}. \tag{1.9}$$

For electrons, $\lambda_{\text{C}} \sim 10^{-12}\,\text{m}$, roughly. At or below this scale, vacuum fluctuations readily produce particle–antiparticle pairs; a description with a fixed number of particles becomes inadequate and a field-based framework – with infinitely many degrees of freedom – is the natural language.

1.4 Relativistic propagator and causality

To understand why particles must be replaced by fields, let us look more closely at relativistic propagation. Although the non–interacting theory based on wavefunctions with the energy given as (1.4) is both relativistic and quantum, it seems to allow information to propagate faster than light. For clarity, let us work in $1+1$ dimensions and consider the propagator

$$G(x_f, t_f; x_i, t_i) := \langle x_f | e^{-iH_0(t_f - t_i)} | x_i \rangle\,. \tag{1.10}$$

Here $|x_i\rangle$ and $|x_f\rangle$ are eigenstates of the position operator. In the momentum basis, the wavefunctions of such eigenstates are

$$\langle p|x\rangle = e^{-ipx}\,. \tag{1.11}$$

[4] This was not possible in non-relativisitc QM since there we didn't have $E = mc^2$.

The physical interpretation of the propagator is that $|G|^2$ is the probability of finding the particle at position x_f at time t_f, given that it was at x_i at time t_i. We focus on the situation where $x_f > x_i$, $t_f > t_i$, and $t_f - t_i < x_f - x_i$. This implies, (x_f, t_f) and (x_i, t_i) are spacelike separated, i.e. outside each other's lightcones. Naively, no signal slower than or equal to light can propagate between them. Yet, as we will show, the propagator in this situation is nonzero.

We can evaluate Eq. (5.21) by inserting a complete set of momentum eigenstates:

$$G(x_f, t_f; x_i, t_i) = \int_{-\infty}^{\infty} \frac{dp}{2\pi}\, e^{ip(x_f - x_i) - i\sqrt{p^2+m^2}\,(t_f - t_i)}\,. \tag{1.12}$$

By translation invariance, set $t_i = 0$, $x_i = 0$, relabel $t_f = t$, $x_f = x$, so the propagator becomes

$$G(x, t) = \int_{-\infty}^{\infty} \frac{dp}{2\pi}\, e^{ipx - i\sqrt{p^2+m^2}\,t}\,. \tag{1.13}$$

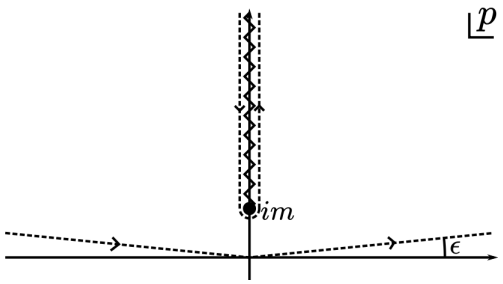


Figure 1: Deforming the contour in the complex p-plane. The dashed line is the original contour, which can be deformed into the upper contour wrapping around the branch cut on the imaginary axis.

This integral is subtle due to oscillatory behavior. To regulate it, we tilt the contour in the complex p-plane, slightly above the real axis (see Fig. 1). At large $|p|$, the integrand behaves as

$$e^{ipx - i|p|t} \approx e^{-t|p| + ipx} \qquad (p \gg 0)\,, \tag{1.14}$$

$$e^{ipx - i|p|t} \approx e^{-t|p| - ipx} \qquad (p \ll 0)\,. \tag{1.15}$$

This ensures convergence by Jordan's lemma. By Cauchy's theorem, the contour can be rotated to run along the imaginary axis $p = i\lambda$, from $\lambda \in [m, \infty)$. Evaluating this gives,

$$G(x, t) = i \int_m^{\infty} \frac{d\lambda}{2\pi}\, e^{-\lambda x} \left(e^{t\sqrt{\lambda^2 - m^2}} - e^{-t\sqrt{\lambda^2 - m^2}} \right) \tag{1.16}$$

$$= \frac{i}{\pi} \int_m^{\infty} d\lambda\, e^{-\lambda x} \sinh\!\left(\sqrt{\lambda^2 - m^2}\, t \right). \tag{1.17}$$

The integrand (ignoring the prefactor i) is strictly positive for $\lambda > m$, so indeed the propagator is nonzero outside the lightcone. By using the monotonicity of $\sinh y$ for $y > 0$ and then ignoring the negative exponential, we can bound the integral as:

$$\int_m^{\infty} d\lambda\, e^{-\lambda x} \sinh\!\left(\sqrt{\lambda^2 - m^2}\, t \right) < \tfrac{1}{2} \int_m^{\infty} d\lambda\, e^{-\lambda(x-t)} = \frac{e^{-m(x-t)}}{2(x-t)}\,, \tag{1.18}$$

which implies

$$0 < |G(x,t)| < \frac{e^{-m(x-t)}}{2\pi(x-t)}\,. \tag{1.19}$$

Thus, for a massive relativistic particle, propagation outside the lightcone is exponentially suppressed, but never exactly zero. In a relativistic setting, this is deeply problematic: if signals could propagate outside the lightcone, then by boosting to another frame one could send messages backwards in time. In interacting theories this is even worse: a signal sent outside the lightcone could return, allowing communication with one's own past. Such violations of *causality* are unacceptable in physics, and therefore demand a resolution.

1.5 From particles to fields and then to operators

In classical mechanics we track finitely many generalised coordinates and conjugate momenta $\{q_a(t), p_a(t)\}$. Canonical quantization promotes these to operators which obey Heisenberg uncertainty relations like (1.6). To handle processes that create/annihilate quanta, we instead start from *fields*. For example, a real scalar field is a map [5]:

$$\phi : \mathbb{R}^{1,3} \to \mathbb{R}, \tag{1.20}$$

that assigns a number to each spacetime point $(t, \vec{x})$; more generally one can have complex scalars or vector fields such as the electric field $\vec{E}(t,\vec{x})$. Because $\phi(t,\vec{x})$ takes a value at every spatial point, the theory has infinitely many degrees of freedom—exactly what we need to describe variable particle number.

In classical field theory, a field $\phi(t,\vec{x})$ comes with a conjugate momentum, denoted by $\pi(t,\vec{x})$, and the pair obeys equations of motion derived from an action. To obtain a *quantum* field theory, we canonically quantize this system, promoting

$$\phi(t,\vec{x}) \;\to\; \hat{\phi}(t,\vec{x}), \qquad \pi(t,\vec{x}) \;\to\; \hat{\pi}(t,\vec{x}),$$

with the usual equal–time commutators [6]. This procedure—quantizing *fields* rather than a finite set of coordinates $q_a(t)$—is historically called "second quantization" [7]. A crucial practical point is that there is now an *operator at each point in space*: we have infinitely many quantum degrees of freedom.

Once we have this framework in hand, a number of deep facts about the world fall into place:

1. There are many distinct yet perfectly indistinguishable copies of a given elementary particle species (e.g. electrons).

2. Particle *statistics* (behavior under exchange) is tied to *spin* (behavior under rotations).

[5] Here, $\mathbb{R}^{1,3}$ is the physical spacetime.

[6] This step is analogous to imposing uncertainty relations and by *equal time* we mean that we define a Hilbert space of *states* at each point in time. We will discuss this in much detail later.

[7] The name is mostly historical; nothing is being quantized twice.

3. Antiparticles are real physical entities.

4. Particle number is not sacred: particles can be created and destroyed.

5. Most profoundly, interactions we call "forces" can be understood as arising from the *exchange of quanta*.

Each of these statements will be derived within QFT; some we have already met in simpler guises, others will appear later.

2 Lorentz Invariance

We now review Lorentz invariance. You will have already seen most of this in your Relativity course,or in the pre-lecture courses, so this section is largely a review, but with an emphasis relevant for us. Before diving in, let us warm up quickly with ordinary rotational invariance.

2.1 2d Rotational Invariance

Consider the Euclidean plane $\mathbb{R}^2$ with coordinates (x, y). A rotated system (x', y') is related by

$$\vec{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \equiv R(\theta)\vec{x}. \tag{2.1}$$

where $\theta \in [0, 2\pi)$. Here, $R(\theta) \in SO(2)$ [8]. Some immediate facts follow:

1. The components x' and y' are not independent but mix.

2. The vector length is invariant:

$$(x')^2 + (y')^2 = (x\cos\theta + y\sin\theta)^2 + (y\cos\theta - x\sin\theta)^2 = x^2 + y^2. \tag{2.2}$$

3. More generally, the dot product is invariant. For $\vec{v} = (v_x, v_y)$, $\vec{w} = (w_x, w_y)$,

$$\vec{v}\cdot\vec{w} = \vec{v}^T\vec{w}. \tag{2.3}$$

   Invariance under rotations requires

$$\vec{v}'\cdot\vec{w}' = \vec{v}^T R^T R\,\vec{w}, \tag{2.4}$$

   implying, $R^T R = 1$, the defining property of orthogonal matrices.

These 2×2 matrices form a group under the binary operation: matrix multiplication.

[8] Here, $SO(2)$ is the group of 2×2 *orthogonal matrices* with determinant equal to unity.

§2: Show that the following holds:

- Closure: $R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2)$
- Associativity: $(R(\theta_1)R(\theta_2))R(\theta_3) = R(\theta_1)(R(\theta_2)R(\theta_3))$
- Identity: Show that there exists an identity element $R(\theta) \in SO(2)$ which does nothing and denote it by $\mathbb{1}$; find the corresponding θ that does this.
- Inverse: Show that there exists $R(\theta)^{-1} \in SO(2)$ for every $R(\theta) \in SO(2)$ such that their composition under the group operation yields the identity element above.

Note that since θ is a continuous parameter and varies smoothly, one can show that the map: $\theta \mapsto R(\theta)$ is a smooth map. Due to this the matrix multiplication and group inverse (or, matrix inverse here) are also smooth operations in (θ_1, θ_2) and in θ respectively. This makes $SO(2)$ a *Lie group* – which is essentially a group which is also a smooth *manifold*.

2.2 Lorentz group

The postulates of special relativity are:

1. The laws of physics are the same in all inertial frames.
2. The speed of light is the same in all inertial frames. In $(1+1)$ dimensions, this implies: $\frac{dx'}{dt'} = c = \frac{dx}{dt}$. Squaring, we get the invariant line element [9]:

$$c^2dt^2 - dx^2 = c^2dt'^2 - dx'^2 = 0. \tag{2.5}$$

A Lorentz transformation is a 4×4 matrix Λ satisfying

$$\Lambda^T \eta \Lambda = \eta, \qquad \Lambda^{\mu}{}_{\rho}\, \eta_{\mu\nu} \Lambda^{\nu}{}_{\sigma} = \eta_{\rho\sigma}. \tag{2.6}$$

So, Lorentz transformations are a set of linear transformations on the spacetime coordinates. These matrices above form the group Lorentz group, denoted by, $O(1,3)$.

§3: Using (2.6) show that Λs indeed form a group.

Now starting with the defining relation (2.6) and taking determinants yields

$$(\det \Lambda)^2 \det \eta = \det \eta \quad \Rightarrow \quad \det \Lambda = \pm 1. \tag{2.7}$$

Thus, Lorentz transformations split into two disjoint sets according to the sign of their determinant. The subgroup with $\det \Lambda = +1$ are continuously connected to the identity since the

[9] Note, this is not a *complete* derivation of the invariant line element yet! Can you tell why? What is missing?

identity lives here. Let us focus on this first. So, to begin, we can consider Lorentz transformations in a small neighbourhood of identity [10]. So, consider transformations of the form,

$$\Lambda = e^{\epsilon M}, \qquad \epsilon \text{ small.} \tag{2.8}$$

Expanding about $\epsilon \to 0$ gives,

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \epsilon M^{\mu}{}_{\nu} + \mathcal{O}(\epsilon^2). \tag{2.9}$$

where δ^{μ}_{ν} is the Kronecker delta. We can view it as just the identity 4×4 matrix and hence it is Lorentz invariant.

§4: We have seen that $\eta_{\mu\nu}$ and δ^{μ}_{ν} are Lorentz invariant. What other non-trivial tensors are invariant under infinitesimal Lorentz transformations?

Plugging above into (2.6) gives

$$(\delta^{\mu}{}_{\alpha} + M^{\mu}{}_{\alpha})\,\eta_{\mu\nu}(\delta^{\nu}{}_{\beta} + M^{\nu}{}_{\beta}) = \eta_{\alpha\beta}. \tag{2.10}$$

Simplifying, we obtain the condition

$$M_{\alpha\beta} + M_{\beta\alpha} = 0, \tag{2.11}$$

i.e. $M_{\mu\nu}$ is antisymmetric (with both indices lowered).

§5: For an $n \times n$ anti-symmetric matrix, how many independent components are there?

There are six linearly independent antisymmetric matrices in $(1+3)$ dimensions. Denote a basis by $M_{\rho\sigma}$, antisymmetric in its labels. An explicit representation is [11]

$$(M_{\rho\sigma})^{\mu}{}_{\nu} = \eta_{\rho\nu}\delta^{\mu}_{\sigma} - \eta_{\sigma\nu}\delta^{\mu}_{\rho}. \tag{2.12}$$

Here μ, ν are spacetime indices of the 4×4 matrix – which label the rows and columns of the matrix M, while ρ, σ label which *generators* we are discussing. Since $M_{\rho\sigma} = -M_{\sigma\rho}$, there are indeed 6 generators.

Let us check explicit examples. For $\rho = 0, \sigma = 1$ and $\rho = 0, \sigma = 2$, we have,

$$(M_{01})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = (-M_{10})^{\mu}{}_{\nu}, \qquad (M_{02})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = (-M_{20})^{\mu}{}_{\nu}. \tag{2.13}$$

[10] We can do this because recall, we are dealing with a *Lie group*.

[11] Note, that $\eta_{\mu\nu}$ and δ^{μ}_{ν} are the only 2-index Lorentz invariant objects and since $M_{\rho\sigma}$ is anti-symmetric in its indices we cannot put both ρ and σ together as indices of the Minkowski metric and hence this gives the index structure as in (2.12) upto an overall normalisation.

respectively. Other components follow analogously. If you write (2.12) with all indices lowered (including the spacetime ones), you get:

$$(M_{\rho\sigma})_{\mu\nu} = \eta_{\rho\nu}\eta_{\sigma\mu} - \eta_{\rho\mu}\eta_{\sigma\nu} \tag{2.14}$$

and in the above way, even the (μ, ν) indices are anti-symmetric and you get an anti-symmetric matrix for each $M_{\rho\sigma}$ matrix.

One can now show that the commutators of the generators follow (from their definition):

$$[M_{\mu\nu}, M_{\rho\sigma}] = \eta_{\nu\rho}M_{\mu\sigma} - \eta_{\mu\rho}M_{\nu\sigma} - \eta_{\nu\sigma}M_{\mu\rho} + \eta_{\mu\sigma}M_{\nu\rho}. \tag{2.15}$$

This is the Lorentz algebra. This is what is called a *Lie algebra* for the Lorentz group and since it was obtained by expanding about the group identity, this defines a vector space at the identity – more precisely, known as the tangent space at the identity [12].

Now we can move away from the identity by considering a finite Lorentz transformation which is like exponentiating as before but no longer with infinitesimal parameters. So, we consider exponential of an arbitrary sum over all of these possible matrices $\mathcal{M} = \omega_{\rho\sigma}M^{\rho\sigma}$, where the parameters ω are finite parameters.

$$\Lambda(\omega_{\rho\sigma}) = \exp\left(\omega_{\rho\sigma}M^{\rho\sigma}\right), \tag{2.16}$$

where $\omega_{\rho\sigma} = -\omega_{\sigma\rho}$ are real parameters.

Physical interpretation: The generators M_{0i} correspond to boosts along the spatial directions – there are three linearly independent boosts. For example,

$$\Lambda(\omega_{01}) = \exp(\omega_{01}M^{01}) = \begin{pmatrix} \cosh(\omega_{01}) & -\sinh(\omega_{01}) & 0 & 0 \\ -\sinh(\omega_{01}) & \cosh(\omega_{01}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.17}$$

This is precisely the Lorentz boost in the x^1 direction. The parameter ω_{01} is the rapidity, related to the boost velocity v by $\tanh\omega_{01} = v$.

The generators M_{ij} with purely spatial indices correspond to spatial rotations – there are three linearly independent purely spatial rotations. For instance,

$$(M_{12})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{exponentiates to}} \Lambda(\omega_{12}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\omega_{12} & \sin\omega_{12} & 0 \\ 0 & -\sin\omega_{12} & \cos\omega_{12} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.18}$$

[12] Vector space structure is easy to understand since now we have matrix addition also. The term *algebra* refers to the fact that this vector space comes with the structure of a *commutator* that satisfies the so-called *Jacobi identity*: $[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0$. Additionally, the vector space should be closed under the commutator structure.

a rotation in the x^1–x^2 plane about the x^3-axis. Thus the familiar rotation group $SO(3)$ is a subgroup of the Lorentz group $O(1,3)$.

§6: Derive (2.17) and (2.18).

Discrete transformations: Transformations with $\det\Lambda = -1$ include time reversal,

$$T : (x^0, x^1, x^2, x^3) \mapsto (-x^0, x^1, x^2, x^3), \qquad \Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.19}$$

and parity,

$$P : (x^0, x^1, x^2, x^3) \mapsto (x^0, -x^1, -x^2, -x^3), \qquad \Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.20}$$

Note, these are not connected to the identity. Let us now briefly discuss some subgroups of the Lorentz group. Focusing on the time-time component in (2.6) we get: $(\Lambda^0{}_0)^2 - \Lambda^i{}_0\Lambda^i{}_0 = 1$ implying $\Lambda^0{}_0 \geq 1$ or, $\Lambda^0{}_0 \leq -1$. Lorentz transformations with $\det = 1$ and $\det = -1$ will be called *proper* and *improper* transformations respectively. While transformations with $\Lambda^0{}_0 \geq 1$ are called *orthochronous* transformations and for $\Lambda^0{}_0 \leq -1$, we get the *non-orthochronous* transformations. So, we get the following table:

Component	$\det\Lambda$	$\Lambda^0{}_0$	Name	Contains
1	$+1$	≥ 1	Proper orthochronous	Identity, boosts, spatial rotations
2	-1	≤ -1	Improper non-orthochronous	Time reversal Λ_T
3	-1	≥ 1	Improper orthochronous	Parity Λ_P
4	$+1$	≤ -1	Proper non-orthochronous	$\Lambda_P\Lambda_T$ (space and time reflection)

Table 2.2.1: The four disconnected components of $O(1,3)$ in the $(+,-,-,-)$ convention.

Note that, the proper Lorentz subgroup is denoted by $SO(1,3)$ [13] and the proper orthochronous Lorentz subgroup is denoted by $SO^+(1,3)$ [14].

Note: From here on, whenever we will say Lorentz transformation or Lorentz group we will mostly be focusing on $SO^+(1,3)$ or the proper orthochronous ones unless otherwise specified.

[13] The S here stands for the word "special" which simply means that the matrices have unit determinant.

[14] Here, '+' denotes $\Lambda^0{}_0 \geq 1$.

2.3 Translations and the Poincaré Group

So far, we have discussed Lorentz transformations, which generalize the idea of rotations to a relativistic spacetime. Next, we also consider spacetime translations, i.e.

$$x^\mu \longrightarrow x'^\mu = x^\mu - a^\mu, \tag{2.21}$$

where a^μ is a constant 4-vector. Translations also preserve the invariant line element. These along with the Lorentz transformations describe the symmetries of empty flat spacetime. Concretely, the time component a^0 corresponds to a translation of the origin of time (asserting that the laws of physics are the same now and in the future), while the spatial components correspond to a translation of the spatial origin (asserting that the laws of physics are the same in this room as in the next).

Combining translations with Lorentz transformations, we arrive at the *Poincaré group*, which is a 10-parameter group of spacetime symmetries:

$$x^\mu \longrightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu - a^\mu. \tag{2.22}$$

The generators of this group are the six antisymmetric Lorentz generators $M_{\rho\sigma}$ together with the four translation generators P_μ and as we will see a^μ serves as the parameter for the generators of translations.

The full Poincaré algebra is then given by:

$$[M_{\mu\nu}, M_{\rho\sigma}] = \eta_{\nu\rho} M_{\mu\sigma} - \eta_{\mu\rho} M_{\nu\sigma} - \eta_{\nu\sigma} M_{\mu\rho} + \eta_{\mu\sigma} M_{\nu\rho}, \tag{2.23}$$

$$[P_\mu, M_{\nu\rho}] = \eta_{\mu\nu} P_\rho - \eta_{\mu\rho} P_\nu, \tag{2.24}$$

$$[P_\mu, P_\nu] = 0. \tag{2.25}$$

Equation (2.23) is the Lorentz algebra we already encountered. The second relation (2.24) shows how translations transform under Lorentz transformations [15], and the last line (2.25) states that spacetime translations commute – which they should.

This algebra encodes the fundamental spacetime symmetries underlying all of special relativity and any generic QFT.

2.4 Realizing transformations on fields

We have seen how coordinate transformations act on the spacetime coordinates x, written compactly as

$$x' = \Lambda x - a, \tag{2.26}$$

Now let us discuss how these transformations act on fields.

[15] This can also be taken as the defining property of a 4-vector.

2.4.1 Scalar fields

Consider a scalar field $\phi(x)$ which will transform under coordinate transformations; in other words, $\phi(x)$ furnishes a *representation* of the Poincaré group. We now make this precise. From the definition of a scalar, we have,

$$\phi'(x') = \phi(x). \tag{2.27}$$

Translations: Consider a pure translation by a constant four–vector a^μ:

$$x'^\mu = x^\mu - a^\mu. \tag{2.28}$$

Then

$$\phi'(x') = \phi(x) \quad \Longrightarrow \quad \phi'(x-a) = \phi(x), \tag{2.29}$$

which leads to,

$$\phi'(x) = \phi(x+a). \tag{2.30}$$

For an infinitesimal translation, expand the right–hand side to first order in a:

$$\phi'(x) = \phi(x) + a^\mu \frac{\partial}{\partial x^\mu}\phi(x) + \mathcal{O}(a^2). \tag{2.31}$$

Hence the variation is

$$\delta\phi(x) = a^\mu \partial_\mu \phi(x). \tag{2.32}$$

For a finite translation, we can exponentiate the generator as before:

$$\phi'(x) = \exp\left(a^\mu P_\mu\right)\phi(x), \tag{2.33}$$

and by matching the linear term we read off

$$P_\mu = \frac{\partial}{\partial x^\mu}. \tag{2.34}$$

From (2.30), we have that for a general Poincaré transformation [16]:

$$\phi'(x) = \phi(\Lambda^{-1}x) \tag{2.35}$$

Lorentz transformations: Now consider Lorentz transformations. Write

$$x'^\mu = \Lambda^\mu{}_\nu\, x^\nu = \left[\exp\left(\omega_{\rho\sigma}\, M^{\rho\sigma}\right)\right]^\mu{}_\nu\, x^\nu, \tag{2.36}$$

with antisymmetric real parameters $\omega_{\rho\sigma} = -\omega_{\sigma\rho}$. For an infinitesimal transformation,

$$x'^\mu = x^\mu + \delta x^\mu = x^\mu + \omega_{\rho\sigma}\,(M^{\rho\sigma})^\mu{}_\nu\, x^\nu + \mathcal{O}(\omega^2). \tag{2.37}$$

Eq.(2.35) implies, at the same spacetime point x,

$$\phi'(x) = \phi(x) - \delta x^\mu\, \partial_\mu\phi(x) + \mathcal{O}(\omega^2) = \phi(x) - \omega_{\rho\sigma}(M^{\rho\sigma})^\mu{}_\nu x^\nu \partial_\mu\phi(x) + \mathcal{O}(\omega^2). \tag{2.38}$$

[16] By an abuse of notation, sometime we will denote arbitrary Poincaré transformations also as Λ.

Using the explicit coordinate–space representation of the Lorentz generators as in (2.12), we get,

$$(M^{\rho\sigma})^{\mu}{}_{\nu}x^{\nu}\partial_{\mu} = \left(\eta^{\sigma\mu}\delta^{\rho}{}_{\nu} - \eta^{\rho\mu}\delta^{\sigma}{}_{\nu}\right)x^{\nu}\partial_{\mu} \tag{2.39}$$

$$= \eta^{\sigma\mu}x^{\rho}\partial_{\mu} - \eta^{\rho\mu}x^{\sigma}\partial_{\mu} \tag{2.40}$$

$$= x^{\rho}\partial^{\sigma} - x^{\sigma}\partial^{\rho}. \tag{2.41}$$

Therefore,

$$\phi'(x) = \phi(x) + \omega_{\rho\sigma}\left(x^{\rho}\partial^{\sigma} - x^{\sigma}\partial^{\rho}\right)\phi(x) \equiv \phi(x) + \omega_{\rho\sigma}\,L^{\rho\sigma}\phi(x), \tag{2.42}$$

where we have defined the orbital Lorentz generators acting on scalars [17],

$$L^{\rho\sigma} = x^{\rho}\partial^{\sigma} - x^{\sigma}\partial^{\rho}. \tag{2.43}$$

For a finite Lorentz transformation continuously connected to the identity,

$$\phi'(x) = \exp\left(\omega_{\rho\sigma}L^{\rho\sigma}\right)\phi(x). \tag{2.44}$$

Now you can check that P_{μ} and $L_{\rho\sigma}$ as obtained above explicitly satisfy the Poincaré algebra as given in (2.23)-(2.25).

§7: Check above.

2.4.2 Vector fields

Vector fields transform as follows:

$$A'^{\mu}(x') = \Lambda^{\mu}{}_{\nu}A^{\nu}(x) \tag{2.45}$$

One can now go through the same analysis as before to determine the form of the Lorentz generator that acts on the vector fields.

§8: Show that the above generator becomes [a]:

$$L^{\rho\sigma} = M^{\rho\sigma} + x^{\rho}\partial^{\sigma} - x^{\sigma}\partial^{\rho} \tag{2.46}$$

[a] Wow! now we see that the $L_{\rho\sigma}$ picks up a spacetime independent piece: $M_{\rho\sigma}$. So, in addition to the orbital piece we now have the intrinsic piece which we will later identify as *spin* of the field. This is the classical analogue of the spin angular momentum whose existence is a consequence of relativity unlike as we were taught that spin is a quantum phenomenon. This is partially true since we were taught this fact while studying quantum mechanics which is a non-relativistic theory!

The extra M here acts on the space of $A^{\mu}(x)$ at a single point x and rotates it just as though it was a vector, the other parts act on the coordinates x similar to the scalar field case.

[17] You might recall from classical mechanics that this is exactly how angular momentum looked like: $\vec{L} = \vec{r} \times \vec{p}$ but now written as derivative operators acting on fields!

2.5 Representation of groups and Lie algebras

An interesting thing to note is that both sets of generators $M_{\alpha\beta}$s and $L_{\rho\sigma}$s satisfy the Lorentz algebra as given in (2.23). However, there is a crucial difference. Note that, $M_{\alpha\beta}$s are linear maps which act on spacetime coordinates and hence are just matrices while $L_{\rho\sigma}$s are linear maps which act on physical fields and hence will be differential operators. So, $M_{\alpha\beta}$ furnish a representation of the Lorentz algebra on the coordinate vectors while $L_{\rho\sigma}$ furnish a representation of the Lie algebra on the scalar fields. In the former case, the vector space on which $M_{\alpha\beta}$ acts is the space of spacetime coordinates and hence $M_{\alpha\beta}$ furnishes a four dimensional representation. While, in the latter case, the vector space on which $L_{\rho\sigma}$ acts is the space of all functions $\mathcal{F}$ of the form $\phi : \mathbb{R}^{1,3} \to \mathbb{R}$ which is infinite dimensional and hence $L_{\rho\sigma}$ furnishes an infinite-dimensional representation.

There is another related concept to the above. Consider the Lorentz group and,

$$R(\Lambda) \equiv \Lambda, \quad \text{and,} \quad R(\Lambda)x = \Lambda x \tag{2.47}$$

$$\phi'(x) \equiv (U(\Lambda)\phi)(x) \equiv \phi(\Lambda^{-1}x) \tag{2.48}$$

Now by Ex.(3), you have already shown that Λs form the Lorentz group and hence $R(\Lambda)$s form what is called the *fundamental representation* of the Lorentz group which is a four dimensional since they act on a four dimensional vector space – the coordinate space. Now you can show the following for $U(\Lambda)$:

§9: Show the following

1. Linearity: For every $\phi, \psi \in \mathcal{F}$ (where $\mathcal{F}$ is the space of all functions from $\mathbb{R}^{1,3} \to \mathbb{R}$) and $a, b \in \mathbb{C}$, show that $(U(\Lambda)(a\phi + b\psi))(x) = (aU(\Lambda)\phi + bU(\Lambda)\psi)(x)$
2. Identity: For every $\phi \in \mathcal{F}$, there exists a unique Λ such that $(U(\Lambda)\phi)(x) = \phi(x)$. What is this Λ?
3. Homomorphism: Take Λ_1, Λ_2 in the Lorentz group and $\phi \in \mathcal{F}$. Show that $(U(\Lambda_1)U(\Lambda_2)\phi)(x) = (U(\Lambda_1\Lambda_2)\phi)(x)$. Hint: It might be useful to denote $U(\Lambda_2)\phi \equiv \psi$ and then evaluate $(U(\Lambda_1)U(\Lambda_2)\phi)(x)$.
4. Inverse: For every $\Lambda \in O(1,3)$ there exists Λ^{-1} such that $(U(\Lambda)U(\Lambda^{-1})\phi)(x) = \phi(x)$. This implies, $U(\Lambda)^{-1} = U(\Lambda^{-1})$.

Ex.(9) defines a group representation. In technical terms, given a group G and a vector space V consider $GL(V)$ as the space of invertible linear maps from $V \to V$ (which also is a group under the binary operation: composition of linear maps). Then, a representation ρ of the group is a group homomorphism:

$$\rho : G \to GL(V), \quad \rho(g) : V \mapsto V$$

where for every $g \in G$, $\rho(g)$ is a linear map from $V \to V$. The idea of representation is that it being a group homomorphism (structure-preserving maps) preserves the group structure and helps us to, in some sense, *linearise* the group action which otherwise is quite an abstract concept. Hence, we end up with a problem in linear algebra which is much more tractable.

Analogously, we can define representation of a Lie algebra as follows. If ρ is a linear map and satisfies:

$$\rho([X,Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X)$$

where X, Y belong to the Lie algebra then ρ is a representation of the Lie algebra. We see that both $M_{\alpha\beta}$s and $L_{\rho\sigma}$s satisfy the Lorentz algebra and hence are different representations of the Lorentz algebra.

3 Classical Field Theory

We now want to construct a Lagrangian for a classical scalar field $\phi(x) = \phi(t, \vec{x})$. To motivate this, recall the situation in classical mechanics where we have generalized coordinates $q_a(t)$. Now imagine placing these degrees of freedom on a cubic lattice with N sites, and then take $N \to \infty$ so that the lattice approaches the continuum $\mathbb{R}^3$. In this limit we identify

$$q_a(t) \longrightarrow \phi(t, \vec{x}), \qquad \sum_{a=1}^{N} \longrightarrow \int d^3x.$$

Thus, the spatial coordinate $\vec{x}$ plays the role of the index a, labelling the degrees of freedom.

This motivates us to consider a Lagrangian of the form

$$L = \int d^3x \, \mathcal{L}\big(\phi(x), \partial_\mu \phi(x)\big), \tag{3.1}$$

where $\mathcal{L}$ is called the *Lagrangian density*. The corresponding action is then

$$S[\phi] = \int dt \, d^3x \, \mathcal{L}\big(\phi(x), \partial_\mu \phi(x)\big). \tag{3.2}$$

This represents the general form of the field theories we will study. A few important features of this construction are worth emphasizing [18]:

1. In the integral we encounter the four-dimensional measure

$$d^4x = dx^0 dx^1 dx^2 dx^3,$$

[18] In addition to the following properties, we will also demand our Lagrangian to be a function, at max, first order derivatives of fields. Also, we will demand our Lagrangian to be real and scalar. Technically, one can consider Lagrangians with second derivatives of the fields as long as you get 2nd order EoMs from them. For instance, one can consider a Lagrangian: $\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{\alpha}{2}\phi \, \partial_\mu \partial^\mu \phi$. Now performing an integration by parts we see that we get a total derivative term from above and hence ultimately we still get second order EoMs. This is exactly what happens in Einstein's GR where the action itself is a second-derivative of the metric – which is the *Ricci scalar* and still we get the corresponding field equations to be second-derivative of the metric as the variation of the Ricci scalar produces 2nd derivatives terms which can be written as total derivatives.

meaning that space and time coordinates appear on equal footing. This symmetry between space and time will later be shown to be crucial for relativistic invariance.

2. The principle of *locality* is automatically incorporated: fields interact only when evaluated at the same spacetime point. Consequently, the action never contains terms of the form $\phi(t,\vec{x})\phi(t,\vec{y})$ with $\vec{x}\neq\vec{y}$. This requirement is essential for having a well-behaved field theory as we had argued heuristically earlier.

3.1 Euler–Lagrange equations

Now let us derive the Euler-Lagrange equation for a given action of classical fields. Let $\phi_a(x)$, $a=1,\ldots,N$, be fields with Lagrangian density $\mathcal{L}(\phi_a,\partial_\mu\phi_a,x)$. The action is

$$S[\phi]=\int \mathrm{d}^4x\,\mathcal{L}(\phi_a,\partial_\mu\phi_a,x)\,. \tag{3.3}$$

Consider a variation $\phi_a\to\phi_a+\delta\phi_a$ with $\delta\phi_a$ vanishing on the boundary of the integration region. Then

$$\delta S=\int \mathrm{d}^4x\left[\frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a+\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta(\partial_\mu\phi_a)\right] \tag{3.4}$$

$$=\int \mathrm{d}^4x\left[\frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a+\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\mu(\delta\phi_a)\right] \tag{3.5}$$

$$=\int \mathrm{d}^4x\left[\frac{\partial\mathcal{L}}{\partial\phi_a}-\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]\delta\phi_a+\underbrace{\int \mathrm{d}^4x\,\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right)}_{\text{boundary term}=0}. \tag{3.6}$$

Stationarity $\delta S=0$ for arbitrary $\delta\phi_a$ yields the Euler–Lagrange equations:

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)-\frac{\partial\mathcal{L}}{\partial\phi_a}=0\,. \tag{3.7}$$

Above we made an assumption that $\delta(\partial_\mu\phi_a)=\partial_\mu(\delta\phi_a)$. This assumption is justified because $\delta\phi$ is an arbitrary but smooth function of spacetime that vanishes on the boundary, so differentiation and variation can be interchanged without issue [19].

Notation: when a field configuration satisfies the Euler-Lagrange equations, we call it *on-shell*. For an on-shell field configuration the action is stationary under small variations about this field configuration.

3.1.1 Action of a real scalar field

Let us now outline some requirements for the action S (equivalently, for the Lagrangian density $\mathcal{L}$):

1. The action should be invariant under Poincaré transformations. We have

$$S=\int d^4x\,\mathcal{L}(\phi,\partial_\mu\phi). \tag{3.8}$$

[19] Obviously at a physics level of rigour only :).

To check this, consider the measure under a Poincaré transformation. For translations $x' = x - a$ we clearly have $d^4x' = d^4x$. For Lorentz transformations $x' = \Lambda x$, we obtain $d^4x' = |\det \Lambda| \, d^4x = d^4x$. Hence the measure is invariant, and Poincaré invariance of the action requires only that the Lagrangian density $\mathcal{L}$ itself be invariant.

2. The action should yield non-trivial dynamics. In practice this requires $\mathcal{L}$ to depend explicitly on $\partial_\mu \phi$.

The simplest choice satisfying these two conditions is the *action of a real scalar field*:

$$\mathcal{L}(\phi, \partial_\mu \phi) = \tfrac{1}{2} \, \partial_\mu \phi \, \partial^\mu \phi - V(\phi). \tag{3.9}$$

Here $V(\phi)$ is a function of a single variable, known as the *potential* of the scalar field. Expanding the expression, we find

$$\mathcal{L} = \tfrac{1}{2} \, \eta^{\mu\nu} \, \partial_\mu \phi \, \partial_\nu \phi - V(\phi), \tag{3.10}$$

$$= \tfrac{1}{2} (\partial_t \phi)^2 - \tfrac{1}{2} \, \vec{\nabla} \phi \cdot \vec{\nabla} \phi - V(\phi). \tag{3.11}$$

This form is manifestly relativistically invariant, since all indices are contracted with the Minkowski metric. By analogy with point particle mechanics, the first term can be interpreted as a type of "kinetic energy," while the gradient term and the potential together play the role of "potential energy."

The simplest choice for the potential is

$$V(\phi) = \tfrac{1}{2} m^2 \phi^2. \tag{3.12}$$

What does the parameter m signify? With this form of potential it turns out that the quantum theory built from this action corresponds to a system of non-interacting particles of mass m. This is a nontrivial statement which we will justify later, but for now we simply take it as motivation.

Let us now explicitly derive the Euler–Lagrange equation for this system. The general field-theoretic Euler–Lagrange equation reads

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0. \tag{3.13}$$

First, compute the derivative with respect to ϕ:

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{dV}{d\phi} = -m^2 \phi. \tag{3.14}$$

Next, compute the derivative with respect to $\partial_\mu \phi$. From $\mathcal{L} = \tfrac{1}{2} \eta^{\rho\sigma} \partial_\rho \phi \, \partial_\sigma \phi - V(\phi)$, we obtain

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = \eta^{\mu\sigma} \partial_\sigma \phi = \partial^\mu \phi. \tag{3.15}$$

where we used the identity [20]:

$$\frac{\partial(\partial_\mu\phi)}{\partial(\partial_\nu\phi)} = \delta^\nu_\mu. \tag{3.16}$$

Thus,

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) = \partial_\mu\partial^\mu\phi. \tag{3.17}$$

Putting everything together, the Euler–Lagrange equation becomes

$$\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) = -m^2\phi - \partial_\mu\partial^\mu\phi = 0, \tag{3.18}$$

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0. \tag{3.19}$$

This is the *Klein–Gordon equation*, a relativistically invariant linear wave equation. Finally, let us expand the d'Alembertian explicitly,

$$\partial_\mu\partial^\mu = \eta^{\mu\nu}\partial_\mu\partial_\nu = \partial_t^2 - \vec{\nabla}\cdot\vec{\nabla} = \partial_t^2 - \nabla^2. \tag{3.20}$$

Thus, the Klein–Gordon equation takes the familiar form

$$\partial_t^2\phi - \nabla^2\phi + m^2\phi = 0, \tag{3.21}$$

which is the relativistic generalization of the wave equation. The speed of propagation is set to $c = 1$ in natural units.

[20] This identity can be derived using *functional derivative* methods, check out the book *QFT for gifted amateurs* for this (section 1.3).

§10: Consider the following Lagrangian:

$$\mathcal{L}_1 = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A^\mu)^2 \tag{3.22}$$

where $(\partial_\mu A^\mu)^2 = (\eta^{\mu\nu}\partial_\mu A_\nu)(\eta^{\rho\sigma}\partial_\rho A_\sigma)$.

Derive the EoMs for this Lagrangian. Hint: Use the identity: $\frac{\partial(\partial_\mu A_\nu)}{\partial(\partial_\rho A_\sigma)} = \delta^\rho_\mu \delta^\sigma_\nu$.

Show that they coincide with the EoMs derived from the following Lagrangian:

$$\mathcal{L}_2 = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{3.23}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. Hence, show that $\mathcal{L}_2 - \mathcal{L}_1$ is a total derivative term and hence they give the same action upto boundary terms which we can discard.

The above Lagrangians give the action for the electromagnetic field without sources. A_μ is the electromagnetic potential and $F_{\mu\nu}$ is its field strength called the *Faraday tensor*. Notice that there is no kinteic term $(\partial_t A_0)^2$ for A_0 in $\mathcal{L}_1$. What does this mean? You will find its use when you do *Quantum electrodynamics*!

Check that you get two out of the four source-free Maxwell equations from their EoMs. What happens to the other two Maxwell equations? What is a Lorentz covariant way to write them?

Hint: Use the definition of $F_{\mu\nu}$ above and consider taking *cyclic-derivatives* of it.

3.2 Noether's theorem

We now turn to Noether's theorem, one of the most profound results in mathematical physics. In essence, it states:

Every continuous symmetry of the action gives rise to a conserved current.

Let us clarify what is meant by a continuous symmetry through an example.

3.2.1 Complex scalar field

As a concrete case, let us consider a complex scalar field Φ (rather than a real scalar field). The action for this theory is

$$S[\Phi, \Phi^*] = \int d^4x \left((\partial_\mu \Phi^*)(\partial^\mu \Phi) - m^2 \Phi^* \Phi\right). \tag{3.24}$$

§11: Derive the equations of motion for the above action.

Here Φ has both real and imaginary parts, encoded together with its conjugate Φ^*. The action is invariant under the $U(1)$ global phase rotation [21].

$$\Phi \to e^{i\alpha}\Phi, \qquad \Phi^* \to e^{-i\alpha}\Phi^*. \tag{3.25}$$

Let us check this explicitly:

$$\begin{aligned} S[\Phi e^{i\alpha}, \Phi^* e^{-i\alpha}] &= \int d^4x \left((\partial_\mu \Phi^* e^{-i\alpha})(\partial^\mu \Phi e^{i\alpha}) - m^2 \Phi^* e^{-i\alpha} \Phi e^{i\alpha} \right) \\ &= \int d^4x \left((\partial_\mu \Phi^*)(\partial^\mu \Phi) - m^2 \Phi^* \Phi \right) \\ &= S[\Phi, \Phi^*]. \end{aligned} \tag{3.26}$$

Thus the action is unchanged. Since α is constant across spacetime, this is called a *continuous global symmetry*. It is continuous because it holds for all $\alpha \in \mathbb{R}$, or equivalently, for all $e^{i\alpha} \in U(1)$.

We can also expand for small α to get,

$$\Phi \to (1 + i\alpha)\Phi + \mathcal{O}(\alpha^2), \qquad \Phi^* \to (1 - i\alpha)\Phi^* + \mathcal{O}(\alpha^2). \tag{3.27}$$

Besides continuous symmetries, there can also be discrete ones [22]. In this case, the action is invariant under charge conjugation:

$$\Phi \to \Phi^*, \qquad \Phi^* \to \Phi. \tag{3.28}$$

This is often denoted by C. In the quantum theory this operation exchanges particles with antiparticles, as we will see later. However, it has no infinitesimal version and so does not directly fall under Noether's theorem. This example shows that, existence of conserved currents is a stronger statement than existence of conserved charges.

Proof of Noether's theorem

Now let us prove the theorem. Consider a field theory with scalar fields $\phi_a(x)$ and action $S[\phi_a]$. Suppose the theory is invariant under a continuous symmetry. Then there exists an infinitesimal transformation

$$\phi_a(x) \to \phi_a(x) + \epsilon\, \Delta\phi_a(x), \tag{3.29}$$

where ϵ is a small symmetry parameter. Invariance of the action implies

$$S[\phi_a + \epsilon \Delta\phi_a] - S[\phi_a] = 0 \quad \Rightarrow \quad \delta_\epsilon S[\phi_a] = 0. \tag{3.30}$$

[21] $U(1)$ is the group of unitary rank 1 matrices. In other words, it is a set of phases: $e^{i\alpha}$ with $\alpha \in \mathbb{R}$ which forms a group.

[22] For real scalar field this is just $\Phi \to -\Phi$ or what is called the $\mathbb{Z}_2$ symmetry – which can be thought of as the group consisting of the set: $\{1, -1\}$ equipped with the ordinary multiplication of integers as the binary operation. Here, note that even powers of Φ transform trivially, that is, with +1 eigenvalue under this symmetry while odd powers of Φ, transform with -1 eigenvalue. Due to this it is said that the former kind of fields live in the *even* representation of $\mathbb{Z}_2$ while the latter type of fields live in its *odd* representation.

To proceed, consider a generalization where the parameter depends on spacetime:

$$\phi_a(x) \;\to\; \phi_a(x) + \epsilon(x)\,\Delta\phi_a(x). \tag{3.31}$$

This trick will allow us to derive the conserved current. The action with a spacetime-dependent parameter $\epsilon(x)$ is no longer invariant. However, for constant ϵ it must vanish (since this reduces to the global case). The variation can therefore be written in the form

$$\delta_\epsilon S[\phi_a] = \int d^4x\, j^\mu\, \partial_\mu \epsilon(x), \tag{3.32}$$

where j^μ is some function of the fields, depending on the particular theory. Integrating by parts and discarding boundary terms gives

$$\delta_\epsilon S[\phi_a] = -\int d^4x\, (\partial_\mu j^\mu)\, \epsilon(x). \tag{3.33}$$

This identity holds for arbitrary field configurations ϕ_a. Now let us specialize to fields that satisfy the Euler–Lagrange equations of motion, i.e. fields that are *on-shell*. In that case the variation of the action vanishes for any $\epsilon(x)$, so we obtain

$$0 = \delta_\epsilon S[\phi_a] = -\int d^4x\, (\partial_\mu j^\mu)\, \epsilon(x). \tag{3.34}$$

Since $\epsilon(x)$ is arbitrary and non-vanishing, it follows that

$$\partial_\mu j^\mu = 0. \tag{3.35}$$

Thus, there exists a divergenceless vector j^μ, called a *conserved current*. This proves Noether's theorem, and also provides an algorithm to construct the current explicitly.

Conserved charge

What does it mean to have a divergenceless current? It implies the existence of a conserved quantity. Writing (3.35) in components:

$$\partial_t j^0(t,\vec{x}) + \partial_i j^i(t,\vec{x}) = 0. \tag{3.36}$$

Integrating over all space, we obtain

$$\partial_t \int d^3x\, j^0(t,\vec{x}) + \int d^3x\, \partial_i j^i(t,\vec{x}) = 0. \tag{3.37}$$

The second term is a total derivative, which vanishes if fields fall off sufficiently rapidly at spatial infinity. Thus we arrive at

$$\frac{d}{dt} Q(t) = 0, \qquad Q \equiv \int d^3x\, j^0(t,\vec{x}). \tag{3.38}$$

In other words, Q is a conserved charge, constant in time.

Complex scalar field again!

Let us compute the conserved current for the complex scalar field. From the infinitesimal $U(1)$ transformation (3.27), we have

$$\Delta\Phi(x) = i\alpha\,\Phi(x), \qquad \Delta\Phi^*(x) = -i\alpha\,\Phi^*(x). \tag{3.39}$$

If we allow $\alpha \to \alpha(x)$ to depend on spacetime, this becomes

$$\Delta\Phi(x) = i\alpha(x)\Phi(x), \qquad \Delta\Phi^*(x) = -i\alpha(x)\Phi^*(x). \tag{3.40}$$

Now compute the variation of the action

$$\delta_\alpha S[\Phi,\Phi^*] = \int d^4x \Big[(\partial_\mu\Delta\Phi^*)(\partial^\mu\Phi) + (\partial_\mu\Phi^*)(\partial^\mu\Delta\Phi) - m^2\Delta\Phi^*\Phi - m^2\Phi^*\Delta\Phi\Big] \tag{3.41}$$

$$= \int d^4x \Big[(\partial_\mu(-i\alpha\Phi^*))(\partial^\mu\Phi) + (\partial_\mu\Phi^*)(\partial^\mu(i\alpha\Phi)) - m^2(-i\alpha\Phi^*)\Phi - m^2\Phi^*(i\alpha\Phi)\Big] \tag{3.42}$$

$$= \int d^4x\, \partial_\mu\alpha \Big[i(\partial^\mu\Phi^*)\Phi - i\Phi^*\partial^\mu\Phi\Big]. \tag{3.43}$$

From this expression we can read off the conserved current:

$$j^\mu = i\Big[(\partial^\mu\Phi^*)\Phi - \Phi^*(\partial^\mu\Phi)\Big]. \tag{3.44}$$

We can now also construct the conserved charge explicitly:

$$Q = \int d^3x\, j^0 = i\int d^3x \Big[(\partial_t\Phi^*)\Phi - \Phi^*(\partial_t\Phi)\Big]. \tag{3.45}$$

§12: Check that this charge is conserved directly using the Euler–Lagrange equations of motion for the complex scalar field.

Alternative way to find the current

We now present another proof of Noether's theorem. Consider infinitesimal transformations. Suppose we have a continuous symmetry. Then a transformation of the fields

$$\delta\phi_a(x) = X_a(\phi) \tag{3.46}$$

is a symmetry if the Lagrangian changes only by a total derivative:

$$\delta\mathcal{L} = \partial_\mu F^\mu, \tag{3.47}$$

for some functions $F^\mu(\phi)$.

To derive Noether's theorem, let us first consider a general variation of the fields $\delta\phi_a$. The corresponding variation of the Lagrangian density is

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\mu(\delta\phi_a) \\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right).\end{aligned} \tag{3.48}$$

On-shell, the Euler–Lagrange equations hold:

$$\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) = 0,$$

so the first term vanishes. Thus,

$$\delta \mathcal{L} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta \phi_a \right). \tag{3.49}$$

For a symmetry transformation $\delta\phi_a = X_a(\phi)$, by assumption $\delta\mathcal{L} = \partial_\mu F^\mu$. Comparing with (3.49), we find the conserved current:

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} X_a(\phi) - F^\mu(\phi), \qquad \partial_\mu j^\mu = 0. \tag{3.50}$$

This provides a quick way to find symmetry currents when the Lagrangian itself is invariant. For the complex scalar field, one may check that this gives the same result as (3.44).

§13: Check this.

Translations and the Energy–Momentum Tensor

Let $\mathcal{L}(\phi_a, \partial_\mu \phi_a)$ be a Lagrangian density (no explicit x–dependence) and

$$S[\phi] = \int d^4x\, \mathcal{L}(\phi_a, \partial_\mu \phi_a), \qquad \Pi_a^{\ \mu} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}, \tag{3.51}$$

Consider a spacetime translation $x^\mu \to x'^\mu = x^\mu + \epsilon^\mu$ acts on the *fields* at the same spacetime point as

$$\phi'_a(x) = \phi_a(x-\epsilon) = \phi_a(x) - \epsilon^\nu \partial_\nu \phi_a(x) + \mathcal{O}(\epsilon^2), \qquad \Rightarrow \qquad \delta\phi_a(x) = -\epsilon^\nu(x)\, \partial_\nu \phi_a(x). \tag{3.52}$$

Promote ϵ^ν to a spacetime-dependent parameter $\epsilon^\nu(x)$. Now let us vary the action. Using $\delta(\partial_\mu \phi_a) = \partial_\mu \delta\phi_a = -(\partial_\mu \epsilon^\nu)\partial_\nu \phi_a - \epsilon^\nu \partial_\mu \partial_\nu \phi_a$, we get

$$\begin{aligned} \delta S &= \int d^4x \Big[\frac{\partial \mathcal{L}}{\partial \phi_a} \delta\phi_a + \Pi_a^{\ \mu}\, \delta(\partial_\mu \phi_a) \Big] \\ &= -\int d^4x \Big[\frac{\partial \mathcal{L}}{\partial \phi_a} \epsilon^\nu \partial_\nu \phi_a + \Pi_a^{\ \mu} (\partial_\mu \epsilon^\nu)\, \partial_\nu \phi_a + \Pi_a^{\ \mu} \epsilon^\nu \partial_\mu \partial_\nu \phi_a \Big]. \end{aligned} \tag{3.53}$$

Now using the first equation in (3.48), we can write:

$$\partial_\nu \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi_a} \partial_\nu \phi_a + \Pi_a^\mu \partial_\nu \partial_\mu \phi_a \tag{3.54}$$

and plugging this in (3.53) we get:

$$\delta S = -\int d^4x \, \left[\Pi_a^{\ \mu} (\partial_\mu \epsilon^\nu)\, \partial_\nu \phi_a + \epsilon^\nu \partial_\nu \mathcal{L} \right]. \tag{3.55}$$

Performing an integration by part above, we get (upto boundary terms):

$$\delta S = \int d^4x\, \partial_\mu \epsilon^\nu(x) \Big[\Pi_a^{\ \mu}\, \partial_\nu \phi_a - \delta^\mu_{\ \nu}\, \mathcal{L} \Big]. \tag{3.56}$$

Hence, identifying $\delta S = \int d^4x\, (\partial_\mu \epsilon^\nu)\, j^\mu{}_{(\nu)}$, the *canonical stress-energy tensor* is,

$$T^\mu{}_\nu = \Pi_a^{\ \mu}\, \partial_\nu \phi_a - \delta^\mu{}_\nu\, \mathcal{L}\ , \tag{3.57}$$

and translation invariance for constant ϵ^ν implies the conservation law

$$\partial_\mu T^\mu{}_\nu = 0\ . \tag{3.58}$$

Energy, momentum, and local densities: Define the conserved charges

$$P_\nu = \int d^3x\, T^0{}_\nu(t, \mathbf{x}), \qquad \frac{dP_\nu}{dt} = 0\,. \tag{3.59}$$

Then $P_0 = \int d^3x\, T^0{}_0$ is the total *energy* (Hamiltonian) and $P_i = \int d^3x\, T^0{}_i$ are the components of the total *momentum*. Locally: T^{00} is the energy density, T^{0i} the momentum density (and energy flux with raised indices), and T^{ij} the stress (momentum flux).

Real scalar field: For a single real scalar with

$$\mathcal{L} = \frac{1}{2}\, \partial_\mu \phi\, \partial^\mu \phi - V(\phi) = \frac{1}{2}\big[(\partial_t \phi)^2 - (\nabla \phi)^2\big] - V(\phi)\,, \tag{3.60}$$

we have $\Pi^{\ \mu} = \partial^\mu \phi$, so

$$T^\mu{}_\nu = \partial^\mu \phi\, \partial_\nu \phi - \delta^\mu{}_\nu\, \mathcal{L}\ , \qquad T^{\mu\nu} = \partial^\mu \phi\, \partial^{\ \nu} \phi - \eta^{\mu\nu}\, \mathcal{L}\ . \tag{3.61}$$

In components:

$$T^{00} = \frac{1}{2}(\partial_t \phi)^2 + \frac{1}{2}(\nabla \phi)^2 + V(\phi) \quad \text{(energy density)}, \tag{3.62}$$

$$T^0{}_i = \dot{\phi}\, \partial_i \phi \quad \text{(momentum density)}, \tag{3.63}$$

$$T^{ij} = \partial^i \phi\, \partial^j \phi - \eta^{ij} \mathcal{L} \quad \text{(stress)}. \tag{3.64}$$

§14: Compute the above quantities for $V(\phi) = \frac{1}{2} m^2 \phi^2$.

Note that, above $T^{\mu\nu}$ happens to be symmetric: $T^{\mu\nu} = T^{\nu\mu}$. This will not always be the case. However, in general one can always construct a symmetric energy–momentum tensor by adding a divergence-free "Belinfante improvement term":

$$\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\rho \Gamma^{\rho\mu\nu}, \tag{3.65}$$

where $\Gamma^{\rho\mu\nu}$ is antisymmetric in its first two indices, $\Gamma^{\rho\mu\nu} = -\Gamma^{\mu\rho\nu}$. This antisymmetry ensures $\partial_\mu \partial_\rho \Gamma^{\rho\mu\nu} = 0$, so the improved tensor $\Theta^{\mu\nu}$ is also conserved. To make $\Theta^{\mu\nu}$ symmetric we can choose $\Gamma^{\rho\mu\nu}$ such that the following holds:

$$T^{\mu\nu} - T^{\nu\mu} + \partial_\rho (\Gamma^{\rho\mu\nu} - \Gamma^{\rho\nu\mu}) = 0 \tag{3.66}$$

A comment on the extra index on the stress tensor

Note that, for the stress tensor we had $j^{\mu}_{(\nu)} \equiv T^{\mu}{}_{\nu}$ and the spacetime derivative acted on the μ index while the translation parameter was contracted with the ν index. From this we see that μ is a spacetime index and ν is a Lie algebra valued index in the sense that it labels the different conserved charges which generate different linearly independent translations. Hence, from this point of view though stress tensor has two indices, it can be viewed as a one index object like usual conserved currents while the other index labels the generators. In other words, since ϵ^{ν} is a constant 4-vector we can simply define the conserved current to be $j^{\mu} \equiv \epsilon^{\nu} T^{\mu}{}_{\nu}$ which will again be conserved [23].

Note: In all the above discussions, we kept the symmetry parameter: ϵ as a global quantity and only made it local: $\epsilon(x)$ as a trick to derive Noether's theorem. Instead if we had *physically* promoted $\epsilon \to \epsilon(x)$ then we would have gotten what is called a *gauge symmetry*. In that case also, the above Noether procedure will work and you will get a conserved charge only when $\epsilon(x) \neq 0$ at the boundary (at infinity) implying we get non-trivial conserved charge only for what are called *large gauge transformations*. In quantum electrodynamics, you will see that the electric charge is essentially a gauge charge but is still an observable since it is tied to the large gauge part of the $U(1)$ gauge group. This is the charge measure in say a lab. A quick way to see this is:

$$\delta_{\epsilon} S = \int d^4x\, \partial_{\mu} \epsilon_{\nu}\, T^{\mu\nu} = \int d^4x\, \partial_{\mu}(\epsilon_{\nu}\, T^{\mu\nu}) - \epsilon_{\nu} \partial_{\mu} T^{\mu\nu} \equiv \int d^4x\, \partial_{\mu} j^{\mu}(\epsilon) \tag{3.67}$$

where the last equation follows from the conservation of the stress tensor and we defined an ϵ-dependent current or a gauge current as: $j^{\mu}(\epsilon) = \epsilon_{\nu}\, T^{\mu\nu}$. This will yield a non-trivial charge only when $\epsilon_{\nu} \neq 0$ at the boundary (at infinity): $\epsilon_{\nu}\, T^{0\nu}$. So, the boundary terms we have been neglecting up until now play a crucial role to define a non-trivial gauge charge for large gauge transformations.

3.2.2 Lorentz Transformations and Angular Momentum

Here we will use the second way to derive conserved charges and currents. The infinitesimal form $\omega^{\mu\nu}$ of a Lorentz transformation for a scalar field yields,

$$\begin{aligned} \phi(x) \;\to\; \phi'(x) &= \phi(\Lambda^{-1} x) \\ &= \phi(x^{\mu} - \omega^{\mu}{}_{\nu} x^{\nu}) \\ &= \phi(x^{\mu}) - \omega^{\mu}{}_{\nu} x^{\nu} \partial_{\mu} \phi(x). \end{aligned} \tag{3.68}$$

[23] In generic spacetime, this is one way in which we can see why Killing vectors are important for symmetries. For instance, consider $\epsilon^{\nu}(x)$ as functions of spacetime now. Then $\partial_{\mu}[\epsilon_{\nu} T^{\mu\nu}] = \epsilon_{\nu} \partial_{\mu} T^{\mu\nu} + T^{\mu\nu}[\partial_{\mu} \epsilon_{\nu}] = 0 + T^{\mu\nu}[\partial_{(\mu} \epsilon_{\nu)}]$, where the first term is zero by conservation of the stress tensor and the second term comes from symmetric indices of the stress tensor. Now demanding that this vanishes means ϵ_{ν} must satisfy: $\partial_{(\mu} \epsilon_{\nu)} = 0$ – which is the Killing's equations and hence ϵ_{ν} are the Killing vectors. Note, for a rank two covariant tensor: $A_{(\mu\nu)}$ refers to complete symmetrisation of indices and $A_{[\mu\nu]}$ refers to complete anti-symmetrisation of indices.

where $\omega^{\mu}_{\ \nu}$ is the infinitesimal Lorentz transformation itself and no the antisymmetric parameter $\omega_{\mu\nu}$ considered previously.

From the above, we read off,

$$\delta\phi = -\omega^{\mu}_{\ \nu} x^{\nu} \partial_{\mu}\phi. \tag{3.69}$$

The Lagrangian density transforms as

$$\delta\mathcal{L} = -\omega^{\mu}_{\ \nu} x^{\nu} \partial_{\mu}\mathcal{L} = -\partial_{\mu}\big(\omega^{\mu}_{\ \nu} x^{\nu}\mathcal{L}\big), \tag{3.70}$$

where the last equality holds because $\omega^{\mu}_{\mu} = 0$ due to antisymmetry. Thus, the variation is a total derivative, and Noether's theorem applies with $F^{\mu} = -\omega^{\mu}_{\ \nu} x^{\nu}\mathcal{L}$.

This yields the conserved current

$$\begin{aligned} j^{\mu} &= \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\,\omega^{\rho}_{\ \nu} x^{\nu}\partial_{\rho}\phi + \omega^{\mu}_{\ \nu} x^{\nu}\mathcal{L} \\ &= -\omega^{\rho}_{\ \nu}\left[\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}\, x^{\nu}\partial_{\rho}\phi - \delta^{\mu}_{\ \rho} x^{\nu}\mathcal{L}\right] \\ &= -\omega^{\rho}_{\ \nu}\, T^{\mu}_{\rho} x^{\nu}, \end{aligned} \tag{3.71}$$

where T^{μ}_{ρ} is the energy–momentum tensor.

Stripping off the infinitesimal parameter $\omega^{\rho}_{\ \nu}$, we define the six conserved currents

$$(J^{\mu})^{\rho\sigma} = x^{\rho} T^{\mu\sigma} - x^{\sigma} T^{\mu\rho}, \tag{3.72}$$

which satisfy $\partial_{\mu}(J^{\mu})^{\rho\sigma} = 0$. These correspond to six conserved charges.

For $\rho, \sigma = 1, 2, 3$, the transformations are spatial rotations, and the conserved charges are the components of total angular momentum:

$$Q^{ij} = \int d^3x \left(x^i T^{0j} - x^j T^{0i}\right). \tag{3.73}$$

For boosts ($\rho = 0$, $\sigma = i$), the conserved charges are

$$Q^{0i} = \int d^3x \left(x^0 T^{0i} - x^i T^{00}\right). \tag{3.74}$$

Conservation of Q^{0i} implies

$$\begin{aligned} 0 = \frac{dQ^{0i}}{dt} &= \int d^3x\, T^{0i} + t\int d^3x\, \frac{\partial}{\partial t} T^{0i} - \frac{d}{dt}\int d^3x\, x^i T^{00} \\ &= P^i + t\frac{dP^i}{dt} - \frac{d}{dt}\int d^3x\, x^i T^{00}. \end{aligned} \tag{3.75}$$

Since P^i is already conserved ($dP^i/dt = 0$), we deduce

$$\begin{aligned} &\frac{d}{dt}\int d^3x\, x^i T^{00} = E\frac{dx^i}{dt} = P^i(\text{constant}), \\ &\frac{dx^i}{dt} = \frac{P^i}{E}(\text{constant}) \end{aligned} \tag{3.76}$$

This is the statement that the center of energy, x^i, of the field travels with a constant velocity: $\frac{P^i}{E}$. It's kind of like a field theoretic version of Newton's first law but, rather surprisingly, appearing here as a conservation law.

Now we see that Newton's first law assumes homogeneity of spacetime and isotropy of space. Hence, it is a law of assumptions and sets the stage for the second law.

§15: Off-shell conserved current

Consider a real scalar field in $(1+1)$ dimension. Define a current as follows:

$$j^\mu(x) \equiv \epsilon^{\mu\nu}\partial_\nu\phi(x) \tag{3.77}$$

where $\epsilon_{\mu\nu}$ is the Levi-Civita symbol in $(1+1)$ dimensions.

Show that the above constructed current is conserved. Note, that we don't need any equations of motion for the divergenceless of this current and hence it is off-shell conserved or conserved as an identity.

We can generalise the construction of Ex.(15) to vector and tensor fields and also to higher dimensions. For instance, in $(1+3)$ D with vector fields we will have,

$$j^{\mu\nu}(x) = \epsilon^{\mu\nu\rho\sigma}\partial_\rho A_\sigma(x) \tag{3.78}$$

which is again an off-shell conserved current. Can you identify where have you seen a similar current? Recall Ex.(10) where you got the remaining pair of Maxwell equations by considering cyclic-derivatives of the field strength $F_{\mu\nu}$. That equation is exactly identical to (3.78) and as we know that we don't need EoMs for this. This is an example of a *higher-form* symmetry, specifically 1-form symmetry in this case. If you write this equation in terms of $\vec{E}$ and $\vec{B}$ fields, you will notice that it corresponds to Gauss law for magnetostatics and Farday's law. Hence, this can be treated as a conservation equation for magnetic field lines since by Gauss law we know that there are no magnetic monopoles. However, this is a conservation law for extended objects – 1-dimensional magnetic field lines in this case – unlike the usual point particle conservation which was constructed by the Noether procedure and are called 0-form symmetries.

3.3 Hamiltonian methods

Now let us use Hamiltonian methods in classical field theory. As we know, the basic dynamical variable is the field $\phi(t,\vec{x})$. We define its canonical momentum $\pi(t,\vec{x})$ from the Lagrangian density as

$$\pi(t,\vec{x}) \equiv \frac{\partial\mathcal{L}}{\partial\dot{\phi}(t,\vec{x})}. \tag{3.79}$$

Real scalar field again!

For the real scalar field, recall the Lagrangian density

$$\begin{aligned}\mathcal{L} &= \tfrac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi - \tfrac{1}{2}m^2\phi^2 \\ &= \tfrac{1}{2}\dot{\phi}^2 - \tfrac{1}{2}\vec{\nabla}\phi\cdot\vec{\nabla}\phi - \tfrac{1}{2}m^2\phi^2.\end{aligned} \tag{3.80}$$

From this we obtain

$$\pi(t,\vec{x}) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}. \tag{3.81}$$

Hamiltonian density

The Hamiltonian density is constructed in the usual way:

$$\mathcal{H} = \pi(t,\vec{x})\,\dot{\phi}(t,\vec{x}) - \mathcal{L} \tag{3.82}$$

$$= \pi^2 - \left(\tfrac{1}{2}\pi^2 - \tfrac{1}{2}\vec{\nabla}\phi\cdot\vec{\nabla}\phi - \tfrac{1}{2}m^2\phi^2\right) \tag{3.83}$$

$$= \tfrac{1}{2}\pi^2 + \tfrac{1}{2}\vec{\nabla}\phi\cdot\vec{\nabla}\phi + \tfrac{1}{2}m^2\phi^2. \tag{3.84}$$

which is the same as T^{00} obtained before and now see that it is in the *Legendre transform* of the Lagrangian density. This justifies that T^{00} is indeed the energy density. Also, the Hamiltonian density is positive definite. The full Hamiltonian is the spatial integral:

$$H = \int d^3x\,\mathcal{H} = \int d^3x\left(\tfrac{1}{2}\pi^2 + \tfrac{1}{2}\vec{\nabla}\phi\cdot\vec{\nabla}\phi + \tfrac{1}{2}m^2\phi^2\right). \tag{3.85}$$

Comments

- As in classical mechanics, H can be interpreted as the *energy* of the system. Being an integral of a positive-definite density, it ensures positivity of the total energy.
- One can also check that H is conserved in time. In fact, it coincides with the Noether charge associated with time-translation invariance as we have seen in Ex.(14) and (3.59).
- A subtle point: although the underlying theory was Lorentz invariant, the Hamiltonian formalism requires picking out a preferred time coordinate t and defining $\dot{\phi} \equiv \partial_t\phi$. This procedure breaks manifest Lorentz covariance. The final physical results are Lorentz invariant, but intermediate steps are not.

§16: Poincaré Algebra from Conserved Charges

The conserved charges obtained for Poincaré transformations in the previous subsection are

$$P^\mu = \int d^3x\, T^{0\mu}, \qquad M^{\mu\nu} = \int d^3x \left(x^\mu T^{0\nu} - x^\nu T^{0\mu}\right) = \int d^3x \left(J^0\right)^{\mu\nu}, \qquad H = P^0. \tag{3.86}$$

Also define,

$$J^i = \tfrac{1}{2}\epsilon^{ijk} M_{jk}, \qquad K^i = M^{i0}. \tag{3.87}$$

Using above show that the Poincaré algebra as given in (2.23)-(2.25) becomes

$$[J^i, J^j] = \epsilon^{ijk} J^k, \qquad [J^i, K^j] = \epsilon^{ijk} K^k, \qquad [K^i, K^j] = -\epsilon^{ijk} J^k, \tag{3.88}$$

$$[J^i, P^j] = \epsilon^{ijk} P^k, \qquad [J^i, H] = 0, \qquad [K^i, H] = P^i, \tag{3.89}$$

$$[K^i, P^j] = \delta^{ij} H, \qquad [P^i, P^j] = 0, \qquad [P^i, H] = 0. \tag{3.90}$$

Explain what do each of the above commutators imply *physically*. Then identify which of the above form sub-algebras of the Poincaré algebra.

Hint: A sub-algebra of a Lie algebra is the set of all generators such that they map among themselves (or, to the zero vector) under the commutator structure. in other words, this sub-set is *closed* under the commutator structure.

4 Canonical quantization

We now move towards constructing quantum theories. We begin with the more familiar Hamiltonian approach, known as the *canonical formalism* (to be distinguished from the path integral formalism).

To recall how this works in ordinary quantum mechanics: we start with a classical system with coordinates q_a and their conjugate momenta p_a. From the Lagrangian $L(q_a, \dot{q}_a)$ one constructs the Hamiltonian $H(q_a, p_a)$. The basic structure is given by the Poisson brackets

$$\{q_a, q_b\} = 0, \qquad \{p_a, p_b\} = 0, \qquad \{q_a, p_b\} = \delta_{ab}. \tag{4.1}$$

Quantization is the algorithm which promotes these classical variables to Hermitian operators $\hat{q}_a$, $\hat{p}_a$ acting on a Hilbert space. The Poisson brackets are replaced by commutators. In the Schrödinger picture the operators are time independent, and the commutator algebra becomes

$$[\hat{q}_a, \hat{q}_b] = 0, \qquad [\hat{p}_a, \hat{p}_b] = 0, \qquad [\hat{q}_a, \hat{p}_b] = i\hbar\, \delta_{ab}. \tag{4.1}$$

Alternatively, in the Heisenberg picture the operators depend on time, $\hat{q}_a(t)$ and $\hat{p}_b(t)$.

The same algebra holds but only at equal times [24]:

$$[\hat{q}_a(t), \hat{q}_b(t)] = 0, \qquad [\hat{p}_a(t), \hat{p}_b(t)] = 0, \qquad [\hat{q}_a(t), \hat{p}_b(t)] = i\hbar\,\delta_{ab}. \tag{4.2}$$

4.1 Canonical commutators

We now extend this construction to fields. Consider a real scalar field with classical variables $\phi(t,\vec{x})$ and their conjugate momenta $\pi(t,\vec{x})$. Quantization promotes these to operators $\hat{\phi}(t,\vec{x})$ and $\hat{\pi}(t,\vec{x})$ satisfying the canonical commutation relations

$$[\hat{\phi}(t,\vec{x}), \hat{\phi}(t,\vec{y})] = 0, \tag{4.3}$$

$$[\hat{\pi}(t,\vec{x}), \hat{\pi}(t,\vec{y})] = 0, \tag{4.4}$$

$$[\hat{\phi}(t,\vec{x}), \hat{\pi}(t,\vec{y})] = i\delta^{(3)}(\vec{x}-\vec{y}). \tag{4.5}$$

where the last relation: the delta function is the continuum analogue of δ_{ab} above.

To make this more concrete, recall the free Klein–Gordon equation for a real scalar field, derived earlier:

$$\left(\partial_\mu \partial^\mu + m^2\right)\phi(t,\vec{x}) = \left(\partial_t^2 - \nabla^2 + m^2\right)\phi(t,\vec{x}) = 0, \tag{4.6}$$

Since this equation is translationally invariant, it is natural to expand the field in momentum space (check appendix A):

$$\phi(t,\vec{x}) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\vec{k}\cdot\vec{x}}\, \tilde{\phi}(t,\vec{k}). \tag{4.7}$$

This may be inverted to give

$$\tilde{\phi}(t,\vec{k}) = \int d^3x\, e^{-i\vec{k}\cdot\vec{x}}\, \phi(t,\vec{x}). \tag{4.8}$$

Thus, momentum space provides the natural basis for quantizing fields in a translationally invariant system.

Applying the Fourier expansion to the Klein–Gordon equation, recall that each Laplacian brings down a factor of $-\vec{k}^2$. Thus the equation of motion in momentum space becomes [25]

$$\left(\partial_t^2 + \vec{k}^2 + m^2\right)\tilde{\phi}(t,\vec{k}) = 0. \tag{4.9}$$

We see that this is identical to the equation of a harmonic oscillator, now with frequency

$$\omega_{\vec{k}}^2 = \vec{k}^2 + m^2.$$

[24] If we considered commutators at unequal times we would find different and generically more complicated commutation relations which depend explicitly on the time delay. Alternatively, in the Hamiltonian formalism we choose a particular time slice and define our Hilbert space of states in it. So, it is natural to consider states and operators defined in that time slice and relations among them for simplicity. Furthermore, we will see an exercise later that imposing these commutation relations at any fixed time means they keep holding for any arbitrary time.

[25] Strictly speaking, one obtains this after multiplying the original real–space equation by $e^{i\vec{k}\cdot\vec{x}}$ and integrating over space, but this step is equivalent since the Fourier transform is invertible.

Hence we may rewrite the equation as

$$\left(\partial_t^2 + \omega_{\vec{k}}^2\right)\tilde{\phi}(t,\vec{k}) = 0. \tag{4.10}$$

Thus, at least classically, each *Fourier mode of the free scalar field* behaves like a *harmonic oscillator*. Also, different modes don't talk with each other and hence they are decoupled! Each mode has their own SHO-like differential equation as in (4.9).

Now since we already know how to quantize the harmonic oscillator, this observation allows us to build the quantum theory of the scalar field by treating it as a collection of oscillators, one for each momentum $\vec{k}$.

It is important to note that this simplification hinges on the quadratic nature of the potential. Had the potential not been quadratic in ϕ, this mapping to oscillators would fail, and we would face the much harder problem of *interacting quantum field theory*. However, any potential $V(\phi)$ when expanded about its minima is to leading orders a SHO potential and hence the quadratic/SHO approximation is very crucial and important in perturbative QFT [26].

We now repeat a construction for the free *complex* scalar field similar to the quantum SHO case, which has the classical action

$$S[\phi,\phi^*] = \int d^4x \left[(\partial_\mu\phi^*)(\partial^\mu\phi) - m^2\phi^*\phi\right], \tag{4.11}$$

4.2 Classical theory

Let us first solve the classical equations of motion. In position space we find

$$(\partial_t^2 - \nabla^2 + m^2)\,\phi(t,\vec{x}) = 0, \tag{4.12}$$

$$(\partial_t^2 + \omega_{\vec{k}}^2)\,\tilde{\phi}(t,\vec{k}) = 0, \qquad \omega_{\vec{k}} = \sqrt{\vec{k}^{\,2} + m^2}. \tag{4.13}$$

The momentum–space equation is identical to that of a harmonic oscillator. Thus the general solution in momentum space is [27]

$$\tilde{\phi}(\vec{k},t) = \frac{1}{\sqrt{2\omega_{\vec{k}}}}\left(a_{\vec{k}}\, e^{-i\omega_{\vec{k}}t} + b^*_{-\vec{k}}\, e^{+i\omega_{\vec{k}}t}\right). \tag{4.14}$$

Here $a_{\vec{k}}$ and $b_{\vec{k}}$ are momentum–dependent Fourier coefficients. Transforming back to position space we have

$$\phi(\vec{x},t) = \int \frac{d^3k}{(2\pi)^3}\,\frac{1}{\sqrt{2\omega_{\vec{k}}}}\left(a_{\vec{k}}\, e^{-i\omega_{\vec{k}}t} + b^*_{-\vec{k}}\, e^{+i\omega_{\vec{k}}t}\right) e^{i\vec{k}\cdot\vec{x}}. \tag{4.15}$$

It is convenient to relabel in the second term by sending $\vec{k} \to -\vec{k}$. This gives a more symmetric and Lorentz-invariant expression:

$$\phi(\vec{x},t) = \int \frac{d^3k}{(2\pi)^3}\,\frac{1}{\sqrt{2\omega_{\vec{k}}}}\left(a_{\vec{k}}\, e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}} + b^*_{\vec{k}}\, e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}}\right). \tag{4.16}$$

[26] To quote Sidney Coleman: "Career of a young theoretical physicist is treating the harmonic oscillator in ever-increasing levels of abstraction".

[27] Recall, in the SHO case: $q(t) = \frac{1}{\sqrt{2\omega}}(ae^{-i\omega t} + be^{i\omega t})$ and by demanding reality of $q(t)$ we had the relation: $b = a^*$ – which will be the case for the real scalar field.

This is analogous to the real scalar case, except that here $a_{\vec{k}}$ and $b_{\vec{k}}$ are independent, since ϕ is complex.

Hamiltonian and canonical momenta

To quantize, we first compute the Hamiltonian. Expanding the Lagrangian density:

$$\mathcal{L} = \dot{\phi}^* \dot{\phi} - (\nabla \phi^*) \cdot (\nabla \phi) - m^2 \phi^* \phi. \tag{4.17}$$

The canonical momenta are obtained by varying with respect to $\dot{\phi}$ and $\dot{\phi}^*$:

$$\pi_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}^*, \tag{4.18}$$

$$\pi_{\phi^*} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^*} = \dot{\phi}. \tag{4.19}$$

Since ϕ and ϕ^* are independent variables, they each carry their own conjugate momentum. For notational simplicity, let us denote the canonical momenta as

$$\pi = \pi_\phi, \qquad\qquad \pi^* = \pi_{\phi^*}. \tag{4.20}$$

The Hamiltonian then takes the form

$$\mathcal{H} \;=\; \pi \dot{\phi} + \pi^* \dot{\phi}^* - \mathcal{L} = \dot{\phi}^* \dot{\phi} + (\nabla \phi^*) \cdot (\nabla \phi) + m^2 \phi^* \phi. \tag{4.21}$$

4.3 Operator-valued Fourier expansions

Now we proceed to quantization: the classical fields $\phi(t, \vec{x})$ and $\pi(t, \vec{x})$ are promoted to operators on a Hilbert space,

$$\phi(t, \vec{x}) \;\to\; \hat{\phi}(t, \vec{x}), \tag{4.22}$$

$$\phi^*(t, \vec{x}) \;\to\; \hat{\phi}^\dagger(t, \vec{x}). \tag{4.23}$$

In the Heisenberg picture, the Fourier expansion of the operator $\hat{\phi}$ is

$$\hat{\phi}(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}_{\vec{k}}\, e^{-i\omega_{\vec{k}} t + i\vec{k}\cdot\vec{x}} + \hat{b}^\dagger_{\vec{k}}\, e^{+i\omega_{\vec{k}} t - i\vec{k}\cdot\vec{x}} \right), \tag{4.24}$$

where the coefficients $\hat{a}_{\vec{k}}, \hat{b}_{\vec{k}}$ are now operators. Its Hermitian conjugate is

$$\hat{\phi}^\dagger(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}^\dagger_{\vec{k}}\, e^{+i\omega_{\vec{k}} t - i\vec{k}\cdot\vec{x}} + \hat{b}_{\vec{k}}\, e^{-i\omega_{\vec{k}} t + i\vec{k}\cdot\vec{x}} \right). \tag{4.25}$$

Next, consider the canonical momenta. Starting with the classical relation $\pi = \dot{\phi}^*$ and promoting it to an operator, we obtain

$$\hat{\pi}(\vec{x}, t) = \dot{\hat{\phi}}^\dagger(\vec{x}, t) = i \int \frac{d^3k}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{k}}}{2}} \left(\hat{a}^\dagger_{\vec{k}}\, e^{+i\omega_{\vec{k}} t - i\vec{k}\cdot\vec{x}} - \hat{b}_{\vec{k}}\, e^{-i\omega_{\vec{k}} t + i\vec{k}\cdot\vec{x}} \right). \tag{4.26}$$

Similarly, its conjugate momentum is

$$\hat{\pi}^\dagger(\vec{x},t) = \dot{\hat{\phi}}(\vec{x},t) = i\int \frac{d^3k}{(2\pi)^3}\sqrt{\frac{\omega_{\vec{k}}}{2}}\left(-\hat{a}_{\vec{k}}\, e^{-i\omega_{\vec{k}}t+i\vec{k}\cdot\vec{x}} + \hat{b}^\dagger_{\vec{k}}\, e^{+i\omega_{\vec{k}}t-i\vec{k}\cdot\vec{x}}\right). \tag{4.27}$$

The significance of these expansions is that the dynamics of the quantum field can be described entirely in terms of the operators $\hat{a}_{\vec{k}}$ and $\hat{b}_{\vec{k}}$. These operators will diagonalize the Hamiltonian and have a simple physical meaning: they create and annihilate particles.

Finally, recall the equal-time commutation relations of the scalar field and its conjugate momentum:

$$[\hat{\phi}(\vec{x},t), \hat{\pi}(\vec{y},t)] = [\hat{\phi}^\dagger(\vec{x},t), \hat{\pi}^\dagger(\vec{y},t)] = i\,\delta^{(3)}(\vec{x}-\vec{y}). \tag{4.28}$$

All other equal-time commutators vanish, e.g.

$$[\hat{\phi},\hat{\phi}] = [\hat{\phi},\hat{\pi}^\dagger] = \cdots = 0. \tag{4.29}$$

4.4 Algebra of creation and annihilation operators

Let us now ask: what do the canonical commutators imply for the operators $\hat{a}_{\vec{k}}$ and $\hat{b}_{\vec{k}}$? It follows that

$$[\hat{a}_{\vec{k}}, \hat{a}^\dagger_{\vec{k}'}] = (2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}'), \tag{4.30}$$

$$[\hat{b}_{\vec{k}}, \hat{b}^\dagger_{\vec{k}'}] = (2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}'), \tag{4.31}$$

with all other commutators vanishing. This means we can think of each Fourier mode as consisting of two independent harmonic oscillators: one set from the a operators, and another from the b operators. The Dirac delta indicates that distinct Fourier modes do not interact with each other at the algebraic level, or in other words, the oscillators are decoupled modes.

Let us now show this explicitly. This type of computation is standard in field theory, where Fourier integrals play a central role. We begin by recalling the Fourier identity

$$\int d^3x\, e^{i\vec{p}\cdot\vec{x}} = (2\pi)^3\delta^{(3)}(\vec{p}). \tag{4.32}$$

Now multiply the expansion of $\hat{\phi}$ in (4.24) by $e^{i\vec{k}\cdot\vec{x}}$ and integrate over all $\vec{x}$:

$$\int d^3x\, e^{i\vec{k}\cdot\vec{x}}\,\hat{\phi}(\vec{x},t) = \int d^3x \int \frac{d^3k'}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{k}'}}}\left(\hat{a}_{\vec{k}'}e^{-i\omega_{\vec{k}'}t+i(\vec{k}+\vec{k}')\cdot\vec{x}} + \hat{b}^\dagger_{\vec{k}'}e^{+i\omega_{\vec{k}'}t+i(\vec{k}-\vec{k}')\cdot\vec{x}}\right) \tag{4.33}$$

$$= \int \frac{d^3k'}{\sqrt{2\omega_{\vec{k}'}}}\left(\hat{a}_{\vec{k}'}e^{-i\omega_{\vec{k}'}t}(2\pi)^3\delta^{(3)}(\vec{k}+\vec{k}') + \hat{b}^\dagger_{\vec{k}'}e^{+i\omega_{\vec{k}'}t}(2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}')\right). \tag{4.34}$$

This yields

$$\int d^3x\, e^{i\vec{k}\cdot\vec{x}}\,\hat{\phi}(\vec{x},t) = \frac{1}{\sqrt{2\omega_{\vec{k}}}}\left(\hat{a}_{-\vec{k}}e^{-i\omega_{\vec{k}}t} + \hat{b}^\dagger_{\vec{k}}e^{+i\omega_{\vec{k}}t}\right), \tag{4.35}$$

where we used: $\omega_{-\vec{k}} = \omega_{\vec{k}}$ (see (4.13)).

Performing a similar manipulation with $\hat{\pi}^\dagger(\vec{x},t)$ using (4.27), we find

$$\int d^3x\, e^{i\vec{k}\cdot\vec{x}}\,\hat{\pi}^\dagger(\vec{x},t) = i\sqrt{\frac{\omega_{\vec{k}}}{2}}\left(-\hat{a}_{-\vec{k}}e^{-i\omega_{\vec{k}}t} + \hat{b}^\dagger_{\vec{k}}e^{+i\omega_{\vec{k}}t}\right). \tag{4.36}$$

Now, combining (4.35) and (4.36), we can isolate $\hat{a}_{\vec{k}}$: multiply the first equation by $\sqrt{\omega_{\vec{k}}}$ and the second by $1/\sqrt{\omega_{\vec{k}}}$ and then add them. This gives (with $\vec{k}\to-\vec{k}$)

$$\hat{a}_{\vec{k}} = \int d^3x\, e^{-i\vec{k}\cdot\vec{x}+i\omega_{\vec{k}}t}\left(\hat{\phi}(\vec{x},t)\sqrt{\frac{\omega_{\vec{k}}}{2}} + i\,\hat{\pi}^\dagger(\vec{x},t)\frac{1}{\sqrt{2\omega_{\vec{k}}}}\right). \tag{4.37}$$

Similarly, taking the opposite linear combination allows us to extract $\hat{b}^\dagger_{\vec{k}}$:

$$\hat{b}^\dagger_{\vec{k}} = \int d^3x\, e^{+i\vec{k}\cdot\vec{x}-i\omega_{\vec{k}}t}\left(\hat{\phi}(\vec{x},t)\sqrt{\frac{\omega_{\vec{k}}}{2}} - i\,\hat{\pi}^\dagger(\vec{x},t)\frac{1}{\sqrt{2\omega_{\vec{k}}}}\right). \tag{4.38}$$

Analogous expressions hold for $\hat{a}^\dagger_{\vec{k}}$ and $\hat{b}_{\vec{k}}$, obtained by taking Hermitian conjugates.

Taking Hermitian conjugates, we obtain

$$\hat{a}^\dagger_{\vec{k}} = \int d^3x\, e^{+i\vec{k}\cdot\vec{x}-i\omega_{\vec{k}}t}\left(\hat{\phi}^\dagger(\vec{x},t)\sqrt{\tfrac{\omega_{\vec{k}}}{2}} - i\hat{\pi}(\vec{x},t)\frac{1}{\sqrt{2\omega_{\vec{k}}}}\right), \tag{4.39}$$

$$\hat{b}_{\vec{k}} = \int d^3x\, e^{-i\vec{k}\cdot\vec{x}+i\omega_{\vec{k}}t}\left(\hat{\phi}^\dagger(\vec{x},t)\sqrt{\tfrac{\omega_{\vec{k}}}{2}} + i\hat{\pi}(\vec{x},t)\frac{1}{\sqrt{2\omega_{\vec{k}}}}\right). \tag{4.40}$$

Commutator of the creation and annihilation operators

Let us now compute the equal-time commutator of $\hat{a}_{\vec{k}}$ with $\hat{a}^\dagger_{\vec{k}'}$. Starting from the definitions, we find

$$\begin{aligned}[\hat{a}_{\vec{k}},\hat{a}^\dagger_{\vec{k}'}] = \tfrac{1}{2}\int d^3x\, d^3x'\, e^{it(\omega_{\vec{k}}-\omega_{\vec{k}'})-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{x}'}\\ \times\Bigg(\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}}[\hat{\phi}(\vec{x},t),\hat{\phi}^\dagger(\vec{x}',t)] - i\sqrt{\tfrac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}}[\hat{\phi}(\vec{x},t),\hat{\pi}(\vec{x}',t)]\\ + i\sqrt{\tfrac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}}[\hat{\pi}^\dagger(\vec{x},t),\hat{\phi}^\dagger(\vec{x}',t)] + \tfrac{1}{\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}}}[\hat{\pi}^\dagger(\vec{x},t),\hat{\pi}(\vec{x}',t)]\Bigg),\end{aligned} \tag{4.41}$$

which becomes,

$$[\hat{a}_{\vec{k}},\hat{a}^\dagger_{\vec{k}'}] = \tfrac{1}{2}\int d^3x\, d^3x'\, e^{i(\omega_{\vec{k}}-\omega_{\vec{k}'})t}\, e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{x}'}\times\Big(\sqrt{\tfrac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}}[\hat{\phi}(\vec{x},t),\hat{\pi}(\vec{x}',t)] + \sqrt{\tfrac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}}[\hat{\pi}^\dagger(\vec{x},t),\hat{\phi}^\dagger(\vec{x}',t)]\Big), \tag{4.42}$$

since the commutators of $\hat{\phi}$ with $\hat{\phi}^\dagger$ and of $\hat{\pi}$ with $\hat{\pi}^\dagger$ vanish identically.

Now the two mixed terms yield delta functions, and we obtain

$$[\hat{a}_{\vec{k}},\hat{a}^\dagger_{\vec{k}'}] = \tfrac{1}{2}\int d^3x\, d^3x'\, e^{i(\omega_{\vec{k}}-\omega_{\vec{k}'})t}\, e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{x}'}\left(\sqrt{\tfrac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}}\,\delta^{(3)}(\vec{x}-\vec{x}') + \sqrt{\tfrac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}}\,\delta^{(3)}(\vec{x}-\vec{x}')\right). \tag{4.43}$$

Using the delta function to collapse the $\vec{x}'$ integral (i.e. setting $\vec{x}' = \vec{x}$), we find

$$[\hat{a}_{\vec{k}}, \hat{a}^\dagger_{\vec{k}'}] = \tfrac{1}{2} \int d^3x \, e^{i(\omega_{\vec{k}} - \omega_{\vec{k}'})t} \, e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}} \left(\sqrt{\frac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}} + \sqrt{\frac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}} \right). \tag{4.44}$$

The spatial integral gives

$$[\hat{a}_{\vec{k}}, \hat{a}^\dagger_{\vec{k}'}] = \tfrac{1}{2}(2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}') \, e^{i(\omega_{\vec{k}} - \omega_{\vec{k}'})t} \left(\sqrt{\frac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}} + \sqrt{\frac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}} \right). \tag{4.45}$$

Since the delta function enforces $\vec{k} = \vec{k}'$, the two square-root terms each become 1, giving a total factor of 2 (see appendix C.1 for more details). Thus the final result is

$$[\hat{a}_{\vec{k}}, \hat{a}^\dagger_{\vec{k}'}] = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}'). \tag{4.46}$$

In exactly the same way, one checks that the only other nonvanishing commutator is

$$[\hat{b}_{\vec{k}}, \hat{b}^\dagger_{\vec{k}'}] = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}'), \tag{4.47}$$

with all other commutators equal to zero.

§17: Derive (4.47).

4.5 Hamiltonian and the energy of the vacuum

We now want to express the Hamiltonian of the quantum theory in terms of the creation and annihilation operators. From (4.21), the classical Hamiltonian is

$$H = \int d^3x \left(\pi\pi^* + \nabla\phi \cdot \nabla\phi^* + m^2 \phi^* \phi \right). \tag{4.48}$$

Promoting each of these classical quantities to operators gives the quantum Hamiltonian

$$\hat{H} = \int d^3x \left(\hat{\pi}\hat{\pi}^\dagger + \nabla\hat{\phi} \cdot \nabla\hat{\phi}^\dagger + m^2 \hat{\phi}^\dagger \hat{\phi} \right). \tag{4.49}$$

Expressing this in terms of the creation and annihilation operators is a straightforward but somewhat tedious algebraic exercise. After simplification one finds

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3} \, \omega_{\vec{k}} \left(\hat{a}^\dagger_{\vec{k}} \hat{a}_{\vec{k}} + \hat{b}_{\vec{k}} \hat{b}^\dagger_{\vec{k}} \right). \tag{4.50}$$

§18: Derive (4.50).

Rewriting this using the commutators of b and $b^\dagger$, we obtain

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3} \, \omega_{\vec{k}} \left(\hat{a}^\dagger_{\vec{k}} \hat{a}_{\vec{k}} + \hat{b}^\dagger_{\vec{k}} \hat{b}_{\vec{k}} + (2\pi)^3 \delta^{(3)}(0) \right). \tag{4.51}$$

Here we encounter the infamous divergence $\delta^{(3)}(0)$, which formally represents the infinite volume of space. This is the first instance of the many infinities that arise in quantum field theory.

To proceed, let us define the *vacuum state* $|0\rangle$, constructed in direct analogy with the harmonic oscillator vacuum. It is annihilated by all the operators $\hat{a}_{\vec{k}}$ and $\hat{b}_{\vec{k}}$:

$$\hat{a}_{\vec{k}}|0\rangle = 0, \qquad \hat{b}_{\vec{k}}|0\rangle = 0, \qquad \forall\, \vec{k}. \tag{4.52}$$

Thus the vacuum is the state with no particles present.

Vacuum energy

Let us now ask: what is the energy of the vacuum? Acting with the Hamiltonian on $|0\rangle$,

$$\hat{H}|0\rangle = \int \frac{d^3k}{(2\pi)^3}\, \omega_{\vec{k}} \Big(\hat{a}^\dagger_{\vec{k}} \hat{a}_{\vec{k}} + \hat{b}^\dagger_{\vec{k}} \hat{b}_{\vec{k}} + (2\pi)^3 \delta^{(3)}(0) \Big) |0\rangle \tag{4.53}$$

$$= \int \frac{d^3k}{(2\pi)^3}\, \omega_{\vec{k}} (2\pi)^3 \delta^{(3)}(0)\, |0\rangle. \tag{4.54}$$

Therefore the vacuum is an eigenstate of $\hat{H}$ with eigenvalue

$$E_0 = \int \frac{d^3k}{(2\pi)^3}\, \omega_{\vec{k}}\, (2\pi)^3 \delta^{(3)}(0). \tag{4.55}$$

This eigenvalue is divergent. To interpret it, recall that $\delta^{(3)}(0)$ corresponds to the spatial volume. In fact, one can regulate it as

$$(2\pi)^3 \delta^{(3)}(p) = \lim_{V\to\infty} \int_V d^3x\, e^{i\vec{p}\cdot\vec{x}}. \tag{4.56}$$

Thus, in a finite box of volume V, we should understand

$$(2\pi)^3 \delta^{(3)}(0) = \lim_{V\to\infty} \int_V d^3x = \lim_{V\to\infty} V. \tag{4.57}$$

Hence the vacuum energy is infinite, scaling with the volume of space, which is the manifestation of the zero-point energy of the quantum field.

4.6 Normal ordering

Since the vacuum energy diverges, one can try to regulate it by dividing out the spatial volume. This leads to the definition of a vacuum energy density

$$\epsilon_0 = \frac{E_0}{V} = \int \frac{d^3k}{(2\pi)^3}\, \omega_{\vec{k}} \left(\tfrac{1}{2} + \tfrac{1}{2} \right). \tag{4.58}$$

This expression is just the zero-point energy of the harmonic oscillator, but with contributions from each momentum mode and from both a- and b-type oscillators. The result is extremely divergent, and there is no natural way to eliminate it.

One common prescription is to adopt *normal ordering*. Given a product of field operators, normal ordering (denoted by $:\mathcal{O}:$) is defined by moving all annihilation operators (a, b) to

the right of all creation operators ($a^\dagger$, $b^\dagger$). Applying this to the Hamiltonian, we obtain the normal-ordered form [28]

$$:\hat{H}:=\int \frac{d^3k}{(2\pi)^3}\,\omega_{\vec{k}}\Big(\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}}+\hat{b}^\dagger_{\vec{k}}\hat{b}_{\vec{k}}\Big). \tag{4.59}$$

In this prescription, the vacuum is annihilated by the Hamiltonian,

$$:\hat{H}:\,|0\rangle=0, \tag{4.60}$$

so the infinite vacuum energy is removed.

This might feel somewhat ad hoc. In fact, the ambiguity arises because passing from the classical Hamiltonian to the quantum one involves operator ordering. For example, classically one might have terms like

$$\tfrac{1}{2}(\omega q-ip)(\omega q+ip)\quad\text{versus}\quad\tfrac{1}{2}(\omega q+ip)(\omega q-ip).$$

Quantizing, the first gives $\omega\hat{a}^\dagger\hat{a}$ while the second gives $\omega(\hat{a}^\dagger\hat{a}+\frac{1}{2})$. The difference corresponds precisely to the zero-point energy. Physical effects such as the *Casimir effect* show that vacuum energy can have observable consequences [29].

4.7 Single-particle states

Let us now move beyond the vacuum and consider single-particle excitations. We first compute the commutator of the Hamiltonian with a creation operator:

$$[\hat{H},\hat{a}^\dagger_{\vec{k}}]=\left[\int\frac{d^3k'}{(2\pi)^3}\,\omega_{\vec{k}'}\,\hat{a}^\dagger_{\vec{k}'}\hat{a}_{\vec{k}'},\;\hat{a}^\dagger_{\vec{k}}\right]=\int\frac{d^3k'}{(2\pi)^3}\,\omega_{\vec{k}'}\,\hat{a}^\dagger_{\vec{k}'}[\hat{a}_{\vec{k}'},\hat{a}^\dagger_{\vec{k}}] \tag{4.61}$$

$$=\int\frac{d^3k'}{(2\pi)^3}\,\omega_{\vec{k}'}\,\hat{a}^\dagger_{\vec{k}'}(2\pi)^3\delta^{(3)}(\vec{k}'-\vec{k})=\omega_{\vec{k}}\,\hat{a}^\dagger_{\vec{k}}. \tag{4.62}$$

Thus

$$[\hat{H},\hat{a}^\dagger_{\vec{k}}]=\omega_{\vec{k}}\,\hat{a}^\dagger_{\vec{k}},\qquad[\hat{H},\hat{a}_{\vec{k}}]=-\omega_{\vec{k}}\,\hat{a}_{\vec{k}}. \tag{4.63}$$

Exactly the same relations hold in the b-sector:

$$[\hat{H},\hat{b}^\dagger_{\vec{k}}]=\omega_{\vec{k}}\,\hat{b}^\dagger_{\vec{k}},\qquad[\hat{H},\hat{b}_{\vec{k}}]=-\omega_{\vec{k}}\,\hat{b}_{\vec{k}}. \tag{4.64}$$

Let us now consider an excited state $\hat{a}^\dagger_{\vec{k}}|0\rangle$. Its energy is computed as

$$:\hat{H}:\,\hat{a}^\dagger_{\vec{k}}|0\rangle=\Big(\omega_{\vec{k}}\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}}+\omega_{\vec{k}}\hat{a}_{\vec{k}}\hat{a}^\dagger_{\vec{k}}\Big)|0\rangle=\omega_{\vec{k}}\,\hat{a}^\dagger_{\vec{k}}|0\rangle. \tag{4.65}$$

Thus the energy of this state is

$$\omega_{\vec{k}}=\sqrt{\vec{k}^{\,2}+m^2}, \tag{4.66}$$

which is precisely the relativistic energy of a particle with momentum $\vec{k}$. *We have created a particle*! Similarly, the state $\hat{b}^\dagger_{\vec{k}}|0\rangle$ also has energy $\omega_{\vec{k}}$. This further justifies that P^0 as created out of the stress tensor is indeed the Hamiltonian.

[28] Note the normal ordered Hamiltonian differs from this by an infinite constant, but that does not affect this commutation relation.

[29] You can read more about them from David Tong's lecture notes (sec. 2.3.2.) or example 11.5 of chapter 11 in the book: *QFT for gifted amateurs*.

Noether charge for $U(1)$ symmetry

Recall that the complex scalar field has a global $U(1)$ symmetry,

$$\phi \;\to\; e^{i\alpha}\phi, \qquad \phi^* \;\to\; e^{-i\alpha}\phi^*. \tag{4.67}$$

The corresponding Noether charge (choosing a convenient sign convention) is

$$Q = -i\int d^3x \Big[(\partial_t\phi^*)\phi - \phi^*(\partial_t\phi)\Big]. \tag{4.68}$$

Promoting this to an operator gives

$$\hat{Q} = -i\int d^3x \Big[\hat{\pi}\,\hat{\phi} - \hat{\phi}^\dagger\,\hat{\pi}^\dagger\Big]. \tag{4.69}$$

After substituting the Fourier expansions and normal ordering, one finds

$$:\hat{Q}: = \int \frac{d^3k}{(2\pi)^3}\Big(\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}} - \hat{b}^\dagger_{\vec{k}}\hat{b}_{\vec{k}}\Big). \tag{4.70}$$

This operator is the $U(1)$ charge. Notice the key difference compared to the Hamiltonian: the b-sector contributes with the opposite sign. Thus, $a^\dagger$ creates particles with charge $+1$, while $b^\dagger$ creates antiparticles with charge -1.

§19: Derive (4.70).

By Noether's theorem, this charge is conserved: in the quantum theory,

$$[\hat{Q}, \hat{H}] = 0, \tag{4.71}$$

so we can simultaneously diagonalize $\hat{H}$ and $\hat{Q}$.

§20: Show that (4.71) holds by using (4.59) and (4.70).

Acting with $\hat{Q}$ on the states, we find

$$:\hat{Q}:\,|0\rangle = 0, \tag{4.72}$$

$$:\hat{Q}:\,\hat{a}^\dagger_{\vec{k}}|0\rangle = (+1)\,\hat{a}^\dagger_{\vec{k}}|0\rangle, \tag{4.73}$$

$$:\hat{Q}:\,\hat{b}^\dagger_{\vec{k}}|0\rangle = (-1)\,\hat{b}^\dagger_{\vec{k}}|0\rangle. \tag{4.74}$$

Thus the vacuum is neutral, $a^\dagger$ creates positively charged particles, and $b^\dagger$ creates negatively charged antiparticles.

Finally, recall that in a previous discussion we introduced the momentum operator via Noether's theorem for spatial translations. In the quantum theory this gives

$$\hat{P}^i\,\hat{a}^\dagger_{\vec{k}}|0\rangle = k^i\,\hat{a}^\dagger_{\vec{k}}|0\rangle, \tag{4.75}$$

showing explicitly that $\hat{a}^\dagger_{\vec{k}}|0\rangle$ is a one-particle state with momentum $\vec{k}$.

§21: Derive (4.75). This shows that indeed $\hat{P}^i$ is the momentum operator which was constructed out of the stress tensor.

Let us summarize what we have learned:

1. The vacuum $|0\rangle$ is empty. With normal ordering, it has no energy and no charge.
2. The state $\hat{a}^\dagger_{\vec{k}}|0\rangle$ is a single-particle state. It has energy $\omega_{\vec{k}}$, mass m, momentum k^i, and charge $+1$.
3. The state $\hat{b}^\dagger_{\vec{k}}|0\rangle$ is a single-*antiparticle* state. It has energy $\omega_{\vec{k}}$, mass m, momentum k^i, and charge -1.

These particles are momentum eigenstates, and therefore completely delocalized in position space.

Spin

Now let us compute the "spin" of the single particle states we constructed above. For this, we first need to define a notion of *intrinsic spin*. From our knowledge of quantum mechanics, we expect the concept of spin to be related with the conserved charge for infinitesimal rotations: $J^i = \frac{1}{2}\epsilon^{ijk}M_{jk}$ as defined in Ex.(16). However, as we shall see, since these capture the notion of angular momentum, there will be an *intrinsic* part and an *orbital* associated with them. This is in line with how we get an intrinsic angular momentum – which we defined *spin* in quantum mechanics and the usual angular momentum was just the orbital angular momentum. So, to isolate the spin part, we shall look at the action of the corresponding operator valued conserved charges on single-particle states and then put their momenta to zero [30]. The eigenvalues we get will be the spin of these states. We will do this in the following exercise.

[30] Since $\vec{L} = \vec{r} \times \vec{p}$ and hence putting the momentum to zero leads to vanishing of the orbital angular momentum, classically. We could also set $\vec{r} = 0$ (or, in other words, just stay at the origin) to get $\vec{L} = 0$ and obtain the intrinsic part or the spin. However, since we will be dealing mostly with stuffs in the momentum space putting $\vec{p} = 0$ is more useful.

§22: Consider the rotational generators as defined in Ex.(16): $J_i = \frac{1}{2}\epsilon_{ijl}M^{jl}$. Write this in terms of x^j and T^{0k}. Then you can show that:

$$\begin{aligned}\hat{J}_i &= \int \frac{d^3k}{(2\pi)^3}\Big[\hat{a}^\dagger_{\vec{k}}\Big(-i\epsilon_{ijl}k^j\partial^l_{(k)}\Big)\hat{a}_{\vec{k}} + \hat{b}^\dagger_{\vec{k}}\Big(-i\epsilon_{ijl}k^j\partial^l_{(k)}\Big)\hat{b}_{\vec{k}}\Big],\\ &= \int \frac{d^3k}{(2\pi)^3}\Big[\hat{a}^\dagger_{\vec{k}}\big(-i\vec{k}\times\nabla_k\big)_i\,\hat{a}_{\vec{k}} + \hat{b}^\dagger_{\vec{k}}\big(-\vec{k}\times\nabla_k\big)_i\,\hat{b}_{\vec{k}}\Big], \qquad (4.76)\end{aligned}$$

where $\partial^l_{(k)} \equiv \frac{\partial}{\partial k_l}$ that is derivative w.r.t. the l^{th} component of the 3-vector $\vec{k}$.

Hint: To get (4.76) use the following delta distribution identity (see appendix C.3):

$$\int d^3x\, x^j e^{i(\vec{q}-\vec{k})\cdot\vec{x}} = -\frac{i}{(2\pi)^3}\frac{\partial}{\partial q_j}\delta^{(3)}(\vec{q}-\vec{k}), \qquad (4.77)$$

and then perform integration by parts (ignoring boundary terms) to move the derivative above away from the delta distribution. After this, do normal ordering too.

Next show that,

$$[\hat{J}_i, \hat{a}^\dagger_{\vec{p}}] = \hat{a}^\dagger_{\vec{p}}\Big(-i\epsilon_{ijl}p^j\partial^l_{(p)}\Big) \qquad (4.78)$$

implying that $\hat{J}_i\left(\hat{a}^\dagger_{\vec{p}}|0\rangle\right)\Big|_{\vec{p}=0} = 0$ and hence the single-particle states are spin 0 states [a].

[a] Do not put $\vec{p}=0$ before acting with $\hat{J}_i$ since that will be just like acting $\hat{J}_i$ on the vacuum itself which is Lorentz invariant as we will see in the next section when we discuss quantum symmetries.

As we saw in the above exercise that the excitations of the complex scalar field have spin 0. These are what are called *bosons* [31]. We shall see soon how this is connected to the commutator structure of the ladder operators and that will reveal to us their statistics under particle exchange. For now, let us note that the action of the Lorentz generators on the classical scalar fields was done by the generators $L^{\rho\sigma}$ as obtained in (2.43). Comparing this to their action on vector fields – as in (2.46) – we notice that the scalar action didn't have the spacetime independent $M^{\rho\sigma}$ part and only had the spacetime dependent (or, the orbital) part: $x^\rho\partial^\sigma - x^\sigma\partial^\rho$. So, it is the $M^{\rho\sigma}$ part of the generators that is the intrinsic part. An intuitive way to see this is to recall that, we argued, the $M^{\rho\sigma}$ part in (2.46) led to the transformation of the vector index in the vector field $A^\mu(x)$ and the orbital part took care of the transformation of the argument x just like the scalar field case. Now, since the scalar field has no index this intrinsic part of $M^{\rho\sigma}$ there was just zero. We also see that to obtain the $M^{\rho\sigma}$ part in (2.46) we have to put the translation generators: $P_\sigma \equiv \partial_\sigma$ to zero which in the quantum theory means that we put the state's momentum to zero after acting the state with the Lorentz generators. This is exactly why we defined spin this way above.

[31] In general, they have integer spins and obviously 0 is an integer.

Noether charge for spacetime symmetries

So, to summarise, we have seen that the Hamiltonian $\hat{P}^0$ is the Noether charge for time translational symmetry whose eigenvalues $\omega_{\vec{k}}$ give the energy of the state. $\hat{P}^i$s are the Noether charge for the spatial translational symmetry whose eigenvalues p^i denote the momentum of the state and $\hat{J}_i$s are the Noether charge for the rotational symmetry whose eigenvalues denote the angular momentum of the states. To get the spin from these eigenvalues we then put the spatial momenta p^i to zero for these states.

4.8 The Fock space

We now extend this to the full multi-particle space. A generic state is obtained by acting repeatedly with creation operators. For example, a state with N_a particles and N_b antiparticles is

$$|\{\vec{p}_1,\vec{p}_2,\ldots,\vec{p}_{N_a}\};\{\vec{q}_1,\vec{q}_2,\ldots,\vec{q}_{N_b}\}\rangle = \hat{a}^\dagger_{\vec{p}_1}\hat{a}^\dagger_{\vec{p}_2}\cdots\hat{a}^\dagger_{\vec{p}_{N_a}}\,\hat{b}^\dagger_{\vec{q}_1}\hat{b}^\dagger_{\vec{q}_2}\cdots\hat{b}^\dagger_{\vec{q}_{N_b}}|0\rangle. \tag{4.79}$$

The operator $\hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}}$ counts the number of particles with momentum $\vec{p}$, while $\hat{b}^\dagger_{\vec{q}}\hat{b}_{\vec{q}}$ counts antiparticles. Thus the charge operator simply counts the difference:

$$Q = N_a - N_b, \tag{4.80}$$

and the Hamiltonian gives the total energy

$$H = \sum_{i=1}^{N_a}\omega_{\vec{p}_i} + \sum_{j=1}^{N_b}\omega_{\vec{q}_j}. \tag{4.81}$$

Indeed above are the eigenvalues of an energy and charge eigenstate. As discussed earlier, we have now constructed a space of states that allows particle number to change.

This defines the *Fock space*: the Hilbert space spanned by all states obtained by applying any number of creation operators to the vacuum. A convenient basis is given by specifying the occupation numbers $\{n_{\vec{p}_i}, n_{\vec{q}_j}\}$.

Two-particle example

As an illustration, consider the two-particle state with momenta $\vec{p}_1$ and $\vec{p}_2$:

$$|\vec{p}_1,\vec{p}_2\rangle = \hat{a}^\dagger_{\vec{p}_1}\hat{a}^\dagger_{\vec{p}_2}|0\rangle. \tag{4.82}$$

Since the $\hat{a}^\dagger$ operators commute, we find

$$|\vec{p}_2,\vec{p}_1\rangle = \hat{a}^\dagger_{\vec{p}_2}\hat{a}^\dagger_{\vec{p}_1}|0\rangle = \hat{a}^\dagger_{\vec{p}_1}\hat{a}^\dagger_{\vec{p}_2}|0\rangle = |\vec{p}_1,\vec{p}_2\rangle. \tag{4.83}$$

Therefore, the order of creation does not matter. The two-particle state is exactly the same regardless of interchange, and there is no distinction between writing "first" or "second". This is the statement that the quanta of the scalar field are *bosons*: particles with commuting creation operators. Also, as we have seen earlier they are spin 0 particles.

4.9 Position-space particle creation

So far, we have described states with definite momentum [32]. In quantum mechanics, one may also be interested in creating states localized at a definite position. This is achieved by superposing momentum eigenstates with the appropriate plane-wave weights.

In our case, this structure is already built into the field operator itself. The field expansion is

$$\hat{\phi}(\vec{x},t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \Big(\hat{a}_{\vec{k}}\, e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}} + \hat{b}^{\dagger}_{\vec{k}}\, e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} \Big). \tag{4.84}$$

Now let us act with $\hat{\phi}(\vec{x},t)$ on the vacuum. We obtain

$$\hat{\phi}(\vec{x},t)|0\rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \hat{b}^{\dagger}_{\vec{k}}\, e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}}\, |0\rangle. \tag{4.85}$$

The $\hat{a}_{\vec{k}}$ term annihilates the vacuum, leaving only the $\hat{b}^{\dagger}_{\vec{k}}$ term. Thus $\hat{\phi}(\vec{x},t)$ creates a superposition of antiparticle states, with wavefunction $e^{-i\vec{k}\cdot\vec{x}}$. In other words, $\hat{\phi}(\vec{x},t)$ creates an *antiparticle* localized at $(\vec{x},t)$.

Similarly, consider the Hermitian conjugate field $\hat{\phi}^{\dagger}(\vec{x},t)$:

$$\hat{\phi}^{\dagger}(\vec{x},t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \Big(\hat{a}^{\dagger}_{\vec{k}}\, e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} + \hat{b}_{\vec{k}}\, e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}} \Big). \tag{4.86}$$

Acting on the vacuum, we find

$$\hat{\phi}^{\dagger}(\vec{x},t)|0\rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \hat{a}^{\dagger}_{\vec{k}}\, e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}}\, |0\rangle. \tag{4.87}$$

Here the $\hat{b}_{\vec{k}}$ term annihilates the vacuum, so only the $\hat{a}^{\dagger}_{\vec{k}}$ piece survives. Thus $\hat{\phi}^{\dagger}(\vec{x},t)$ creates a *particle* localized at position $(\vec{x},t)$. Now, recall that under a discrete C transformation $\phi \to \phi^*$ which in operator language imply: $\hat{\phi} \to \hat{\phi}^{\dagger}$ and thus under charge conjugation, as the name suggests, particles are mapped to anti-particles and vice-versa [33].

Operator-valued distributions

Although we have referred to the states $|\vec{k}\rangle$ as "particles", they are not localized in space. They are momentum eigenstates, which are not normalizable in the Hilbert space sense, but only in the sense of delta-functions. In ordinary quantum mechanics, neither momentum eigenstates nor position eigenstates are normalizable; instead, they are idealized distributions. The same

[32] We did this simply becasue solving the EoM was easier to solve in momentum space. Often, in QFT this is the case: things strongly coupled in position space will be decoupled and non-interacting in momentum space. When you will do interacting QFT you will notice that momentum space Feynman rules are more intuitive and easier to write down than their position space counterparts. Hence the slogan: *think in position space but compute in momentum space*.

[33] Under charge conjugation ϕ is mapped to itself for a real scalar field ϕ and hence as we know in that case a particle is its own anti-particle. Here, charge conjugation is the trivial as in just the identity transformation. Note, this is not the $\mathbf{Z}_2$ symmetry of the real scalar field which maps: $\phi \to -\phi$ instead.

situation arises in quantum field theory. The field operators $\hat{\phi}(\vec{x})$ and the annihilation operators $\hat{a}_{\vec{k}}$ are not good operators on the Fock space because they do not create normalizable states. For example,

$$\langle 0|\hat{a}_{\vec{k}}\hat{a}^{\dagger}_{\vec{k}}|0\rangle = \langle \vec{k}|\vec{k}\rangle = (2\pi)^3\delta^{(3)}(0), \tag{4.88}$$

and, even though,

$$\langle 0|\hat{\phi}(\vec{x})|0\rangle = 0, \tag{4.89}$$

we still have:

$$\langle 0|\hat{\phi}(\vec{x})\hat{\phi}^{\dagger}(\vec{x})|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\omega_{\vec{k}}} \to \infty. \tag{4.90}$$

Thus, these objects are best regarded as *operator-valued distributions*, not ordinary functions. Even though $\hat{\phi}(\vec{x})$ has a well-defined vacuum expectation value the fluctuations at a fixed point are infinite [34]:

§23: Derive (4.89) and (4.90).

To construct well-defined states, we must *smear* these distributions against smooth functions. For example, we may create a wavepacket

$$|\varphi\rangle = \int \frac{d^3k}{(2\pi)^3}\, e^{-i\vec{k}\cdot\vec{x}}\, \varphi(\vec{k})\,|\vec{k}\rangle, \tag{4.91}$$

which is localized in both position and momentum space. A common choice is a Gaussian wavepacket, for instance

$$\varphi(\vec{k}) \sim \exp\Big(-\tfrac{\vec{k}^2}{2m^2}\Big).$$

Note: Smearing does not change the physics. It simply makes rigorous what is already implicit: in experiments, detectors always have finite resolution in both space and momentum. Thus, the idealized delta-function states are never directly realized, but any physically measurable quantity is unchanged by replacing them with smooth, localized wavepackets. In this sense, smearing is a mathematical tool to regulate distributions, without affecting the observable content of the theory.

5 Propagators and Causality

We now turn to the question of how particles *propagate* from one point to another. As it turns out, there are several notions of a propagator. Before introducing them, let us pause for a mathematical interlude: how does one define a Lorentz-invariant measure when integrating over 3-momenta?

[34] Recall, for a random variable X, its mean is: $\langle X\rangle$ and its fluctuations are: $\langle X^2\rangle$ - $\langle X\rangle^2$. Hence, 2-pt. functions capture quantum fluctuations. Here these divergent quantum fluctuations are again a result of the infinite degrees of freedom in QFT.

From this point onward, I will drop the hats on operators, since essentially all the objects we consider will be operators. It should be clear from the context whether a symbol denotes an operator or a number.

5.1 Lorentz-invariant integration measures

It is important to note that the naive integral $\int d^3k$ is not Lorentz invariant. If we imagine $\vec{k}$ as part of a 4-vector k_μ, then under a Lorentz boost the measure transforms with additional γ factors. Thus we seek a better definition of an invariant integration measure.

A manifestly Lorentz-invariant expression is

$$\int d^4k\, \delta(k_\mu k^\mu - m^2), \tag{5.1}$$

which is an integral over all 4-momenta, with the delta-function enforcing the *on-shell* condition that the 4-momentum has invariant magnitude m^2. Geometrically, this restricts the integration to a 3-dimensional hyperboloid in momentum space (see fig. 3).

Let us evaluate the integral over k_0 first:

$$\int d^3k\, dk_0\, \delta(k_0^2 - \omega_{\vec{k}}^2), \tag{5.2}$$

where $\omega_{\vec{k}} = \sqrt{\vec{k}^2 + m^2}$. The delta-function can be solved by picking the positive- or negative-energy root:

$$k_0 = \pm\sqrt{\vec{k}^2 + m^2} = \pm\omega_{\vec{k}}. \tag{5.3}$$

Taking the positive-energy root $k^0 = +\omega_{\vec{k}}$, we can use the delta-function pole identity [35] (see appendix C.4)

$$\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|} \tag{5.4}$$

to evaluate:

$$\int d^3k\, dk^0\, \delta(k_0^2 - \omega_{\vec{k}}^2) = \int d^3k \left[\frac{\delta(k_0 + \omega_{\vec{k}}) + \delta(k_0 - \omega_{\vec{k}})}{|2k_0|}\right]_{k_0 = \omega_{\vec{k}}} = \int \frac{d^3k}{2\omega_{\vec{k}}}. \tag{5.5}$$

where in the last eqaulity we used the fact that $k_0 = +\omega_{\vec{k}}$. Thus the Lorentz-invariant measure is

$$\int \frac{d^3k}{2\omega_{\vec{k}}}, \tag{5.6}$$

which represents integration over the hyperboloid in momentum space defined by the mass-shell condition $k^2 = m^2$ as in Fig. 3.

[35] where, $f(x_i) = 0$ and hence x_is are roots of $f(x)$

5.2 Commutators and the light-cone

Let us now note a subtle point. In the Hamiltonian formalism, which we used for quantization, we picked out a preferred time coordinate. For example, our starting commutator was

$$[\phi(\vec{x},t),\pi(\vec{y},t)] = i\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{5.7}$$

which holds only at *equal times* in a given reference frame. This seems to clash with Lorentz invariance. To check this, we can use causality as a diagnostic.

Recall the idea of the light-cone: if an event occurs at x, then the set of events y that can be influenced by it are those with timelike or lightlike separation: $(x-y)^2 \geq 0$. While if two events are spacelike separated, $(x-y)^2 < 0$, then no information should be able to propagate between them. In operator terms this means

$$[\mathcal{O}_1(x),\mathcal{O}_2(y)] = 0 \qquad \text{for } (x-y)^2 < 0 \text{ (spacelike separation)}. \tag{5.8}$$

This requirement is called *microcausality*. It ensures that operators localized at spacelike separation commute, so measurements performed outside each other's light-cones cannot interfere.

We are now in a position to check this explicitly for the complex scalar field. From the Fourier expansion of ϕ and $\phi^\dagger$ as given in (4.24)-(4.25), one finds,

$$[\phi(x),\phi^\dagger(y)] = \int \frac{d^3k}{(2\pi)^3\,2\omega_{\vec{k}}}\left(e^{-ik\cdot(x-y)} - e^{+ik\cdot(x-y)}\right), \tag{5.9}$$

where we have introduced the 4-momentum k^μ with components

$$k^\mu = (\omega_{\vec{k}},\vec{k}), \qquad k_\mu = (\omega_{\vec{k}},-\vec{k}), \qquad k\cdot(x-y) = \omega_{\vec{k}}(t_x-t_y) - \vec{k}\cdot(\vec{x}-\vec{y}). \tag{5.10}$$

§24: Derive (5.9).

The result above is not an operator but a c-number [36]. It is also noteworthy that the expression is built from Lorentz-invariant ingredients: the invariant measure $\frac{d^3k}{2\omega_{\vec{k}}}$ and the invariant scalar product $k\cdot(x-y)$. Thus, although our quantization procedure initially appeared to break manifest Lorentz invariance, the resulting commutators are in fact consistent with Lorentz covariance and, as we will see, with causality.

Timelike vs. spacelike separation

Let us now analyze the commutator (5.9) in the two distinct cases: timelike and spacelike separations.

Suppose $(x-y)^\mu$ is a timelike vector. Then there exists a Lorentz boost Λ which rotates it to point purely along the time axis:

$$\Lambda^\mu_{\ \nu}(x-y)^\nu = (x^0-y^0,0,0,0)^\mu. \tag{5.11}$$

[36] This is only true for free theories and not for interacting theories in which case commutators in general give rise to operators.

Evaluating the commutator in this frame gives

$$\int \frac{d^3k}{(2\pi)^3\,2\omega_{\vec{k}}}\Big(e^{-i\omega_{\vec{k}}(x^0-y^0)}-e^{+i\omega_{\vec{k}}(x^0-y^0)}\Big)=-2i\int \frac{d^3k}{(2\pi)^3\,2\omega_{\vec{k}}}\sin(\omega_{\vec{k}}(x^0-y^0))\neq 0, \quad (5.12)$$

which is a nonzero oscillatory function of (x^0-y^0). Thus the commutator does not vanish for timelike separations, as expected: operators at timelike separation can influence one another.

Now suppose $(x-y)^\mu$ is spacelike. Then there exists a Lorentz boost Λ which rotates it to point purely in a spatial direction:

$$\Lambda^\mu_{\ \nu}(x-y)^\nu=(0,\,\vec{x}-\vec{y})^\mu. \quad (5.13)$$

In this frame the commutator becomes

$$\int \frac{d^3k}{(2\pi)^3}\,\frac{1}{2\sqrt{\vec{k}^2+m^2}}\Big(e^{+i\vec{k}\cdot(\vec{x}-\vec{y})}-e^{-i\vec{k}\cdot(\vec{x}-\vec{y})}\Big). \quad (5.14)$$

Changing variables $\vec{k}\to-\vec{k}$ in the second exponential shows that the two terms cancel:

$$\int \frac{d^3k}{(2\pi)^3}\,\frac{1}{2\sqrt{\vec{k}^2+m^2}}\Big(e^{+i\vec{k}\cdot(\vec{x}-\vec{y})}-e^{+i\vec{k}\cdot(\vec{x}-\vec{y})}\Big)=0. \quad (5.15)$$

§25: Convince yourself that the above change of variables $\vec{k}\to-\vec{k}$ indeed keeps the integration measure invariant and hence the second integral above remains the same under this change. Recall, we did a similar change of variables while evaluating the commutator for the ladder operators $\hat{a}_{\vec{k}}$ and $\hat{a}^\dagger_{\vec{k}}$.

Therefore, for any spacelike separation $(x-y)^2<0$ [37],

$$[\phi(x),\phi^\dagger(y)]=0, \qquad \text{if } (x-y) \text{ is spacelike.} \quad (5.16)$$

This confirms the principle of *microcausality*: field operators commute at spacelike separation. Hence quantum field theory, even though quantized in a Hamiltonian framework with a preferred time, does not allow information to propagate outside the light-cone. This property is Lorentz invariant and holds for all local operator commutators even in interacting QFTs!

5.3 Particle propagation

Let us now ask a different question. Somewhat imprecisely: if we create an anti-particle at one spacetime point $x^\mu=(x^0,\vec{x})$, what is the amplitude of finding it at a different spacetime point $y^\mu=(y^0,\vec{y})$?

There are actually several different possible answers to this question. The most natural candidate is the vacuum expectation value

$$\langle 0|\phi^\dagger(y)\phi(x)|0\rangle. \quad (5.17)$$

[37] It's instructive to compare this to the above computation and ask what the exact difference is: there is no change of variables that take: $\omega_{\vec{k}}\to-\omega_{\vec{k}}$ since the time direction is unidirectional inside the light cone unlike the spatial directions.

From the expansions we already derived, recall

$$\phi(\vec{x},t)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{k}}}}\, b^\dagger_{\vec{k}}\, e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}}\,|0\rangle, \tag{5.18}$$

$$\langle 0|\phi^\dagger(\vec{x},t) = \langle 0|\int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{k}}}}\, b_{\vec{k}}\, e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}}. \tag{5.19}$$

Putting these into (5.17), we find

$$\langle\phi^\dagger(y)\phi(x)\rangle = \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{k}_1}}\sqrt{2\omega_{\vec{k}_2}}}\, e^{ik_1\cdot x - ik_2\cdot y}\,\langle 0|b_{\vec{k}_2}\, b^\dagger_{\vec{k}_1}\,|0\rangle, \tag{5.20}$$

Commuting $b_{\vec{k}_2}$ past $b^\dagger_{\vec{k}_1}$ and using the commutator, this simplifies to [38]

$$\langle\phi^\dagger(y)\phi(x)\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\vec{k}}}\, e^{ik\cdot(x-y)}. \tag{5.21}$$

This object is called a *propagator* or *correlator*, since it is a two-point correlation function of fields [39]. Note again that it is manifestly Lorentz invariant: the measure $d^3k/(2\omega_{\vec{k}})$ is Lorentz invariant, and $k\cdot(x-y)$ is a Lorentz scalar.

Although we won't evaluate this integral explicitly here [40], it can be done in terms of Bessel functions after writing the above integral in polar coordinates. At large spacelike separation $|\vec{x}-\vec{y}|\to\infty$, one finds asymptotically

$$\langle\phi^\dagger(y)\phi(x)\rangle \;\sim\; e^{-m|\vec{x}-\vec{y}|}. \tag{5.22}$$

Unlike the commutator, which vanishes outside the light-cone, the propagator does not vanish for spacelike separation. This is not in conflict with causality: it reflects quantum fluctuations of the vacuum and the fact that the two fields are correlated due to *quantum entanglement*.

Vacuum entanglement

Suppose the vacuum factorized across two spacelike-separated regions A, B: $|0\rangle = |0\rangle_A \otimes |0\rangle_B$. Then correlators of A-local and B-local operators would factorize,

$$\langle 0|\, O_A(x)\, O_B(y)\,|0\rangle = \langle 0|O_A(x)|0\rangle\,\langle 0|O_B(y)|0\rangle, \qquad (x\in A,\ y\in B,\ (x-y)^2<0).$$

For a scalar field we know from Ex.(22) that $\langle 0|\phi(x)|0\rangle = 0$, so the RHS above is 0. But the LHS is generally nonzero at spacelike separation,

$$\langle 0|\phi(x)\phi(y)|0\rangle \neq 0 \quad \text{(it decays with distance but is not identically zero)},$$

[38] Here, we use the notation that expectations values are usually evaluated on the vacuum state: $\langle 0|X|0\rangle \equiv \langle X\rangle$.

[39] Recall, we have already seen a version of this in the beginning of this course.

[40] The computation here is very similar to what we did with square-root branch cuts in the beginning of the course.

contradicting factorization. Hence the vacuum cannot be a product state; it is entangled across any spatial cut. This tells us that the vacuum is a *highly entangled state*.

Thus, we see quantum field theory allows correlations to extend outside the light-cone, but not the transmission of information or signals.

We cannot use entanglement to transmit information faster than light. In a free theory (where the field commutator is a c-number or *commuting number*), we have

$$[\phi(x), \phi^\dagger(y)] = \big\langle [\phi(x), \phi^\dagger(y)] \big\rangle = \big\langle \phi(x)\, \phi^\dagger(y) \big\rangle - \big\langle \phi^\dagger(y)\, \phi(x) \big\rangle. \tag{5.23}$$

The first equality holds because, in a free theory, the commutator is a number (so it can be evaluated in any state). The second equality is just the expansion of the commutator.

Note, the two terms on the right are Lorentz-invariant amplitudes. The first, $\langle \phi(x)\phi^\dagger(y)\rangle$, is the amplitude for a *particle* to propagate from y to x, while the second, $\langle \phi^\dagger(y)\phi(x)\rangle$, is the amplitude for an *antiparticle* to propagate from x to y. For spacelike separations $(x-y)^2 < 0$, these two contributions cancel exactly, yielding $[\phi(x), \phi^\dagger(y)] = 0$ outside the light cone. Heuristically: the impossibility of superluminal signalling via entanglement is tied to the presence of antiparticles, whose contribution cancels the particle one outside the light cone, preserving microcausality.

5.4 Feynman propagator

The most useful propagator in quantum field theory is not the correlator we studied above, but a slightly different object called the *Feynman propagator*. To introduce it, let us first recall the notion of *time-ordering*.

Time-ordering

The time-ordered product of operators $\mathcal{O}_1(x_1), \mathcal{O}_2(x_2), \dots$ is denoted

$$T(\mathcal{O}_1(x_1)\mathcal{O}_2(x_2)\cdots), \tag{5.24}$$

and is defined as the correlation function obtained by ordering the operators according to their time arguments, with the latest times placed to the left and the earliest to the right. Schematically,

$$\mathcal{O}_a(x_a)\, \mathcal{O}_b(x_b) \cdots, \tag{5.25}$$

such that

$$t_a > t_b > t_c > \cdots > t_n. \tag{5.26}$$

As an example, the time-ordered product of ϕ and $\phi^\dagger$ is

$$T\big(\phi(x')\phi^\dagger(x)\big) = \theta(t'-t)\, \phi(x')\phi^\dagger(x) + \theta(t-t')\, \phi^\dagger(x)\phi(x'), \tag{5.27}$$

where $\theta(x)$ is the Heaviside step function, defined as

$$\theta(t) = \begin{cases} 1, & t > 0, \\ 0, & t < 0. \end{cases} \tag{5.28}$$

Note that the derivative of the step function is the delta function (see appendix C.2).

Definition and evaluation

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

$$G(x,x') = \langle 0|T\big(\phi(x')\phi^\dagger(x)\big)|0\rangle. \tag{5.29}$$

This object is of central importance because it is the correct notion of particle propagation [41] in interacting quantum field theories, where it naturally arises in scattering amplitudes.

Using the decomposition (5.27), we can compute the two cases separately. For $t' > t$, we need $\phi(x')\phi^\dagger(x)$, while for $t > t'$, we need $\phi^\dagger(x)\phi(x')$. Both integrals take the form

$$G(x,x') = \theta(t'-t)\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\vec{k}}}\, e^{-ik\cdot(x'-x)} + \theta(t-t')\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\vec{k}}}\, e^{ik\cdot(x'-x)}, \tag{5.30}$$

with $k^\mu = (\omega_{\vec{k}}, \vec{k})$ as usual and $k\cdot x = \omega_{\vec{k}} t - \vec{k}\cdot\vec{x}$.

Expanding explicitly, this can be written as

$$\begin{aligned} G(x,x') = \theta(t'-t)\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\vec{k}}}\, e^{-i\omega_{\vec{k}}(t'-t)+i\vec{k}\cdot(\vec{x}'-\vec{x})} \\ + \theta(t-t')\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\vec{k}}}\, e^{+i\omega_{\vec{k}}(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}. \end{aligned} \tag{5.31}$$

Now let us interpret the above. The point is that if $t' > t$, then the propagator evolves as $e^{-i\omega_{\vec{k}}(t'-t)}$, while if $t > t'$, it evolves as $e^{+i\omega_{\vec{k}}(t'-t)}$. This is the defining property of the Feynman propagator: it implements *causal time-ordering*. In particular, while the spatial dependence may look different between the two cases, this is only an artifact; in the integrals we can always relabel $\vec{k} \to -\vec{k}$ in one of the terms but no similar manipulation can be done for the time integral.

Covariant form

An elegant way to rewrite the Feynman propagator is

$$G(x,x') = i\int \frac{d^4k}{(2\pi)^4}\frac{e^{ik\cdot(x'-x)}}{k^2-m^2}, \tag{5.32}$$

where $k^\mu = (k^0,\vec{k})$ and $k^2 = k^\mu k_\mu = (k^0)^2 - \vec{k}^{\,2}$.

At first glance this looks quite different: we are now integrating over all four components of k^μ. To check equivalence with (5.31), let us explicitly perform the k^0 integral.

$$G(x,x') = i\int \frac{d^3k}{(2\pi)^3}\int_{-\infty}^{+\infty}\frac{dk^0}{2\pi}\frac{e^{ik^0(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}}{(k^0)^2-\omega_{\vec{k}}^2}, \tag{5.33}$$

where $\omega_{\vec{k}} = \sqrt{\vec{k}^{\,2}+m^2}$. Holding $\vec{k}$ fixed, the k^0-integral runs over the real line, but the integrand has poles at $k^0 = \pm\omega_{\vec{k}}$. Thus,

$$I = i\int \frac{dk^0}{2\pi}\frac{e^{ik^0(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}}{(k^0+\omega_{\vec{k}})(k^0-\omega_{\vec{k}})}. \tag{5.34}$$

[41] We will argue why this is the case shortly

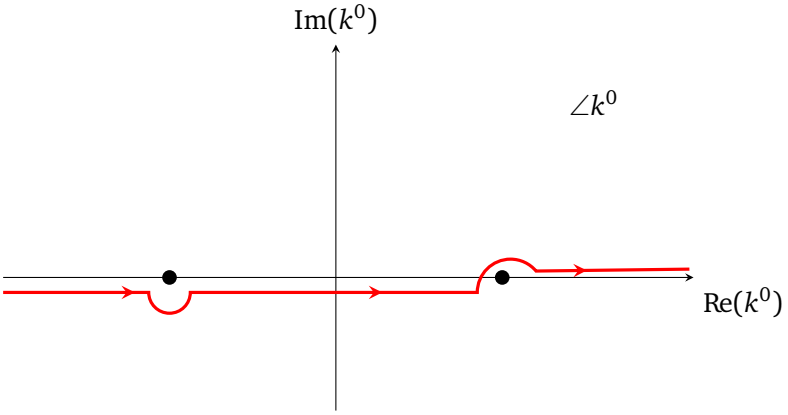


Figure 2: Contour prescription for the Feynman propagator

This integral is ill-defined as a standard real integral. To make sense of it, we deform k^0 into the complex plane and prescribe how to go around the poles.

By definition, the Feynman propagator corresponds to going *below* the pole at $k^0 = -\omega_{\vec{k}}$ and *above* the pole at $k^0 = +\omega_{\vec{k}}$ (see Fig. 2), we will see that this definition indeed is compatible with (5.31), that is, this prescription ensures consistency with the time-ordered form.

Contour evaluation of the Feynman propagator

Suppose $t'-t > 0$. Then we should close the contour in the upper half-plane, where $k^0 \to i\infty$, in the anti-clockwise direction, so that the factor

$$e^{ik^0(t'-t)} \sim e^{-\infty(t'-t)}$$

decays exponentially as $|k^0| \to \infty$. This picks up the pole at $k^0 = -\omega_{\vec{k}}$. The residue at this pole is

$$\text{Res}\Big|_{k^0=-\omega_{\vec{k}}} = \frac{i}{2\pi}\frac{e^{-i\omega_{\vec{k}}(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}}{-2\omega_{\vec{k}}}. \tag{5.35}$$

Using the residue theorem, we find

$$I = (2\pi i)\frac{-i}{2\pi}\frac{e^{-i\omega_{\vec{k}}(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}}. \tag{5.36}$$

Thus,

$$I = \frac{e^{-i\omega_{\vec{k}}(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}}, \qquad t' > t. \tag{5.37}$$

Now let us consider the opposite situation, $t'-t < 0$. In this case we must close the contour in the lower half-plane, where $k^0 \to -i\infty$, in the clockwise direction [42], so that

$$e^{ik^0(t'-t)} \sim e^{-\infty|t'-t|}$$

[42] Hence, we will pick an extra minus sign here

decays exponentially. This picks up the pole at $k^0 = +\omega_{\vec{k}}$. The residue is then

$$\text{Res}\Big|_{k^0=+\omega_{\vec{k}}} = \frac{i}{2\pi}\,\frac{e^{+i\omega_{\vec{k}}(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}}. \tag{5.38}$$

Since the contour orientation is reversed, the sign changes, and we obtain

$$I = \frac{e^{+i\omega_{\vec{k}}(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}}, \qquad t' < t. \tag{5.39}$$

Let us put everything together now. We see that the only difference between the two cases is the sign in the exponential, which depends on whether: $t'-t$ is positive or negative. Combining both branches, we recover the time-ordered form:

$$\begin{aligned} G(x,x') = i\int \frac{d^4k}{(2\pi)^4}\,\frac{e^{ik\cdot(x'-x)}}{k^2-m^2} &= \theta(t'-t)\int \frac{d^3k}{(2\pi)^3}\,\frac{1}{2\omega_{\vec{k}}}\,e^{-i\omega_{\vec{k}}(t'-t)+i\vec{k}\cdot(\vec{x}'-\vec{x})} \\ &+ \theta(t-t')\int \frac{d^3k}{(2\pi)^3}\,\frac{1}{2\omega_{\vec{k}}}\,e^{+i\omega_{\vec{k}}(t'-t)-i\vec{k}\cdot(\vec{x}'-\vec{x})}. \end{aligned} \tag{5.40}$$

In deriving the first term above we changed integration variable $\vec{k} \to -\vec{k}$. This matches exactly what we had in (5.31), confirming the equivalence. Thus we see that the covariant integral form (5.32), together with the contour prescription (going below $-\omega_{\vec{k}}$ and above $+\omega_{\vec{k}}$), gives a consistent Lorentz-covariant definition of the Feynman propagator. This will be the standard expression used throughout the rest of quantum field theory.

5.5 The $i\epsilon$ prescription

Often the contour prescription for the Feynman propagator is captured compactly as

$$G(x,x') = i\int \frac{d^4k}{(2\pi)^4}\,\frac{e^{ik\cdot(x'-x)}}{k^2-m^2+i\epsilon}, \tag{5.41}$$

where $\epsilon > 0$ is an infinitesimal positive number. Its role is to shift the poles slightly off the real axis in exactly the way required by the Feynman contour prescription.

To see this, consider the pole equation with the ϵ included:

$$(k^0)^2 - \vec{k}^{\,2} - m^2 + i\epsilon = 0 \quad \Rightarrow \quad (k^0)^2 = \omega_{\vec{k}}^2 - i\epsilon, \tag{5.42}$$

so that for small ϵ, we get:

$$k^0 = \pm\omega_{\vec{k}}\left(1 - \frac{i\epsilon}{2\omega_{\vec{k}}^2}\right). \tag{5.43}$$

Thus the negative and positive poles are shifted slightly above or below the real axis respectively and hence, for the negative pole we evaluated it in the upper half plane and for the positive pole we evaluated it in the lower half plane. This ensures that the contour avoids them consistently. This compact expression (5.41) is what is usually referred to as the *Feynman $i\epsilon$ prescription*.

Green's function property

The Feynman propagator also has another important property: it is the Green's function for the Klein–Gordon operator. Recall the Klein–Gordon equation:

$$(\Box + m^2)\phi(x) = 0, \tag{5.44}$$

with $\Box = \partial_t^2 - \nabla^2$.

Let us act with $(\Box_x + m^2)$ on the propagator:

$$(\Box_x + m^2)G(x,x') = i(\Box_x + m^2)\int \frac{d^4k}{(2\pi)^4}\,\frac{e^{ik\cdot(x'-x)}}{k^2 - m^2 + i\epsilon} \tag{5.45}$$

$$= i\int \frac{d^4k}{(2\pi)^4}\,\frac{-(k^2-m^2)\,e^{ik\cdot(x'-x)}}{k^2 - m^2 + i\epsilon} \tag{5.46}$$

$$= -i\int \frac{d^4k}{(2\pi)^4}\,e^{ik\cdot(x'-x)} = -i\,\delta^{(4)}(x'-x). \tag{5.47}$$

Thus the Feynman propagator satisfies:

$$(\Box_x + m^2)G(x,x') = i\delta^{(4)}(x'-x), \tag{5.48}$$

which means it is precisely the Green's function of the Klein–Gordon operator. This fact will be crucial when discussing interacting quantum field theory.

On-shell vs. off-shell (again)

The *mass shell* (or "shell") in momentum space is the two-sheeted hyperboloid (see Fig. 3),

$$k^2 = m^2 \quad \Longleftrightarrow \quad \omega = \pm\sqrt{\vec{k}^{\,2} + m^2}. \tag{5.49}$$

A four-momentum lying on this locus is called *on-shell*; otherwise it is *off-shell*. Note, for a free real scalar field the equations of motion are

$$(\Box + m^2)\,\phi(x) = 0, \qquad \Box \equiv \partial_t^2 - \nabla^2. \tag{5.50}$$

which in momentum space becomes,

$$(-k^2 + m^2)\,\tilde{\phi}(k) = 0. \tag{5.51}$$

Nontrivial solutions, that is $\tilde{\phi}(k) \neq 0$, therefore have support only on the shell,

$$\tilde{\phi}(k) \;\propto\; \delta(k^2 - m^2)\big[\theta(k^0)a(\vec{k}) + \theta(-k^0)a^\dagger(-\vec{k})\big], \tag{5.52}$$

so solutions of the (free) EoM are on-shell field configurations.

External vs. internal lines

Asymptotic in/out particles are free [43], hence their momenta satisfy $k^2 = m^2$ ("on-shell"). Internal lines are *propagators*, i.e. Green's functions of the differential operator,

$$(\Box + m^2)\, G(x) = -i\, \delta^{(4)}(x) \qquad \Longleftrightarrow \qquad \tilde{G}(k) = \frac{i}{k^2 - m^2 + i\varepsilon}. \tag{5.53}$$

A Green's function solves the EoM *with a source*, so it is not restricted to the shell; internal loop and exchange momenta are integrated over all k and are generically off-shell.

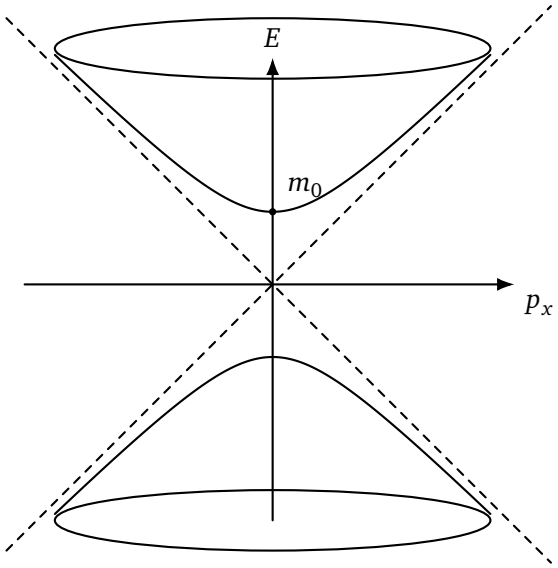


Figure 3: Mass-shell hyperboloid

6 Symmetries in QM

In this section we will discuss how symmetries of the classical field theory are represented in the quantum theory and how they act on states in the Hilbert space and on operators.

Let us first examine the properties required of a translation operator. Translating the particle should not change the probability density. Thus,

$$\langle\psi(x)|\psi(x)\rangle = \langle\psi(x+a)|\psi(x+a)\rangle = \langle\psi(x)|\hat{U}^\dagger(a)\hat{U}(a)|\psi(x)\rangle, \tag{6.1}$$

which implies

$$\hat{U}^\dagger(a)\hat{U}(a) = \mathbb{1}. \tag{6.2}$$

Therefore the translation operator is *unitary*. So, symmetry transformations acting on states must be implemented by unitary operators so as to preserve probabilities in the quantum the-

[43] You will learn all of these including what do we mean by *asymptotic states* in interacting QFT via what is called the *LSZ theorem*.

ory. However, action of symmetry transformations on fields need not be captured by unitaries as we will see in the next section.

Now let us see how translations act on quantum fields. If we translate both a state and an operator by the same vector distance $\vec{a}$, then nothing should change. In equations,

$$\langle y|\hat{\phi}(x)|y\rangle = \langle y+a|\hat{\phi}(x+a)|y+a\rangle. \tag{6.3}$$

Since $|y+a\rangle = \hat{U}(a)|y\rangle$, we must have

$$\hat{U}(a)\,\hat{\phi}(x)\,\hat{U}^{\dagger}(a) = \hat{\phi}(x+a). \tag{6.4}$$

This demonstrates that there are two consistent ways to picture translations, and we should not confuse them:

- Acting on states:
$$\hat{U}(a)|x\rangle = |x+a\rangle, \tag{6.5}$$
which moves a localized state from x to $x+\vec{a}$.

- Acting on locally defined operators:
$$\hat{U}^{\dagger}(a)\,\hat{\phi}(x)\,\hat{U}(a) = \hat{\phi}(x-a). \tag{6.6}$$

Thus, for a generic Lorentz transformations we will have (from (6.6)):

$$\hat{U}^{\dagger}(\Lambda)\,\hat{\phi}(x)\,\hat{U}(\Lambda) = \hat{\phi}(\Lambda^{-1}x), \tag{6.7}$$
$$\hat{U}^{\dagger}(\Lambda)\hat{A}^{\mu}(x)\,\hat{U}(\Lambda) = \Lambda^{\mu}{}_{\nu}\hat{A}^{\nu}(\Lambda^{-1}x). \tag{6.8}$$

§26: Show that it is enough to impose the canonical commutation relations at $t=0$. That is, once we impose them at $t=0$, then the relations at general t are automatically satisfied.

Hint: Think of the commutator as an operator itself: $\hat{C}(t,\vec{x}) \equiv [\hat{A}(t,\vec{x}), \hat{B}(t,\vec{x})]$. Let $\hat{A}$ and $\hat{B}$ be scalar operators. Consider the above commutator relation at $t=0$ and relate it to generic time t relation using (6.7).

§27: Consider translation generators P_μ and a field operator $\hat{\phi}(x)$. Let the unitary operator implementing a spacetime translation by a small four–vector a^μ be

$$U(a) \equiv e^{ia^\mu P_\mu}, \qquad U(a)^\dagger = e^{-ia^\mu P_\mu}.$$

The translated field satisfies

$$U(a)\,\phi(x)\,U(a)^\dagger \;=\; \phi(x+a). \tag{6.9}$$

Expand the right–hand side to first order in a^μ and write it in index–free form. Then use the following identity [a]:

$$e^A B e^{-A} = B + [A,B] + \frac{1}{2!}[A,[A,B]] + \frac{1}{3!}[A,[A,[A,B]]] + \cdots .$$

Take $A = ia^\mu P_\mu$ and $B = \phi(x)$ in the formula above and expand $U(a)\phi(x)U(a)^\dagger$ to first order in a^μ. Next equate RHS to LHS of (6.9). Read off the commutator $[P_\mu, \phi(x)]$. Show that you get:

$$[P_\mu,\, \phi(x)] = -i\,\partial_\mu \phi(x) \tag{6.10}$$

which implies that at the quantum level the momentum operator acts on fields via commutators to generate translations.

[a] Expanding the exponential for small A and keeping the order of operations right you should be able to see this identity. This is the adjoint form of the Baker–Campbell–Hausdorff expansion.

7 Dirac equation

We have seen thus far, how Lorentz transformations act on scalar and vector fields. We have also seen how scalar field produce excitations which have spin 0. These were called bosons and they satisfied equal time commutator relations after canonical quantisation. Furthermore, a multi-particle state was symmetric under exchange of their momenta label as was seen in (4.83). The purpose of this section is to consider a different set of particles which we will call *fermions* and they will have half-integral spin in general. Specifically, we will deal with spin $\frac{1}{2}$ particles. In nature, these describe the electrons, protons, etc.

To begin with let us start by noting a different representation of the Lorentz algebra which as we will see later will act on something called the *spinor field*. Let us recall that, a generic field $\phi^a(x)$ transforms under Lorentz transformations as:

$$\phi^a(x) \to U(\Lambda)^a{}_b\, \phi^b(\Lambda^{-1}x) \tag{7.1}$$

where $U(\Lambda)$ is a representation of the Lorentz group such that: $U(\Lambda_1)U(\Lambda_2) = U(\Lambda_1\Lambda_2)$ and

$U(e) = \mathbb{1}$ [44]. Note that, for scalar fields we had $U(\Lambda) = \mathbb{1}$ and for vector fields we had $U(\Lambda) = \Lambda$. The reason being a scalar field had no index while a vector field, by definition, had a vector index and hence its index transformed just like the index of a coordinate vector. Let us now try to find a representation of the Lorentz group under which a spinor field transforms. Recall, the Lorentz algebra is:

$$[M^{\rho\sigma}, M^{\mu\nu}] = \eta^{\sigma\mu}M^{\rho\nu} - \eta^{\rho\mu}M^{\sigma\nu} + \eta^{\rho\nu}M^{\sigma\mu} - \eta^{\sigma\nu}M^{\rho\mu}$$

7.1 The spinor representation

To construct the spinor representation, we start by defining the *Clifford algebra*,

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{1}, \tag{7.2}$$

where $\{A, B\} = AB + BA$ denotes the anti-commutator of square matrices. γ^μ ($\mu = 0, 1, 2, 3$) are four matrices and $\eta^{\mu\nu}$ above should be interpreted as a symbol with $\eta^{00} = 1$ and $\eta^{ii} = -1$ for all $i = 1, 2, 3$. This implies that

$$\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu \qquad \text{for } \mu \neq \nu, \tag{7.3}$$

and

$$(\gamma^0)^2 = +\mathbb{1}, \qquad (\gamma^i)^2 = -\mathbb{1}, \quad i = 1, 2, 3. \tag{7.4}$$

It is straightforward to check that there are no representations of the Clifford algebra using 2×2 or 3×3 matrices. Let us see this now.

Case: 2×2**:** Suppose γ^μ is a 2×2 matrix. Note that there are four generators of the Clifford algebra. Define $\widetilde{\gamma} \equiv \gamma^0\gamma^1\gamma^2\gamma^3$.

§28: Show that $\{\widetilde{\gamma}, \gamma^\mu\} = 0$. Next, using the cyclicity of trace show that $\text{tr}(\gamma^\mu) = 0$.

Note that if γ^μ is a 2×2 traceless matrix then the no. of independent components in it is $3 < 4$. Hence, 2×2 is ruled out for γ^μ.

Case: 3×3**:** Suppose γ^μ is a 3×3 matrix. Note that $(\gamma^0)^2 = \mathbb{1}$ and hence it is an involutory matrix. So, it is diagonalisable. Furthermore, it has 3 independent components which are its eigenvalues: ± 1 all along its diagonal. Denote V_+ as the eigenspace of $+1$ and V_- as the eigenspace of -1. Let $\dim(V_+) = m$, then $\dim(V_-) = 3 - m$. Also, $m \neq 3 - m$ for non-negative integral m. Now consider $v \in V_+$. By the Clifford algebra γ^i anti-commutes with γ^0. Hence,

$$\gamma^0(\gamma^i v) = -\gamma^i\gamma^0 v = -(\gamma^i v) \tag{7.5}$$

[44] Here, we are denoting e as the identity element for the Lorentz group and $\mathbb{1}$ as the identity matrix in the set $\{U(\Lambda) | \Lambda \in O(1,3)\}$. Furthermore, here $U(\Lambda)$ is just a notation and doesn't mean that $U(\Lambda)$ is unitary.

where in the last equality we used: $\gamma^0 v = v$ as $v \in V_+$. So, we see from above that $\gamma^i v \in V_-$. Since γ^is is invertible, we get that: $\gamma^i : V_+ \to V_-$ is a linear isomorphism, implying $\dim(V_+) = \dim(V_-)$, leading to a contradiction. Hence, γ^μ can't be a 3×3 matrix.

Thus, the simplest representation is 4×4 [45]. There are many such representations; one example is, called the *chiral representation*,

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{7.6}$$

where each block is a 2×2 matrix and σ^i are the Pauli matrices,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.7}$$

These satisfy

$$\{\sigma^i, \sigma^j\} = 2\delta^{ij}. \tag{7.8}$$

Now to see what the Clifford algebra has to do with the Lorentz algebra, define:

$$S^{\rho\sigma} = \frac{1}{4}[\gamma^\rho, \gamma^\sigma] \tag{7.9}$$

§29: Show that: $S^{\rho\sigma} = \frac{1}{2}(\gamma^\rho \gamma^\sigma - \eta^{\rho\sigma})$.

§30: Show that: $[S^{\mu\nu}, \gamma^\rho] = \gamma^\mu \eta^{\nu\rho} - \gamma^\nu \eta^{\rho\mu}$.

§31: Show that: $[S^{\rho\sigma}, S^{\mu\nu}] = \eta^{\sigma\mu} S^{\rho\nu} - \eta^{\rho\mu} S^{\sigma\nu} + \eta^{\rho\nu} S^{\sigma\mu} - \eta^{\sigma\nu} S^{\rho\mu}$.

Thus, from the above exercises we see that $S^{\mu\nu}$ form a representation of the Lorentz algebra.

7.1.1 Spinor fields

The $S^{\mu\nu}$ are 4×4 matrices, because the γ^μ are 4×4 matrices. We introduce the Dirac spinor field $\psi^\alpha(x)$, with four complex components labeled by $\alpha = 1, 2, 3, 4$. Under Lorentz transformations,

$$\psi^\alpha(x) \longrightarrow S[\Lambda]^\alpha{}_\beta \, \psi^\beta(\Lambda^{-1}x), \tag{7.10}$$

where [46]

$$\Lambda(\omega_{\rho\sigma}) = \exp\Big(\tfrac{1}{2}\,\omega_{\rho\sigma}\, M^{\rho\sigma}\Big), \tag{7.11}$$

$$\mathcal{S}[\Lambda(\omega_{\rho\sigma})] = \exp\Big(\tfrac{1}{2}\,\omega_{\rho\sigma}\, S^{\rho\sigma}\Big). \tag{7.12}$$

[45] Now we can go back to treating $\eta^{\mu\nu}$ as the Minkowski metric again since γ^μ is now a 4×4 matrix.

[46] the factor $\frac{1}{2}$ is for convention only

Although the bases of generators $M^{\rho\sigma}$ and $S^{\rho\sigma}$ are different, we use the same six parameters $\omega_{\rho\sigma}$ in both Λ and $\mathcal{S}[\Lambda]$; this ensures that we are describing the same Lorentz transformation on spacetime and on the spinor field.

Both Λ and $\mathcal{S}[\Lambda]$ are 4×4 matrices. To see that the spinor representation is truly different from the vector representation $\Lambda^{\mu}_{\ \nu}$, let us examine some specific transformations.

Rotations

The rotation generators are

$$S^{ij} = \tfrac{1}{2}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} = -\tfrac{i}{2}\,\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}, \qquad (i \neq j). \tag{7.13}$$

If we write the rotation parameters as $\omega_{ij} = -\epsilon_{ijk}\varphi^k$ (so that $\omega_{12} = -\varphi^3$, etc.), then the spinor representation of the Lorentz transformation is

$$\mathcal{S}[\Lambda] = \exp\Big(\tfrac{1}{2}\,\omega_{\rho\sigma}\,S^{\rho\sigma}\Big) = \begin{pmatrix} e^{+i\,\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{+i\,\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix}. \tag{7.14}$$

For example, a rotation by 2π around the x^3-axis is obtained by taking $\vec{\varphi} = (0,0,2\pi)$.

The spinor rotation matrix is

$$\mathcal{S}[\Lambda] = \begin{pmatrix} e^{+i\pi\sigma^3} & 0 \\ 0 & e^{+i\pi\sigma^3} \end{pmatrix} = -\mathbb{1}. \tag{7.15}$$

Therefore, under a 2π rotation,

$$\psi^{\alpha}(x) \;\longrightarrow\; -\psi^{\alpha}(x), \tag{7.16}$$

which is not what happens to a vector. Thus, $\mathcal{S}[\Lambda]$ is clearly a different representation from the familiar vector representation $\Lambda^{\mu}_{\ \nu}$.

Boosts

The boost generators are

$$S^{0i} = \tfrac{1}{2}\begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}. \tag{7.17}$$

Writing the boost parameter as $\omega_{i0} = -\omega_{0i} = \chi_i$, we obtain

$$\mathcal{S}[\Lambda] = \begin{pmatrix} e^{+\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix}. \tag{7.18}$$

Representations of the Lorentz Group are not Unitary

For rotations (7.14), $S[\Lambda]$ is unitary, satisfying $\mathcal{S}[\Lambda]^\dagger \mathcal{S}[\Lambda] = 1$. However, for boosts (7.18), $\mathcal{S}[\Lambda]$ is not unitary. In fact, there is a theorem due to Wigner that there are no finite-dimensional unitary representations of the Lorentz group.

Now quantum states belonging to a Hilbert space (defined on a time slice) describe physical excitations. Hence, they must transform in unitary representations of the Poincaré group so as to preserve *probabilities*. Since, states are labelled by momenta which is a continuous variable they furnish a infinite dimensional unitary representation.

Now let us discuss fields. Fields when seen as a function of spacetime obviously transform in an infinite dimensional representation (as we have seen earlier). However, let us consider evaluating the fields at one spacetime point say the origin. In this case, scalar fields which do not carry any index transform in a 1-dimensional representation of the Lorentz group while vector fields which carry the vector index transform in the 4-dimensional vector representation of the Lorentz group. These are allowed to transform under finite dimensional representations since they can be non-unitary (generically) as we saw earlier [47].

7.2 Constructing the Dirac action

We have now introduced a new field, the Dirac spinor $\psi(x)$ which is a spinor field. Our goal is to construct a Lorentz-invariant equation of motion by building a Lorentz-invariant action.

We begin in a naive way, which ultimately fails but will guide us. Define

$$\psi^\dagger(x) \;=\; (\psi^*)^T(x), \tag{7.19}$$

which is the usual adjoint of a multi-component object. One might try to form a Lorentz scalar by $\psi^\dagger\psi$, summing over spinor indices. Let us see how this transforms under Lorentz transformations:

$$\begin{aligned} \psi(x) \;&\longrightarrow\; \mathcal{S}[\Lambda]\psi(\Lambda^{-1}x), \\ \psi^\dagger(x) \;&\longrightarrow\; \psi^\dagger(\Lambda^{-1}x)\mathcal{S}[\Lambda]^\dagger. \end{aligned} \tag{7.20}$$

Hence

$$\psi^\dagger(x)\psi(x) \;\longrightarrow\; \psi^\dagger(\Lambda^{-1}x)\mathcal{S}[\Lambda]^\dagger\mathcal{S}[\Lambda]\psi(\Lambda^{-1}x). \tag{7.21}$$

But, as we have seen, for some Lorentz transformations $\mathcal{S}[\Lambda]^\dagger\mathcal{S}[\Lambda] \neq \mathbb{1}$, since the spinor representation is not unitary. Therefore $\psi^\dagger\psi$ is not a Lorentz scalar and does not transform nicely under the Lorentz group.

This failure, however, points the way forward. Choose a representation of the Clifford algebra (7.6) satisfying

$$(\gamma^0)^\dagger = \gamma^0, \qquad (\gamma^i)^\dagger = -\gamma^i, \qquad i = 1,2,3.$$

[47] Recall, Λ a generic Lorentz transformation is not orthogonal.

Then for all $\mu = 0, 1, 2, 3$,

$$\gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger. \tag{7.22}$$

This, in turn, implies

$$(S^{\mu\nu})^\dagger = \tfrac{1}{4}\left[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger\right] = -\gamma^0 S^{\mu\nu} \gamma^0. \tag{7.23}$$

From the Hermiticity property we had

$$\mathcal{S}[\Lambda]^\dagger = \exp\Big(\tfrac{1}{2}\omega_{\rho\sigma}(S^{\rho\sigma})^\dagger\Big) = \gamma^0 \mathcal{S}[\Lambda]^{-1} \gamma^0. \tag{7.24}$$

With this in mind, we define the *Dirac adjoint*:

$$\bar{\psi}(x) = \psi^\dagger(x)\gamma^0. \tag{7.25}$$

Claim: $\bar{\psi}\psi$ is a Lorentz scalar.

Proof. Under a Lorentz transformation,

$$\begin{aligned} \bar{\psi}(x)\psi(x) &= \psi^\dagger(x)\gamma^0\psi(x) \\ &\longrightarrow \ \psi^\dagger(\Lambda^{-1}x)\mathcal{S}[\Lambda]^\dagger\gamma^0\mathcal{S}[\Lambda]\psi(\Lambda^{-1}x) \\ &= \psi^\dagger(\Lambda^{-1}x)\gamma^0\psi(\Lambda^{-1}x) \\ &= \bar{\psi}(\Lambda^{-1}x)\psi(\Lambda^{-1}x). \end{aligned} \tag{7.26}$$

This is exactly the transformation law for a Lorentz scalar.

Claim: $\bar{\psi}\gamma^\mu\psi$ is a Lorentz vector.

This means

$$\bar{\psi}(x)\gamma^\mu\psi(x) \ \longrightarrow \ \Lambda^\mu_{\ \nu}\bar{\psi}(\Lambda^{-1}x)\gamma^\nu\psi(\Lambda^{-1}x). \tag{7.27}$$

Thus the index $\mu = 0, 1, 2, 3$ on the γ^μ matrices can be interpreted as a true Lorentz vector index, and Lorentz scalars can be formed by contracting it with other Lorentz indices.

Proof. Suppressing the x-argument, under a Lorentz transformation:

$$\bar{\psi}\,\gamma^\mu\psi \ \longrightarrow \ \bar{\psi}\,\mathcal{S}[\Lambda]^{-1}\gamma^\mu\mathcal{S}[\Lambda]\psi. \tag{7.28}$$

For this to transform as a vector we require

$$\mathcal{S}[\Lambda]^{-1}\gamma^\mu\mathcal{S}[\Lambda] \ = \ \Lambda^\mu_{\ \nu}\gamma^\nu. \tag{7.29}$$

We can show this by working infinitesimally. Expand

$$\Lambda = \exp\Big(\tfrac{1}{2}\omega_{\rho\sigma}M^{\rho\sigma}\Big) \ \approx \ 1 + \tfrac{1}{2}\omega_{\rho\sigma}M^{\rho\sigma} + \ldots, \tag{7.30}$$

$$\mathcal{S}[\Lambda] = \exp\Big(\tfrac{1}{2}\omega_{\rho\sigma}S^{\rho\sigma}\Big) \ \approx \ 1 + \tfrac{1}{2}\omega_{\rho\sigma}S^{\rho\sigma} + \ldots. \tag{7.31}$$

Thus the requirement (7.29) becomes

$$-[S^{\rho\sigma}, \gamma^\mu] = (M^{\rho\sigma})^\mu_{\ \nu}\gamma^\nu, \tag{7.32}$$

where we have suppressed the spinor indices on γ^μ and $S^{\rho\sigma}$, but kept all Lorentz indices explicit. In fact, (7.32) follows from Ex.(28), where we showed

$$[S^{\rho\sigma}, \gamma^\mu] = \gamma^\rho \eta^{\sigma\mu} - \gamma^\sigma \eta^{\rho\mu}.$$

To see this, expand out the vector representation M:

$$\begin{aligned}(M^{\rho\sigma})^\mu{}_\nu \gamma^\nu &= (\eta^{\rho\mu}\delta^\sigma_\nu - \eta^{\sigma\mu}\delta^\rho_\nu)\gamma^\nu \\ &= \eta^{\rho\mu}\gamma^\sigma - \eta^{\sigma\mu}\gamma^\rho,\end{aligned} \tag{7.33}$$

which means the proof reduces to showing

$$-[S^{\rho\sigma}, \gamma^\mu] = \eta^{\rho\mu}\gamma^\sigma - \eta^{\sigma\mu}\gamma^\rho, \tag{7.34}$$

exactly as established in Ex.(28).

Claim: $\bar\psi\gamma^\mu\gamma^\nu\psi$ transforms as a Lorentz tensor.

Proof. The argument follows as before. More precisely, the symmetric part, proportional to $\eta^{\mu\nu}\bar\psi\psi$, transforms as a Lorentz scalar while the antisymmetric part, proportional to $\bar\psi S^{\mu\nu}\psi$, transforms as a Lorentz tensor.

We now have three bilinears of the Dirac field which transform covariantly under the Lorentz group:

$$\bar\psi\psi, \qquad \bar\psi\gamma^\mu\psi, \qquad \bar\psi\gamma^\mu\gamma^\nu\psi.$$

From these we can build a Lorentz-invariant action. In fact, only the first two are needed. We choose

$$S = \int d^4x\, \bar\psi(x)\big(i\gamma^\mu\partial_\mu - m\big)\psi(x). \tag{7.35}$$

This is the *Dirac action*. The factor of i ensures the action is real: upon complex conjugation, it cancels the minus sign that arises from integration by parts. Equivalently, it plays the same role as in quantum mechanics, where the Hermitian momentum operator is $-i\nabla$.

After quantization, as we will see, this action describes particles and anti-particles of mass $|m|$ and spin $1/2$. Notice also that the Lagrangian is first-order in derivatives, unlike the second-order scalar field Lagrangian. Furthermore, the mass parameter m can appear with either sign in the Lagrangian [48].

7.2.1 Equations of Motion

The equation of motion follows from the action (7.35) by varying with respect to ψ and $\bar\psi$ independently. Varying with respect to $\bar\psi$ gives

$$\big(i\gamma^\mu\partial_\mu - m\big)\psi = 0. \tag{7.36}$$

[48] For a scalar field, we had m^2 instead

This is the *Dirac equation*. Varying with respect to ψ gives the conjugate equation

$$i\partial_\mu\bar{\psi}\,\gamma^\mu + m\bar{\psi} = 0. \tag{7.37}$$

The Dirac equation is first order in derivatives, yet Lorentz invariant. If we tried to write down a first-order equation of motion for a scalar field, it would look like $v^\mu\partial_\mu\phi = 0$, which involves a privileged vector v^μ and breaks Lorentz invariance. For spinors, the γ^μ matrices ensure Lorentz invariance.

The Dirac equation couples the components of ψ through the γ^μ matrices. Nevertheless, each component individually satisfies the Klein–Gordon equation. To see this, act on (7.36) with $(i\gamma^\nu\partial_\nu + m)$:

$$\begin{aligned}(i\gamma^\nu\partial_\nu + m)(i\gamma^\mu\partial_\mu - m)\psi &= -\big(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2\big)\psi \\ &= -(\partial_\mu\partial^\mu + m^2)\psi.\end{aligned} \tag{7.38}$$

Here we used $\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu = \frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu = \eta^{\mu\nu}\partial_\mu\partial_\nu = \partial_\mu\partial^\mu$. Thus each component ψ^α ($\alpha = 1,2,3,4$) obeys

$$(\partial_\mu\partial^\mu + m^2)\psi^\alpha = 0, \tag{7.39}$$

the Klein–Gordon equation.

The slash notation

It is useful to introduce the shorthand

$$\not{A} \equiv A_\mu\gamma^\mu. \tag{7.40}$$

In this notation, the Dirac equation (7.36) becomes

$$(i\not{\partial} - m)\psi = 0. \tag{7.41}$$

7.3 Weyl spinors

When we needed an explicit form of the γ^μ matrices, we used the *chiral representation*

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{7.42}$$

where σ^i are the Pauli matrices.

In this representation, the spinor rotation transformation $\mathcal{S}[\Lambda_{\text{rot}}]$ and boost transformation $\mathcal{S}[\Lambda_{\text{boost}}]$ computed earlier in (7.14) and (7.18) are block diagonal:

$$S[\Lambda_{\text{rot}}] = \begin{pmatrix} e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix}, \qquad S[\Lambda_{\text{boost}}] = \begin{pmatrix} e^{+\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix}. \tag{7.43}$$

Thus the Dirac spinor representation of the Lorentz group is *reducible*: it decomposes into two irreducible representations, acting only on two-component spinors $u_\pm$. In the chiral representation, we define

$$\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix}. \tag{7.44}$$

The two-component objects $u_\pm$ are called *Weyl spinors* or *chiral spinors*. They transform under rotations as

$$u_\pm \longrightarrow e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} u_\pm, \tag{7.45}$$

and under boosts as

$$u_\pm \longrightarrow e^{\pm\vec{\chi}\cdot\vec{\sigma}/2} u_\pm. \tag{7.46}$$

7.3.1 The Weyl Equation

Let us decompose the Dirac Lagrangian in terms of Weyl spinors. Substituting (7.44) into the Dirac Lagrangian, we find

$$\begin{aligned} \mathcal{L} &= \bar{\psi}\,(i\partial\!\!\!/ - m)\psi \\ &= i\,u_+^\dagger\,\sigma^\mu\partial_\mu u_+ + i\,u_-^\dagger\,\bar{\sigma}^\mu\partial_\mu u_- - m\big(u_+^\dagger u_- + u_-^\dagger u_+\big), \end{aligned} \tag{7.47}$$

where we introduced a compact notation for the Pauli matrices with Lorentz index $\mu = 0, 1, 2, 3$:

$$\sigma^\mu = (1, \sigma^i), \qquad \bar{\sigma}^\mu = (1, -\sigma^i). \tag{7.48}$$

From (7.47), we see that a massive fermion requires both u_+ and u_-, since they couple through the mass term. However, a massless fermion can be described by u_+ or u_- alone, with equations of motion

$$i\,\bar{\sigma}^\mu\partial_\mu u_+ = 0, \qquad i\,\sigma^\mu\partial_\mu u_- = 0. \tag{7.49}$$

These are the *Weyl equations*.

Degrees of Freedom

Let us comment on the degrees of freedom of a spinor. The Dirac fermion ψ has 4 complex components = 8 real components. How do we count its true degrees of freedom?

In classical mechanics, the number of degrees of freedom of a system equals the dimension of the configuration space, or equivalently half the dimension of the phase space. In field theory we have infinitely many degrees of freedom (one per spatial point), but it is still sensible to count *independent* degrees of freedom per point.

A real scalar field ϕ has one degree of freedom. At the quantum level this means one type of particle. A complex scalar field has two degrees of freedom, corresponding to the particle and antiparticle.

Now consider a Dirac spinor. One might naively think there are 8 degrees of freedom (from its components), but this is not correct. Unlike the scalar case, the equation of motion for a spinor is *first order*. In particular, for the Dirac Lagrangian, the momentum conjugate to ψ is

$$\pi_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger. \tag{7.50}$$

This is not proportional to $\dot{\psi}$. Hence the phase space for a spinor is parametrized by ψ and $\psi^\dagger$, rather than ψ and $\dot{\psi}$.

Thus the phase space of the Dirac spinor ψ has 8 real dimensions, corresponding to 4 real degrees of freedom. In the quantum theory, this manifests as: two spin states (up, down) for the particle and two spin states (up, down) for the antiparticle. Similarly, for a Weyl spinor, the counting shows there are two degrees of freedom in total.

7.4 γ^5 and Parity

The Lorentz group matrices $\mathcal{S}[\Lambda]$ came out block diagonal in (7.43) because we chose the specific representation (7.42). This is why the representation (7.42) is called the *chiral representation*: the decomposition of the Dirac spinor ψ was simply given by (7.44).

But what if we choose a different representation γ^μ of the Clifford algebra, so that

$$\gamma^\mu \longrightarrow U\gamma^\mu U^{-1}, \qquad \psi \longrightarrow U\psi \tag{7.51}$$

Then $\mathcal{S}[\Lambda]$ will not be block diagonal. To remedy this let us introduce the *fifth gamma matrix*,

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{7.52}$$

One checks that this satisfies

$$\{\gamma^5, \gamma^\mu\} = 0, \qquad (\gamma^5)^2 = +\mathbb{1}. \tag{7.53}$$

Moreover, one checks that $[S^{\mu\nu}, \gamma^5] = 0$, so γ^5 is a Lorentz scalar under rotations and boosts. Since $(\gamma^5)^2 = \mathbb{1}$, we can form Lorentz-invariant projection operators

$$P_\pm = \tfrac{1}{2}(\mathbb{1} \pm \gamma^5), \tag{7.54}$$

with properties $P_\pm^2 = P_\pm$ and $P_+P_- = 0$.

§32: Show that: $[P_\pm, \mathcal{S}[\Lambda]] = 0$ for rotations and boosts. This is what is meant by $\mathcal{S}[\Lambda]$ being *reducible*.

In the chiral representation (7.42), we have

$$\gamma^5 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \tag{7.55}$$

so $P_\pm$ project directly onto the Weyl spinors $u_\pm$. Thus, for an arbitrary representation of the Clifford algebra, we may define chiral spinors via

$$\psi_\pm = P_\pm \psi, \tag{7.56}$$

which furnish irreducible representations of the Lorentz group.

The field ψ_+ is called a *left-handed* spinor, and ψ_- is a *right-handed* spinor. This terminology arises from the way massless fermions propagate; we will return to this when we discuss *helicity*.

Parity

The spinors $\psi_\pm$ are related by the discrete symmetry of *parity*. The Lorentz group is defined by

$$x^\mu \longrightarrow \Lambda^\mu_{\ \nu} x^\nu, \qquad \Lambda^\mu_{\ \nu}\Lambda^\rho_{\ \sigma}\eta^{\nu\sigma} = \eta^{\mu\rho}. \tag{7.57}$$

So far we have considered only transformations Λ continuously connected to the identity, i.e. those with infinitesimal generators. But the Lorentz group also contains two discrete transformations:

$$\text{Time Reversal:} \quad T:\ x^0 \mapsto -x^0,\ \ x^i \mapsto x^i, \qquad \text{Parity:} \quad P:\ x^0 \mapsto x^0,\ \ x^i \mapsto -x^i. \tag{7.58}$$

We will not discuss time reversal further here [49]. Parity, however, plays an important role in the Standard Model, particularly in weak interactions, as you will see later.

Under parity, left- and right-handed spinors are exchanged. This follows because in the chiral representation, the rotation and boost transformations for Weyl spinors $u_\pm$ were

$$u_\pm \xrightarrow{\text{rot}} e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} u_\pm, \qquad u_\pm \xrightarrow{\text{boost}} e^{\pm\vec{\chi}\cdot\vec{\sigma}/2} u_\pm. \tag{7.59}$$

Rotations are unaffected, but boosts flip sign.

Now recall that under a Lorentz transformation $\bar\psi\gamma^\mu\psi$ behaves as a vector, implying: $\bar\psi\,\gamma^\mu\psi \to \Lambda^\mu_{\ \nu}\,\bar\psi\,\gamma^\nu\psi$ and for this to hold we showed that:

$$\mathcal{S}[\Lambda]^{-1}\gamma^\mu\mathcal{S}[\Lambda] = \Lambda^\mu_{\ \nu}\gamma^\nu \tag{7.60}$$

All this discussion resulted in identifying that the index in γ^μ can be treated as a spacetime index. Now we want to include parity. Note, that $(\Lambda_P)^\mu_{\ \nu} = \text{diag}(+1,-1,-1,-1)$. So, we can generalise (7.60) for parity, to get,

$$\mathcal{S}[\Lambda_P]^{-1}\gamma^\mu\mathcal{S}[\Lambda_P] = (\Lambda_P)^\mu_{\ \nu}\gamma^\nu \tag{7.61}$$

§33: Show that: $\mathcal{S}[\Lambda_P] = \gamma^0$ satisfies (7.61).

[49] it is implemented by an anti-unitary operator on states, see e.g. sec. 15.3 of the book *QFT for gifted amateurs*

Hence, we get, under parity:

$$\psi(t,\vec{x}) \to \gamma^0\psi(t,-\vec{x}) = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} u_+ \\ u_- \end{pmatrix} = \begin{pmatrix} u_- \\ u_+ \end{pmatrix} \tag{7.62}$$

Thus parity exchanges right- and left-handed spinors:

$$P:\ u_+ \longleftrightarrow u_-, \qquad \text{or equivalently } P:\ \psi_\pm(\vec{x},t) \mapsto \psi_\mp(-\vec{x},t). \tag{7.63}$$

Finally, note that if $\psi(\vec{x},t)$ satisfies the Dirac equation, then the parity-transformed spinor $\gamma^0\psi(-\vec{x},t)$ also satisfies it. Indeed,

$$(i\gamma^0\partial_t + i\gamma^i\partial_i - m)\gamma^0\psi(-\vec{x},t) = \gamma^0(i\gamma^0\partial_t - i\gamma^i\partial_i - m)\psi(-\vec{x},t) = 0, \tag{7.64}$$

where the minus sign from acting on $-\vec{x}$ is compensated by γ^0 conjugation.

7.5 Chiral interaction terms

Let us examine how interaction terms transform under parity. Consider the bilinears from which the Dirac action was built.

For the scalar,

$$P:\ \bar\psi\psi(\vec{x},t) \longrightarrow \bar\psi\psi(-\vec{x},t), \tag{7.65}$$

which is the correct transformation law for a scalar.

For the vector $\bar\psi\gamma^\mu\psi$, separate temporal and spatial components:

$$\begin{aligned} P:&\ \bar\psi\gamma^0\psi(\vec{x},t) \longrightarrow \bar\psi\gamma^0\psi(-\vec{x},t), \\ P:&\ \bar\psi\gamma^i\psi(\vec{x},t) \longrightarrow \bar\psi\gamma^0\gamma^i\gamma^0\psi(-\vec{x},t) = -\bar\psi\gamma^i\psi(-\vec{x},t). \end{aligned} \tag{7.66}$$

Thus $\bar\psi\gamma^\mu\psi$ transforms as a Lorentz vector: the time component is invariant, the spatial components flip sign. Similarly, $\bar\psi S^{\mu\nu}\psi$ transforms as an antisymmetric tensor.

Now that we have introduced γ^5, we can form additional bilinears:

$$\bar\psi\gamma^5\psi, \qquad \bar\psi\gamma^5\gamma^\mu\psi. \tag{7.67}$$

How do these transform under parity? Check:

$$P:\ \bar\psi\gamma^5\psi(\vec{x},t) \longrightarrow \bar\psi\gamma^0\gamma^5\gamma^0\psi(-\vec{x},t) = -\bar\psi\gamma^5\psi(-\vec{x},t), \tag{7.68}$$

$$\begin{aligned} P:\ \bar\psi\gamma^5\gamma^\mu\psi(\vec{x},t) &\longrightarrow \bar\psi\gamma^0\gamma^5\gamma^\mu\gamma^0\psi(-\vec{x},t) \\ &= \begin{cases} -\bar\psi\gamma^5\gamma^0\psi(-\vec{x},t), & \mu = 0, \\ +\bar\psi\gamma^5\gamma^i\psi(-\vec{x},t), & \mu = i, \end{cases} \end{aligned} \tag{7.69}$$

which shows that $\bar\psi\gamma^5\psi$ is a *pseudoscalar*, while $\bar\psi\gamma^5\gamma^\mu\psi$ is an *axial vector*.

To summarize, the independent spinor bilinears are:

$$\begin{aligned} \bar{\psi}\psi &: \text{ scalar}, \\ \bar{\psi}\gamma^{\mu}\psi &: \text{ vector}, \\ \bar{\psi}S^{\mu\nu}\psi &: \text{ tensor}, \\ \bar{\psi}\gamma^{5}\psi &: \text{ pseudoscalar}, \\ \bar{\psi}\gamma^{5}\gamma^{\mu}\psi &: \text{ axial vector}. \end{aligned} \tag{7.70}$$

Counting: 1+4+6+1+4 = 16 bilinears in total, which is exactly the number of independent components for a 4-component Dirac spinor. So, we can write any 4×4 matrix as a linear combinations of elements in the following set:

$$\Gamma = \{\mathbb{1}, \gamma^{\mu}, \gamma^{5}, \gamma^{\mu}\gamma^{5}, S^{\rho\sigma}\}, \tag{7.71}$$

and we can construct Dirac bilinears as $\bar{\psi}\,\Gamma\,\psi$ [50].

Dirac basis

There also exists another basis that is more useful when considering massive fermions called the *Dirac basis*. In this basis, γ^{i}s remain the same but $\gamma^{0} \leftrightarrow \gamma^{5}$. This justifies the relevance of γ^{5}. We will use this to solve for the Dirac equation later.

Helicity

§34: Plug plane-wave solutions

$$\psi(x) = u(p)\,e^{-ip\cdot x} = \begin{pmatrix} u_{+} \\ u_{-} \end{pmatrix} e^{-ip\cdot x}$$

into the massless Dirac equation (note that, $p^{0} = |\vec{p}|$ when $m = 0$). Show that u_{-} and u_{+} are eigenvectors of the helicity operator

$$h := \tfrac{1}{2}\begin{pmatrix} \hat{p}\cdot\vec{\sigma} & 0 \\ 0 & \hat{p}\cdot\vec{\sigma} \end{pmatrix}, \tag{7.72}$$

where $\hat{p}$ is the unit vector in the direction of $\vec{p}$. Without loss of generality, one may take $\hat{p}$ to lie in the z-direction. So, *helicity* is defined as the projection of spin operator along the direction of momentum.
List the eigenvalues of the left- and right-handed spinors. (After quantisation, these correspond to the physical situation in which the particle's spin is either aligned or anti-aligned with the direction of motion.)

[50] You might wonder, what about $\bar{\psi}\gamma^{5}S^{\mu\nu}\psi$? Turns out $\bar{\psi}\gamma^{5}S^{\mu\nu}\psi \propto \frac{i}{2}\epsilon^{\mu\nu\rho\sigma}\bar{\psi}S_{\rho\sigma}\psi$ and hence is not linear independent from $\bar{\psi}S_{\rho\sigma}\psi$ which we have already considered. To show this you can use: $\gamma^{5} \equiv -\frac{i}{4!}\epsilon_{\alpha\beta\theta\delta}\gamma^{\alpha}\gamma^{\beta}\gamma^{\theta}\gamma^{\delta}$.

Note that, massless particles, strictly speaking, don't have spin since to define spin we have to go to a rest frame and massless particles don't have a rest frame [51]. So, for them *helicity* is the analogue of spin.

7.6 Symmetries and Conserved Currents

The Dirac Lagrangian enjoys several continuous symmetries. We list them and compute the associated conserved currents.

Spacetime translations. Under translations, the spinor varies as

$$\delta\psi = \epsilon^\mu \partial_\mu \psi. \tag{7.73}$$

Since $\mathcal{L}$ depends on $\partial_\mu\psi$ but not on $\partial_\mu\bar{\psi}$, the canonical (Noether) energy–momentum tensor is

$$T^{\mu\nu} = i\,\bar{\psi}\,\gamma^\mu \partial^{\,\nu}\psi - \eta^{\mu\nu}\mathcal{L}. \tag{7.74}$$

For spinors the EOM are first order, $(i\partial\!\!\!/ - m)\psi = 0$, so on shell we can set $\mathcal{L} \to 0$ inside $T^{\mu\nu}$, yielding

$$T^{\mu\nu} = i\,\bar{\psi}\,\gamma^\mu \partial^{\,\nu}\psi. \tag{7.75}$$

The total energy is therefore

$$E = \int d^3x\, T^{00} = \int d^3x\, i\,\bar{\psi}\gamma^0\partial^0\psi = \int d^3x\, \psi^\dagger\gamma^0\big(-i\gamma^i\partial_i + m\big)\psi, \tag{7.76}$$

where the last equality uses the Dirac equation.

Lorentz transformations. For an infinitesimal Lorentz transformation,

$$\delta x^\mu = \Omega^\mu{}_\nu\, x^\nu, \qquad \delta\psi^\alpha = -\Omega^\mu{}_\nu x^\nu \partial_\mu\psi^\alpha + \tfrac{1}{2}\,\omega_{\rho\sigma}\,(S^{\rho\sigma})^\alpha{}_\beta\,\psi^\beta, \tag{7.77}$$

with $\Omega^\mu{}_\nu = \frac{1}{2}\omega_{\rho\sigma}(M^{\rho\sigma})^\mu{}_\nu$ and

$$(\mathcal{M}^{\rho\sigma})^\mu{}_\nu = \eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu. \tag{7.78}$$

After direct substitution, we find that $\Omega^{\mu\nu} = \omega^{\mu\nu}$. Thus,

$$\delta\psi^\alpha \;=\; -\Omega^{\mu\nu}\big[\, x_\nu\partial_\mu\psi^\alpha - \tfrac{1}{2}(S_{\mu\nu})^\alpha{}_\beta\psi^\beta\,\big]. \tag{7.79}$$

The conserved current associated with Lorentz transformations follows from the same calculation as for the scalar field, with two differences: using the Dirac equation we can set $\mathcal{L} = 0$, and there is an extra contribution from the second term in (7.79).

Thus the current is

$$(J^\mu)^{\rho\sigma} \;=\; x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} + i\bar{\psi}\gamma^\mu S^{\rho\sigma}\psi. \tag{7.80}$$

Upon quantization, $(J^\mu)^{\rho\sigma}$ becomes an operator. The additional spin term ensures that the single-particle states of the Dirac field carry intrinsic angular momentum, i.e. spin-1/2.

[51] Since light has the same speed c for inertial observers!

Internal Vector Symmetry

The Dirac Lagrangian is invariant under global phase rotations of the spinor,

$$\psi \;\to\; e^{-i\alpha}\psi,$$

which yields the conserved vector current

$$j_V^\mu \;=\; \bar{\psi}\gamma^\mu\psi. \tag{7.81}$$

Here "V" stands for vector, since both left- and right-handed components transform identically under this symmetry.

The conservation law $\partial_\mu j_V^\mu = 0$ follows from the equations of motion:

$$\partial_\mu j_V^\mu = (\partial_\mu\bar{\psi})\gamma^\mu\psi + \bar{\psi}\gamma^\mu(\partial_\mu\psi) \tag{7.82}$$

$$= im\bar{\psi}\psi - im\bar{\psi}\psi = 0. \tag{7.83}$$

Here we used the Dirac equations $i\partial\!\!\!/\psi = m\psi$ and $i\partial_\mu\bar{\psi}\gamma^\mu = -m\bar{\psi}$.

The corresponding conserved charge is

$$Q \;=\; \int d^3x\,\bar{\psi}\gamma^0\psi \;=\; \int d^3x\,\psi^\dagger\psi, \tag{7.84}$$

which is interpreted as the electric charge (or fermion number) in quantum field theory.

Axial Symmetry

For $m = 0$, the Dirac Lagrangian has an additional internal (axial) symmetry under which left- and right-handed fermions transform oppositely:

$$\psi \;\to\; e^{i\alpha\gamma^5}\psi, \qquad \bar{\psi} \;\to\; \bar{\psi}e^{i\alpha\gamma^5}.$$

The second transformation follows from the first, since

$$e^{-i\alpha\gamma^5}\gamma^0 = \gamma^0 e^{+i\alpha\gamma^5}.$$

This leads to the conserved current

$$j_A^\mu \;=\; \bar{\psi}\gamma^\mu\gamma^5\psi, \tag{7.85}$$

where A stands for "axial," reflecting that j_A^μ is an axial vector. This current is conserved only in the massless case $m = 0$.

With the full Dirac Lagrangian we can compute

$$\partial_\mu j_A^\mu = (\partial_\mu\bar{\psi})\,\gamma^\mu\gamma^5\psi + \bar{\psi}\gamma^\mu\gamma^5(\partial_\mu\psi) \tag{7.86}$$

$$= 2im\,\bar{\psi}\gamma^5\psi, \tag{7.87}$$

which clearly vanishes only if $m = 0$.

In quantum theory, however, the situation becomes more subtle for the axial current. When the theory is coupled to gauge fields, the axial transformation is still a symmetry of the classical Lagrangian, but it does not survive quantization. This is the archetypal example of an *anomaly*: a symmetry of the classical theory that fails to be preserved in the quantum theory.

7.7 Plane wave solutions

We now study plane–wave solutions of the Dirac equation

$$\big(i\gamma^\mu\partial_\mu - m\big)\psi = 0. \tag{7.88}$$

Take the ansatz

$$\psi(x) = u(\vec{p})\, e^{-ip\cdot x}, \tag{7.89}$$

where $u(\vec{p})$ is a 4–component spinor depending on the 3–momentum $\vec{p}$. Then

$$\big(\gamma^\mu p_\mu - m\big)u(\vec{p}) = \begin{pmatrix} -m & p_\mu\sigma^\mu \\ p_\mu\bar{\sigma}^\mu & -m \end{pmatrix} u(\vec{p}) = 0, \tag{7.90}$$

with

$$\sigma^\mu = (1, \sigma^i), \qquad \bar{\sigma}^\mu = (1, -\sigma^i), \tag{7.91}$$

so that (in this signature) $p_\mu\sigma^\mu = p^0 - \vec{p}\cdot\vec{\sigma}$ and $p_\mu\bar{\sigma}^\mu = p^0 + \vec{p}\cdot\vec{\sigma}$.

The above are called *positive frequency* solutions. There are also negative frequency solutions given by:

$$\psi(x) = v(\vec{p})\, e^{ip\cdot x}, \tag{7.92}$$

§35: Show that: $\big(\gamma^\mu p_\mu + m\big)v(\vec{p}) = 0$. Note, the difference in sign in the m term.

We can now use the previous results to obtain explicit forms of $u(\vec{p},s)$ and $v(\vec{p},s)$ in a chosen representation. We work in the Dirac basis. So, γ^0 here will mean γ^5 of the chiral basis. Since the Dirac equation is Lorentz invariant, we can choose any convenient frame – take the rest frame:

$$\not{p} = (m\gamma^0, 0, 0, 0). \tag{7.93}$$

The Dirac equation becomes

$$0 = m(\gamma^0 - \mathbb{1}_{4\times4})\,u(\vec{p},s) = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_{2\times2} \end{pmatrix} u(\vec{p},s), \tag{7.94}$$

which gives two independent solutions (the two spin values s):

$$u(\vec{p},1) = \sqrt{2m}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad u(\vec{p},2) = \sqrt{2m}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \tag{7.95}$$

Here we normalized the states with $\sqrt{2m}$.

Similarly, for the negative–frequency solutions we obtain

$$v(\vec{p},1) = \sqrt{2m}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad v(\vec{p},2) = \sqrt{2m}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{7.96}$$

The chiral basis/Weyl basis is related to the Dirac basis by the similarity transformation

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbb{1}_{2\times 2} & -\mathbb{1}_{2\times 2} \\ \mathbb{1}_{2\times 2} & \mathbb{1}_{2\times 2} \end{pmatrix}. \tag{7.97}$$

For instance if you compare above solutions to the one present in David Tong's lecture notes, setting $\vec{p} = \vec{0}$ in his solutions, the spinors match via

$$u^{\text{Tong}}(\vec{p},1) = S\,u^{\text{Us}}(\vec{p},1) = \sqrt{m}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad u^{\text{Tong}}(\vec{p},2) = S\,u^{\text{Us}}(\vec{p},2) = \sqrt{m}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}. \tag{7.98}$$

Claim. The normalisation for the u/v spinors is

$$\bar{u}(\vec{p},s)\,u(\vec{p},s) = 2m, \qquad \bar{v}(\vec{p},s)\,v(\vec{p},s) = -2m. \tag{7.99}$$

Proof. Since $u(\vec{p},s)$ transforms in the same way as $\psi(x)$ [52], the bilinears $\bar{u}(\vec{p},s)u(\vec{p},s)$ and $\bar{u}(\vec{p},s)\gamma^\mu p_\mu u(\vec{p},s)$ are Lorentz invariant and hence we can evaluate them in the above rest frame. From (7.90) we have,

$$\bar{u}(\vec{p},s)\not{p}\,u(\vec{p},s) - m\bar{u}(\vec{p},s)u(\vec{p},s) = 0. \tag{7.100}$$

In the rest frame $p = (m,\vec{0})$, this gives

$$m\,\bar{u}(\vec{p},s)u(\vec{p},s) = m\,u^\dagger(\vec{p},s)(\gamma^0)^2u(\vec{p},s) = m\,u^\dagger(\vec{p},s)u(\vec{p},s) = 2m^2, \tag{7.101}$$

so that

$$\bar{u}(\vec{p},s)u(\vec{p},s) = 2m. \tag{7.102}$$

A similar argument gives $\bar{v}(\vec{p},s)v(\vec{p},s) = -2m$. It is also straightforward to show $\bar{v}(\vec{p},s)u(\vec{p},s) = 0$. The general result is therefore

$$\bar{u}(\vec{p},s)u(\vec{p},r) = 2m\,\delta^{rs}, \qquad \bar{v}(\vec{p},s)v(\vec{p},r) = -2m\,\delta^{rs}. \tag{7.103}$$

This normalisation is Lorentz invariant and basis independent, and so holds in general. There are also non-Lorentz invariant expressions – which will be useful later:

$$u^\dagger(\vec{p},r)u(\vec{p},s) = 2p_0\delta^{rs}, \qquad v^\dagger(\vec{p},r)v(\vec{p},s) = 2p_0\delta^{rs}, \qquad u^\dagger(\vec{p},r)v(-\vec{p},s) = v^\dagger(\vec{p},r)u(-\vec{p},s) = 0 \tag{7.104}$$

(7.103)-(7.104) are called spinor *inner product identities*.

Additionally, we have, what are called *spin sums* relations (or spinor completeness relations/outer product formulae):

$$\sum_{s=1}^{2} u(\vec{p},s)\,\bar{u}(\vec{p},s) = \not{p} + m, \qquad \sum_{s=1}^{2} v(\vec{p},s)\,\bar{v}(\vec{p},s) = \not{p} - m. \tag{7.105}$$

[52] Since $e^{-ip\cdot x}$ are Lorentz invariant as they involve an inner product

8 Quantization of Dirac equation

We now wish to quantize the Dirac Lagrangian,

$$\mathcal{L} = \bar{\psi}(x)\big(i\partial\!\!\!/ - m\big)\psi(x). \tag{8.1}$$

We will at first proceed naively and treat ψ just like the scalar field. However, we will soon see that this fails, and we must reconsider how to properly quantize the Dirac theory.

8.1 A Glimpse at the Spin–Statistics Theorem

We begin by defining the canonical momentum:

$$\pi \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0\psi)} = i\bar{\psi}\gamma^0 = i\psi^\dagger. \tag{8.2}$$

For the Dirac Lagrangian, the conjugate momentum is $i\psi^\dagger$. It does not involve a time derivative of ψ, which is natural: the Dirac equation is first-order in time derivatives. Thus, specifying ψ and $\psi^\dagger$ on an initial time slice suffices to determine the full evolution.

To quantize the theory, we promote ψ and $\psi^\dagger$ to operators and impose canonical commutation relations:

$$[\psi_\alpha(\vec{x}),\, \psi_\beta(\vec{y})] = [\psi^\dagger_\alpha(\vec{x}),\, \psi^\dagger_\beta(\vec{y})] = 0, \tag{8.3}$$

$$[\psi_\alpha(\vec{x}),\, \psi^\dagger_\beta(\vec{y})] = \delta_{\alpha\beta}\,\delta^{(3)}(\vec{x}-\vec{y}). \tag{8.4}$$

It is precisely this step that we will soon need to reconsider.

Since this is a free theory, any classical solution is a superposition of plane waves. Thus, the field operators can be expanded as

$$\psi(\vec{x}) = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\Big[\, b^s_{\vec{p}}\, u^s(\vec{p})\, e^{i\vec{p}\cdot\vec{x}} + c^{s\dagger}_{\vec{p}}\, v^s(\vec{p})\, e^{-i\vec{p}\cdot\vec{x}}\,\Big], \tag{8.5}$$

$$\psi^\dagger(\vec{x}) = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\Big[\, b^{s\dagger}_{\vec{p}}\, u^s(\vec{p})\, e^{-i\vec{p}\cdot\vec{x}} + c^s_{\vec{p}}\, v^s(\vec{p})\, e^{i\vec{p}\cdot\vec{x}}\,\Big]. \tag{8.6}$$

Here the operators $b^{s\dagger}_{\vec{p}}$ create particles associated with the spinors $u^s(\vec{p})$, while $c^{s\dagger}_{\vec{p}}$ create particles associated with $v^s(\vec{p})$. As in the scalar case, the commutation relations of the fields imply commutation relations for the creation and annihilation operators.

Claim: The field commutators (8.4) are equivalent to

$$[b^r_{\vec{p}},\, b^{s\dagger}_{\vec{q}}] = (2\pi)^3\delta^{rs}\,\delta^{(3)}(\vec{p}-\vec{q}), \tag{8.7}$$

$$[c^r_{\vec{p}},\, c^{s\dagger}_{\vec{q}}] = -(2\pi)^3\delta^{rs}\,\delta^{(3)}(\vec{p}-\vec{q}), \tag{8.8}$$

with all other commutators vanishing.

Notice the unusual minus sign in the $[c, c^\dagger]$ commutator. This implies that we cannot define the ground state $|0\rangle$ as the state annihilated by $c^r_{\vec{p}}$, since then the excited states $c^{s\dagger}_{\vec{p}}|0\rangle$ would

carry negative norm [53]. To avoid this, one flips the interpretation of c and $c^\dagger$, instead defining the vacuum by

$$c^{s\dagger}_{\vec{p}}|0\rangle = 0,$$

with the excited states created by $c^r_{\vec{p}}|0\rangle$. This, however, will turn out to be the problem.

Proof

Let us verify that the $[b, b^\dagger]$ and $[c, c^\dagger]$ commutators reproduce the field commutator relations (8.4).

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \sum_{r,s} \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{\sqrt{4E_{\vec{p}}E_{\vec{q}}}} \Big([b^r_{\vec{p}}, b^{s\dagger}_{\vec{q}}] u^r(\vec{p}) u^s(\vec{q})^\dagger e^{i(\vec{p}\cdot\vec{x} - \vec{q}\cdot\vec{y})} \tag{8.9}$$

$$+ [c^{r\dagger}_{\vec{p}}, c^s_{\vec{q}}] v^r(\vec{p}) v^s(\vec{q})^\dagger e^{-i(\vec{p}\cdot\vec{x} - \vec{q}\cdot\vec{y})} \Big) \tag{8.10}$$

$$= \sum_s \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \Big(u^s(\vec{p}) u^s(\vec{p})^\dagger e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + v^s(\vec{p}) v^s(\vec{p})^\dagger e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \Big). \tag{8.11}$$

At this point we employ the spinor completeness relations:

$$\sum_s u^s(\vec{p}) \bar{u}^s(\vec{p}) = \not{p} + m, \qquad \sum_s v^s(\vec{p}) \bar{v}^s(\vec{p}) = \not{p} - m,$$

so that

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \Big((\not{p} + m)\gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + (\not{p} - m)\gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \Big). \tag{8.12}$$

Rearranging,

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \Big((p_0\gamma^0 + p_i\gamma^i + m)\gamma^0 + (p_0\gamma^0 - p_i\gamma^i - m)\gamma^0 \Big) e^{i\vec{p}\cdot(\vec{x}-\vec{y})}. \tag{8.13}$$

Now, using $p_0 = E_{\vec{p}}$, the spatial γ^i terms cancel, leaving

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}-\vec{y})} = \delta^{(3)}(\vec{x} - \vec{y}), \tag{8.14}$$

as required.

The key step was the cancellation of the $p_i\gamma^i$ terms, which depends crucially on the minus sign in the $[c, c^\dagger]$ commutator of (8.8).

[53] To see this note that in SHO, we had $[a, a^\dagger] = +1$ and heence if we denote $|1\rangle = a^\dagger|0\rangle$ then $\langle 1|1\rangle = \langle 0|aa^\dagger|0\rangle = \langle 0|0\rangle \geq 0$. However, if we had instead: $[a, a^\dagger] = -1$ then the above argument shows that $\langle 1|1\rangle = -\langle 0|0\rangle \leq 0$. That is if the vacuum has positive norm then excited states have negative norm!

8.2 The Hamiltonian

To construct the Hamiltonian for the theory, recall that the canonical momentum is $\pi = i\psi^\dagger$. We then have

$$\mathcal{H} \;=\; \pi\psi - \mathcal{L} = \bar{\psi}(-i\gamma^i\partial_i + m)\psi. \tag{8.15}$$

Thus the Hamiltonian

$$H = \int d^3x\, \mathcal{H}$$

is consistent with the conserved energy computed via Noether's theorem. We now proceed to turn the Hamiltonian into an operator. First consider

$$(-i\gamma^i\partial_i + m)\psi = \int \frac{d^3p}{(2\pi)^3}\,\frac{1}{\sqrt{2E_{\vec{p}}}}\Big[b^s_{\vec{p}}(-\gamma^i p_i + m)u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}} + c^{s\dagger}_{\vec{p}}(\gamma^i p_i + m)v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\Big]. \tag{8.16}$$

Here we have left the sum over $s = 1, 2$ implicit. Note that the minus signs arise from the Minkowski metric conventions, since $\vec{p}\cdot\vec{x} \equiv \sum_i x^i p^i = -x^i p_i$.

Now use the spinor equations (see (7.90) and the exercies therein for $v(\vec{p})$'s EoM):

$$(-\gamma^i p_i + m)u^s(\vec{p}) = \gamma^0 p_0 u^s(\vec{p}), \qquad (\gamma^i p_i + m)v^s(\vec{p}) = -\gamma^0 p_0 v^s(\vec{p}),$$

to rewrite

$$(-i\gamma^i\partial_i + m)\psi = \int \frac{d^3p}{(2\pi)^3}\,\frac{E_{\vec{p}}}{2}\,\gamma^0\Big[b^s_{\vec{p}}u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}} - c^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\Big]. \tag{8.17}$$

Substituting into the Hamiltonian gives

$$H = \int d^3x\, \psi^\dagger\gamma^0(-i\gamma^i\partial_i + m)\psi \tag{8.18}$$

$$= \int \frac{d^3x\, d^3p\, d^3q}{(2\pi)^6}\sqrt{\frac{E_{\vec{p}}}{4E_{\vec{q}}}}\Big[b^{r\dagger}_{\vec{q}}u^r(\vec{q})^\dagger e^{-i\vec{q}\cdot\vec{x}} + c^r_{\vec{q}}v^r(\vec{q})^\dagger e^{i\vec{q}\cdot\vec{x}}\Big] \times\Big[b^s_{\vec{p}}u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}} - c^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\Big]. \tag{8.19}$$

Carrying out the d^3x integration yields

$$H = \int \frac{d^3p}{(2\pi)^3}\,\frac{1}{2}\Big[b^{r\dagger}_{\vec{p}}b^s_{\vec{p}}\big(u^r(\vec{p})^\dagger u^s(\vec{p})\big) - c^r_{\vec{p}}c^{s\dagger}_{\vec{p}}\big(v^r(\vec{p})^\dagger v^s(\vec{p})\big) - b^{r\dagger}_{\vec{p}}c^{s\dagger}_{-\vec{p}}\big(u^r(\vec{p})^\dagger v^s(-\vec{p})\big) + c^r_{\vec{p}}b^s_{-\vec{p}}\big(v^r(\vec{p})^\dagger u^s(-\vec{p})\big)\Big]. \tag{8.20}$$

In the last two terms we relabelled $\vec{p} \to -\vec{p}$. Using the inner product identities

$$u^r(\vec{p})^\dagger u^s(\vec{p}) = v^r(\vec{p})^\dagger v^s(\vec{p}) = 2p_0\delta^{rs}, \qquad u^r(\vec{p})^\dagger v^s(-\vec{p}) = v^r(\vec{p})^\dagger u^s(-\vec{p}) = 0,$$

we will simplify the Hamiltonian further to get,

$$H = \int \frac{d^3p}{(2\pi)^3}\,E_{\vec{p}}\left(b^{s\dagger}_{\vec{p}}b^s_{\vec{p}} - c^s_{\vec{p}}c^{s\dagger}_{\vec{p}}\right), \tag{8.21}$$

$$= \int \frac{d^3p}{(2\pi)^3}\,E_{\vec{p}}\left(b^{s\dagger}_{\vec{p}}b^s_{\vec{p}} - c^{s\dagger}_{\vec{p}}c^s_{\vec{p}} + (2\pi)^3\delta^{(3)}(0)\right). \tag{8.22}$$

The $\delta^{(3)}(0)$ term is familiar and removed by normal ordering. The $b^\dagger b$ term is also standard: one verifies that $b^\dagger$ creates positive energy states,

$$[H,\, b^{s\dagger}_{\vec{p}}\,] = E_{\vec{p}}\, b^{s\dagger}_{\vec{p}}. \tag{8.23}$$

The problem lies in the minus sign in front of the $c^\dagger c$ term. If we interpret $c^\dagger$ as a creation operator, the commutation relations (8.8) still imply that $c^\dagger$ creates positive energy states,

$$[H,\, c^{s\dagger}_{\vec{p}}\,] = E_{\vec{p}}\, c^{s\dagger}_{\vec{p}}. \tag{8.24}$$

However, as already noted after (8.8), such states carry negative norm. To ensure a sensible Hilbert space we must reinterpret c (not $c^\dagger$) as the creation operator. But then the Hamiltonian is no longer bounded below, since

$$[H,\, c^{s}_{\vec{p}}\,] = -E_{\vec{p}}\, c^{s}_{\vec{p}}. \tag{8.25}$$

This is a disaster: taken literally, it implies that the system can cascade down to ever lower energies by producing more and more c-particles. In other words, the theory is unstable if we maintain commutators for fermionic fields.

8.3 Fermionic Quantization

The crucial piece of physics we missed is that spin-$\frac{1}{2}$ particles are fermions, and hence they obey Fermi–Dirac statistics: the quantum state acquires a minus sign upon the interchange of any two particles. This principle is embedded in relativistic quantum field theory via the spin–statistics theorem: integer-spin fields must be quantized as bosons, while half-integer spin fields must be quantized as fermions. Any attempt to do otherwise leads to inconsistencies, such as the unbounded Hamiltonian found in (8.22).

How then do we quantize a field as a fermion? Recall that for the scalar field, the resulting quanta obeyed bosonic statistics because the creation and annihilation operators satisfied

$$[a_{\vec{p}}, a^\dagger_{\vec{q}}] = 0 \quad \Rightarrow \quad a^\dagger_{\vec{p}} a^\dagger_{\vec{q}} |0\rangle = |\vec{p},\vec{q}\rangle = |\vec{q},\vec{p}\rangle. \tag{8.26}$$

To obtain states obeying fermionic statistics, we instead impose *anticommutation* relations:

$$\{A, B\} \equiv AB + BA.$$

Rather than the commutators in (8.4), we now require that the spinor fields satisfy

$$\{\psi_\alpha(\vec{x}),\, \psi_\beta(\vec{y})\} = \{\psi^\dagger_\alpha(\vec{x}),\, \psi^\dagger_\beta(\vec{y})\} = 0, \qquad \{\psi_\alpha(\vec{x}),\, \psi^\dagger_\beta(\vec{y})\} = \delta_{\alpha\beta}\, \delta^{(3)}(\vec{x}-\vec{y}). \tag{8.27}$$

We still use the same expansion (8.6) of ψ and $\psi^\dagger$ in terms of $b, b^\dagger, c, c^\dagger$. But now the same proof that previously gave (8.8) yields

$$\{b^r_{\vec{p}}, b^{s\dagger}_{\vec{q}}\} = (2\pi)^3 \delta^{rs}\, \delta^{(3)}(\vec{p}-\vec{q}), \tag{8.28}$$

$$\{c^r_{\vec{p}}, c^{s\dagger}_{\vec{q}}\} = (2\pi)^3 \delta^{rs}\, \delta^{(3)}(\vec{p}-\vec{q}), \tag{8.29}$$

with all other *anticommutators* vanishing:

$$\{b^r_{\vec{p}}, b^s_{\vec{q}}\} = \{c^r_{\vec{p}}, c^s_{\vec{q}}\} = \{b^r_{\vec{p}}, c^s_{\vec{q}}\} = \{b^r_{\vec{p}}, c^{s\dagger}_{\vec{q}}\} = \cdots = 0. \tag{8.30}$$

The calculation of the Hamiltonian proceeds as before, up to the penultimate line of (8.21). At that stage, we obtain

$$H = \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} \left(b^{s\dagger}_{\vec{p}} b^s_{\vec{p}} - c^s_{\vec{p}} c^{s\dagger}_{\vec{p}} \right) \tag{8.31}$$

$$= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} \left(b^{s\dagger}_{\vec{p}} b^s_{\vec{p}} + c^{s\dagger}_{\vec{p}} c^s_{\vec{p}} - (2\pi)^3 \delta^{(3)}(0) \right). \tag{8.32}$$

The use of anticommutators has rescued us from the disaster of an unbounded Hamiltonian. Note, however, that when normal ordering we now discard a *negative* vacuum contribution, $-(2\pi)^3\delta^{(3)}(0)$.

8.3.1 Fermi–Dirac Statistics

As in the bosonic case, we define the vacuum $|0\rangle$ by

$$b^s_{\vec{p}}|0\rangle = c^s_{\vec{p}}|0\rangle = 0. \tag{8.33}$$

Although b and c satisfy anticommutation relations, the Hamiltonian (8.32) has neat commutators with them. Specifically,

$$[H,\, b^r_{\vec{p}}] = -E_{\vec{p}}\, b^r_{\vec{p}}, \qquad [H,\, b^{r\dagger}_{\vec{p}}] = E_{\vec{p}}\, b^{r\dagger}_{\vec{p}}, \tag{8.34}$$

$$[H,\, c^r_{\vec{p}}] = -E_{\vec{p}}\, c^r_{\vec{p}}, \qquad [H,\, c^{r\dagger}_{\vec{p}}] = E_{\vec{p}}\, c^{r\dagger}_{\vec{p}}. \tag{8.35}$$

Thus we can once again build a tower of positive-energy eigenstates by acting on the vacuum with $b^{r\dagger}_{\vec{p}}$ and $c^{r\dagger}_{\vec{p}}$, just as in the bosonic case. For instance, the one-particle states are

$$|\vec{p}, r\rangle = b^{r\dagger}_{\vec{p}}|0\rangle. \tag{8.36}$$

Two-particle states, however, now satisfy

$$|\vec{p}_1, r_1; \vec{p}_2, r_2\rangle = b^{r_1\dagger}_{\vec{p}_1} b^{r_2\dagger}_{\vec{p}_2}|0\rangle = -\, b^{r_2\dagger}_{\vec{p}_2} b^{r_1\dagger}_{\vec{p}_1}|0\rangle = -|\vec{p}_2, r_2; \vec{p}_1, r_1\rangle. \tag{8.37}$$

This confirms that the quanta obey Fermi–Dirac statistics. In particular, the Pauli exclusion principle follows:

$$|\vec{p}, r; \vec{p}, r\rangle = 0.$$

Finally, to verify that these excitations indeed carry spin-$\frac{1}{2}$, we can act with the angular momentum operator (see (7.80)). For example, a stationary particle $|\vec{p} = 0, r\rangle$ carries intrinsic angular momentum $\frac{1}{2}$ as expected.

8.4 Propagators

We now move to the Heisenberg picture. The spinors $\psi(\vec{x},t)$ are defined at every point in spacetime such that they satisfy the operator equation

$$\frac{\partial\psi}{\partial t} = i[H,\psi]. \tag{8.38}$$

The solution is given by the mode expansion

$$\psi(x) = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\Big[b^s_{\vec{p}}u^s(\vec{p})e^{-ip\cdot x} + c^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{+ip\cdot x}\Big], \tag{8.39}$$

$$\psi^\dagger(x) = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\Big[b^{s\dagger}_{\vec{p}}u^s(\vec{p})^\dagger e^{+ip\cdot x} + c^{s}_{\vec{p}}v^s(\vec{p})^\dagger e^{-ip\cdot x}\Big]. \tag{8.40}$$

Next, let us consider the anticommutators of these fields. We define the fermionic propagator as

$$iS_{\alpha\beta}(x-y) = \{\psi_\alpha(x),\bar{\psi}_\beta(y)\}. \tag{8.41}$$

In what follows we will often suppress indices and simply write $iS(x-y)=\{\psi(x),\bar{\psi}(y)\}$, but one should keep in mind that $S(x-y)$ is a 4×4 matrix.

Using the expansion (8.40), we find

$$\begin{aligned} iS(x-y) &= \int \frac{d^3p\,d^3q}{(2\pi)^6}\frac{1}{\sqrt{4E_{\vec{p}}E_{\vec{q}}}}\Big\{[b^s_{\vec{p}},b^{r\dagger}_{\vec{q}}]u^s(\vec{p})\bar{u}^r(\vec{q})\,e^{-i(p\cdot x-q\cdot y)} \\ &\qquad + [c^s_{\vec{p}},c^{r\dagger}_{\vec{q}}]v^s(\vec{p})\bar{v}^r(\vec{q})\,e^{+i(p\cdot x-q\cdot y)}\Big\} \qquad (8.42)\\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\Big[u^s(\vec{p})\bar{u}^s(\vec{p})\,e^{-ip\cdot(x-y)} + v^s(\vec{p})\bar{v}^s(\vec{p})\,e^{+ip\cdot(x-y)}\Big] \qquad (8.43)\\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\Big[(\not{p}+m)e^{-ip\cdot(x-y)} + (\not{p}-m)e^{+ip\cdot(x-y)}\Big]. \qquad (8.44) \end{aligned}$$

where to get to the last line we have used the outer product formulae/spin sums. The final line of the previous calculation can be written as

$$iS(x-y) = (i\not{\partial}_x + m)\big(\Delta^+(x-y)-\Delta^-(x-y)\big), \tag{8.45}$$

in terms of the scalar propagator/Wightman functions $\Delta^+(x-y)$ [54], which for a real scalar field is

$$\Delta^+(x-y) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\,e^{-ip\cdot(x-y)}, \tag{8.46}$$

and, $\Delta^-(x-y)=\Delta^+(y-x)$.

[54] Here, we are assuming the state where they are evaluated on, say the vacuum is translational invariant and hence $\Delta^+(x,y)=\Delta^+(x-y)$.

Comments

- For spacelike separation $(x-y)^2 < 0$, we know $D(x-y)-D(y-x)=0$. In the bosonic theory, this was crucial because it implied

$$[\phi(x),\phi(y)]=0, \qquad (x-y)^2<0, \tag{8.47}$$

which established causality.

For fermions, however, we find

$$\{\psi_\alpha(x),\psi_\beta(y)\}=0, \qquad (x-y)^2<0, \tag{8.48}$$

outside the light cone.

At first this may appear troubling, but note: all observables are bilinear in fermions, such as the Hamiltonian (8.32). These bilinears commute at spacelike separation, so the theory remains causal.

Furthermore, note that since the spinor field itself is not a Lorentz scalar but only a bilinear $\bar\psi\psi$, we can rephrase the causality requirement in this case to be:

$$\big[\,\bar\psi(x)\psi(x),\,\bar\psi(y)\psi(y)\,\big] = 0, \qquad \forall (x-y)^2<0. \tag{8.49}$$

So, if we can show that the following holds then from (8.48) the above would vanish.

$$\big[\,\bar\psi(x)\psi(x),\,\bar\psi(y)\psi(y)\,\big] \sim \{\psi_\alpha(x),\,\bar\psi_\beta(y)\}. \tag{8.50}$$

§36: Show that: $[\bar\psi_\alpha\psi_\alpha,\bar\psi_\beta,\psi_\beta]=\bar\psi_\alpha\{\psi_\alpha,\bar\psi_\beta\}\psi_\beta-\bar\psi_\beta\{\psi_\beta,\bar\psi_\alpha\}\psi_\alpha$.

Hint: Write $\bar\psi\psi=\bar\psi_\alpha\psi_\alpha$ and use the following identities: $[A,B]=\{A,B\}-2BA$, $[AB,C]=A[B,C]+[A,C]B$ and $[A,CD]=C[A,D]+[A,C]D$.

- Away from singularities, the propagator satisfies

$$(i\partial\!\!\!/_x - m)S(x-y)=0, \tag{8.51}$$

which follows from the fact that $(\partial_x^2+m^2)D(x-y)=0$ on the mass shell $p^2=m^2$.

8.4.1 The Feynman Propagator

By a similar calculation to the above, we may evaluate the vacuum expectation values:

$$\langle 0|\psi_\alpha(x)\bar\psi_\beta(y)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec p}}(p\!\!\!/+m)_{\alpha\beta}\,e^{-ip\cdot(x-y)}, \tag{8.52}$$

$$\langle 0|\bar\psi_\beta(y)\psi_\alpha(x)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec p}}(p\!\!\!/-m)_{\alpha\beta}\,e^{+ip\cdot(x-y)}. \tag{8.53}$$

We now define the *Feynman propagator* $S_F(x-y)$, which is again a 4×4 matrix, as the time-ordered product:

$$S_F(x-y) \;=\; \langle 0|T\{\psi(x)\,\bar{\psi}(y)\}|0\rangle \;\equiv\; \begin{cases} \langle 0|\psi(x)\,\bar{\psi}(y)|0\rangle, & x^0 > y^0, \\ \langle 0|-\bar{\psi}(y)\,\psi(x)|0\rangle, & y^0 > x^0. \end{cases} \tag{8.54}$$

The minus sign here is essential for Lorentz invariance. When $(x-y)^2 < 0$, there is no invariant way to decide whether $x^0 > y^0$ or $y^0 > x^0$. The minus sign ensures consistency, since $\{\psi(x), \bar{\psi}(y)\} = 0$ outside the light cone.

The Feynman propagator admits the momentum-space representation

$$S_F(x-y) = i\int \frac{d^4p}{(2\pi)^4}\,\frac{\not{p}+m}{p^2-m^2+i\epsilon}\,e^{-ip\cdot(x-y)}. \tag{8.55}$$

Finally, the propagator satisfies

$$(i\not{\partial}_x - m)S_F(x-y) = i\,\delta^{(4)}(x-y), \tag{8.56}$$

so that S_F is indeed a Green's function for the Dirac operator.

The minus sign that appears in (8.54) also occurs for any string of operators inside a time-ordered product $T(\cdots)$. While bosonic operators commute inside T, fermionic operators *anti-commute*. The same rule applies for normal-ordered products as well, with fermionic operators obeying

$$:\psi_1\psi_2: \;=\; -:\psi_2\psi_1:.$$

With the understanding that all fermionic operators anti-commute inside both T.

Two-operator time ordering.

Let $A(x), B(y)$ be *bosonic* (even) operators and $F(x), G(y)$ be *fermionic* (odd) operators. The time–ordered products are

$$T(A(x)B(y)) = \theta(x^0-y^0)A(x)B(y) + \theta(y^0-x^0)\,B(y)A(x), \tag{8.57}$$

$$T(F(x)G(y)) = \theta(x^0-y^0)\,F(x)G(y) - \theta(y^0-x^0)\,G(y)F(x). \tag{8.58}$$

The minus sign in (8.58) appears because exchanging two odd operators inside $T(\cdot)$ introduces a factor -1.

A Translational Symmetry and Fourier Transform

Consider a field $\psi(x)$ in one spatial dimension. The translation operator $T(a)$ acts as

$$T(a)\,\psi(x) \;=\; \psi(x+a).$$

If the system is invariant under translations, then $T(a)$ commutes with the Hamiltonian. Thus we can choose eigenfunctions of $T(a)$. Let us look for solutions of the form

$$T(a)\,\psi_k(x) = \lambda(a)\,\psi_k(x).$$

This means

$$\psi_k(x+a) = \lambda(a)\,\psi_k(x).$$

Eigenvalue structure

Requiring consistency for two successive translations gives

$$\psi_k(x+a+b) = \lambda(a+b)\,\psi_k(x) = \lambda(a)\,\lambda(b)\,\psi_k(x).$$

Therefore

$$\lambda(a+b) = \lambda(a)\,\lambda(b).$$

The only continuous solutions to this functional equation are exponentials:

$$\lambda(a) = e^{ika}.$$

Hence the eigenfunctions of translations are plane waves

$$\psi_k(x) = e^{ikx}.$$

Fourier decomposition

Since $\{e^{ikx}\}$ form a complete orthogonal basis on $\mathbb{R}$, any field $\psi(x)$ can be expanded as

$$\psi(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi}\,\tilde{\psi}(k)\,e^{ikx},$$

with inverse transform

$$\tilde{\psi}(k) = \int_{-\infty}^{\infty} dx\,e^{-ikx}\,\psi(x).$$

where, we have chosen a specific convention for putting all the 2π factors in the inverse/momentum space.

Physical meaning

- Translation invariance $\Rightarrow$ momentum conservation, as we have seen earlier.
- The plane waves e^{ikx} are simultaneously eigenfunctions of $T(a)$ and the momentum operator $\hat{p} = -i\hbar\,\frac{d}{dx}$, with eigenvalue $\hbar k$.
- Operators that commute with translations (e.g. free-particle Hamiltonians) become diagonal in momentum space.

Thus, continuous translational symmetry naturally leads to the Fourier transform: plane waves are the irreducible building blocks of translation-invariant systems.

B Very brief complex analysis

A complex function $f(z)$ is *analytic* on a region if it has a complex derivative at every point of that region (equivalently: it admits a convergent power series there). A *contour* Γ is a piecewise smooth closed loop. If $f(z)$ is analytic on and inside Γ, the contour encloses no singularities of $f(z)$.

B.1 Cauchy's three theorems

1) Cauchy's Integral Theorem If $f(z)$ is analytic on and inside a closed contour Γ, then

$$\oint_\Gamma f(z)\,dz \;=\; 0. \tag{B.1}$$

Takeaway: you can continuously deform Γ within the analytic region without changing the integral.

2) Cauchy's Integral Formula If $f(z)$ is analytic near $z=a$ and Γ winds once counterclockwise around $z=a$, then

$$f(a)=\frac{1}{2\pi i}\oint_\Gamma \frac{f(z)}{z-a}\,dz, \tag{B.2}$$

$$f^{(n)}(a)=\frac{n!}{2\pi i}\oint_\Gamma \frac{f(z)}{(z-a)^{n+1}}\,dz. \tag{B.3}$$

Takeaway: This formula relates the value of an analytic function $f(z)$ at a point $z=a$ inside a contour Γ to an integral of $f(z)$ over that contour. Hence, can be used to extract values/derivatives of $f(z)$ from its integral over the contour.

3) Cauchy's residue theorem If $f(z)$ is analytic on and inside Γ except for isolated poles z_k inside Γ, then

$$\oint_\Gamma f(z)\,dz \;=\; 2\pi i\sum_k \mathrm{Res}\big(f;z_k\big), \tag{B.4}$$

with $\mathrm{Res}(f;z_0)=\lim_{z\to z_0}(z-z_0)f(z)$ for a simple pole.

For a function $f(z)$, a simple zero at $z=z_0$ means $f(z_0)=0$ but $f'(z_0)\neq 0$. So, $f(z)\sim c_1(z-z_0)$ near $z=z_0$. A simple pole at $z=z_1$ means $\frac{1}{f(z_1)}=0$ but $\frac{1}{f'(z_1)}\neq 0$. So, $f(z)\sim\frac{c_2}{(z-z_1)}$ near $z=z_1$.

B.2 Rules for real integrals via residues

- **Close a contour** in the upper or lower half-plane to capture the poles that make the exponential decay.
- **Arc contribution** $\to 0$ when the integrand decays on the large semicircle.

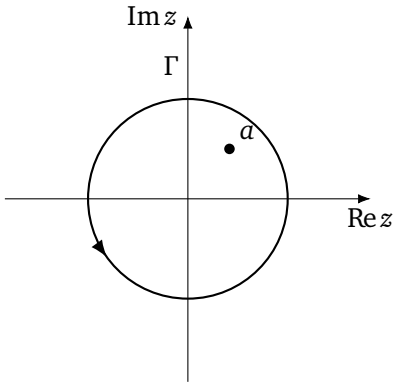


Figure 4: A simple closed contour Γ winding once counterclockwise around a.

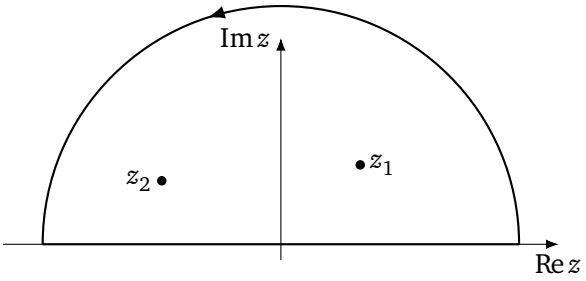


Figure 5: Closing in the upper half-plane to pick up poles z_1, z_2 with $\operatorname{Im} z > 0$.

- Use the residue theorem and take the radius $\to \infty$ to relate the desired real integral to $2\pi i$ times the sum of residues.

As a simple example,

$$\oint_\Gamma dz\, \frac{1}{z} = 2\pi i, \tag{B.5}$$

since the residue is 1. If we reverse the contour orientation (clockwise), the result picks up a minus sign.

B.3 Branch points and branch cuts (logs and square roots)

Some functions are multi-valued under a 2π turn around a point. That point is a *branch point*. To make a single-valued branch, remove a curve (a *branch cut*) so your contour never crosses it.

Logarithm. $\log z = \ln|z| + i \arg z$. A full turn changes $\arg z$ by 2π, so $\log z$ changes by $2\pi i$.

Square root. $\sqrt{z^2+m^2} = \exp\!\left(\tfrac{1}{2}\log(z^2+m^2)\right)$. The branch point is at $z = \pm im$ (and the branch cuts extend till $\pm\infty$ respectively). Crossing the cut flips the sign.

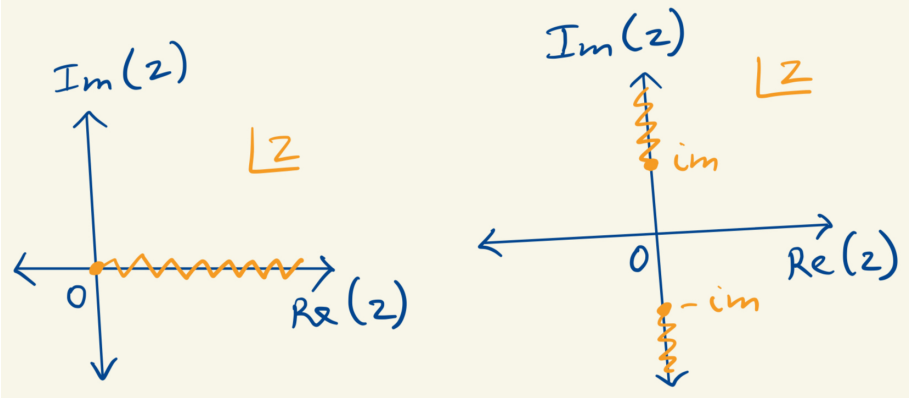


Figure 6: On the left figure we have the branch cut for $\log z$ with branch point at $z = 0$ and on the right side we have the disconnected branch cuts for $\sqrt{z^2 + m^2}$ with branch points at $z = \pm im$.

C Fun with delta distribution identities

C.1 A nice delta identity

Let us discuss a distribution identity which we will use to write (4.46) using (4.45). Let us see below that for any sufficiently regular function $F(\vec{k}, \vec{k}')$,

$$\delta^{(3)}(\vec{k} - \vec{k}')\, F(\vec{k}, \vec{k}') = \delta^{(3)}(\vec{k} - \vec{k}')\, F(\vec{k}, \vec{k})$$

as distributions. To see this note that two distributions T_1 and T_2 are equal iff they act identically on every smooth test function $\varphi(\vec{k}, \vec{k}')$ [55]:

$$\langle T_1, \varphi \rangle = \langle T_2, \varphi \rangle. \tag{C.1}$$

To prove (C.1) note that its LHS is:

$$\begin{aligned} \left\langle \delta^{(3)}(\vec{k} - \vec{k}')\, F(\vec{k}, \vec{k}'),\, \varphi \right\rangle &= \int \mathrm{d}^3k\, \mathrm{d}^3k'\, \delta^{(3)}(\vec{k} - \vec{k}')\, F(\vec{k}, \vec{k}')\, \varphi(\vec{k}, \vec{k}') \\ &= \int \mathrm{d}^3k \left[\int \mathrm{d}^3k'\, \delta^{(3)}(\vec{k} - \vec{k}') \big(F\varphi\big)(\vec{k}, \vec{k}') \right]. \end{aligned}$$

Set $\vec{u} = \vec{k}' - \vec{k}$ so that $\delta^{(3)}(\vec{k} - \vec{k}') = \delta^{(3)}(-\vec{u}) = \delta^{(3)}(\vec{u})$ [56] and $\mathrm{d}^3k' = \mathrm{d}^3u$. Then

$$\begin{aligned} \text{LHS} &= \int \mathrm{d}^3k \left[\int \mathrm{d}^3u\, \delta^{(3)}(\vec{u}) \big(F\varphi\big)(\vec{k}, \vec{k} + \vec{u}) \right] \\ &= \int \mathrm{d}^3k\, F(\vec{k}, \vec{k})\, \varphi(\vec{k}, \vec{k}). \end{aligned}$$

[55] where the below notation means integration as worked out below

[56] by evenness of the delta function

Now RHS of (C.1) is,

$$\begin{aligned}\left\langle \delta^{(3)}(\vec{k}-\vec{k}')\,F(\vec{k},\vec{k}),\,\varphi\right\rangle &= \int \mathrm{d}^3k\,\mathrm{d}^3k'\;\delta^{(3)}(\vec{k}-\vec{k}')\,F(\vec{k},\vec{k})\,\varphi(\vec{k},\vec{k}')\\ &= \int \mathrm{d}^3k\;F(\vec{k},\vec{k})\,\varphi(\vec{k},\vec{k}),\end{aligned}$$

using the same change of variables.

Since both sides act identically on every φ, the distributions are equal:

$$\delta^{(3)}(\vec{k}-\vec{k}')\,F(\vec{k},\vec{k}') = \delta^{(3)}(\vec{k}-\vec{k}')\,F(\vec{k},\vec{k}).$$

Now we can derive the mode-commutator relations by using:

- With $F = e^{it(\omega_{\vec{k}}-\omega_{\vec{k}'})}$,
$$\delta^{(3)}(\vec{k}-\vec{k}')\,e^{it(\omega_{\vec{k}}-\omega_{\vec{k}'})} = \delta^{(3)}(\vec{k}-\vec{k}'),$$
because on the support of δ one has $\omega_{\vec{k}} = \omega_{\vec{k}'}$.
- With $F = \sqrt{\omega_{\vec{k}}/\omega_{\vec{k}'}} + \sqrt{\omega_{\vec{k}'}/\omega_{\vec{k}}}$,
$$\delta^{(3)}(\vec{k}-\vec{k}')\Big(\sqrt{\tfrac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}} + \sqrt{\tfrac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}}\Big) = 2\,\delta^{(3)}(\vec{k}-\vec{k}').$$

Regularity note: it suffices that F be continuous on the diagonal $\{(\vec{k},\vec{k}') : \vec{k}' = \vec{k}\}$, since the delta only probes that set.

C.2 Derivative of the Heaviside step function

Let us show that the derivative of the Heaviside step $\theta(x)$ is the Dirac delta: $\theta'(x) = \delta(x)$ (in the sense of distributions). For any smooth test function φ and using the definition of the delta distribution, we get,

$$\langle \theta', \varphi\rangle = -\langle \theta, \varphi'\rangle = -\int_{-\infty}^{\infty} \theta(x)\,\varphi'(x)\,dx = -\int_{0}^{\infty} \varphi'(x)\,dx = \varphi(0) = \langle \delta, \varphi\rangle,$$

hence $\theta' = \delta$.

C.3 Another delta identity

Start from the fact that the Fourier transform of identity is the delta distribution,

$$(2\pi)^3\delta^{(3)}(\vec{q}-\vec{k}) = \int d^3x\, e^{i(\vec{q}-\vec{k})\cdot\vec{x}} \tag{C.2}$$

and differentiating w.r.t. q_j we get:

$$(2\pi)^3\frac{\partial}{\partial q_j}\delta^{(3)}(\vec{q}-\vec{k}) = \int d^3x\, i x^j e^{i(\vec{q}-\vec{k})\cdot\vec{x}}, \tag{C.3}$$

which is essentially the same identity as in (4.77).

C.4 Delta function pole identity

Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function and suppose its real zeros $\{x_i\}$ are *simple*: $f(x_i) = 0$ and $f'(x_i) \neq 0$. Now let us consider the following delta representation:

$$\delta(y) = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \frac{\varepsilon}{y^2 + \varepsilon^2}. \tag{C.4}$$

The above works since,

$$\lim_{\varepsilon \to 0^+} \int_{-\infty}^{\infty} dy\, f(y) \frac{1}{\pi} \frac{\varepsilon}{y^2 + \varepsilon^2} \overbrace{=}^{(y=\varepsilon t)} \frac{1}{\pi} \lim_{\varepsilon \to 0^+} \int_{-\infty}^{\infty} dt\, \frac{f(\varepsilon t)}{t^2 + 1} \overbrace{=}^{(\varepsilon \to 0^+)} \frac{f(0)}{\pi} \int_{-\infty}^{\infty} \frac{dt}{t^2 + 1} = f(0), \tag{C.5}$$

and so it satisfies the definition of a delta function/distribution.

Now for any smooth test function $\varphi(x)$,

$$\int_{-\infty}^{\infty} \delta\big(f(x)\big)\, \varphi(x)\, dx = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\varepsilon}{f(x)^2 + \varepsilon^2}\, \varphi(x)\, dx. \tag{C.6}$$

Split the integral into small disjoint neighborhoods I_i around each simple zero x_i (away from these points the integrand vanishes in the limit, as $f(x) \neq 0$ there):

$$\int \frac{\varepsilon}{f(x)^2 + \varepsilon^2}\, \varphi(x)\, dx = \sum_i \int_{I_i} \frac{\varepsilon}{f(x)^2 + \varepsilon^2}\, \varphi(x)\, dx. \tag{C.7}$$

On each I_i the map $u = f(x)$ is monotonic, so we may change variables: $u = f(x)$, $dx = du/|f'(x)|$ (we only need $|f'|$ at x_i in the limit). Then

$$\int_{I_i} \frac{\varepsilon}{f(x)^2 + \varepsilon^2}\, \varphi(x)\, dx = \int_{-\delta}^{\delta} \frac{\varepsilon}{u^2 + \varepsilon^2}\, \varphi\big(x(u)\big)\, \frac{du}{|f'(x(u))|}. \tag{C.8}$$

As $u \to 0$, $x(u) \to x_i$ and $f'(x(u)) \to f'(x_i)$, so by continuity we may replace $\varphi(x(u)) \to \varphi(x_i)$ and $|f'(x(u))| \to |f'(x_i)|$ under the narrow peak:

$$\int_{I_i} \frac{\varepsilon}{f(x)^2 + \varepsilon^2}\, \varphi(x)\, dx \xrightarrow[\varepsilon \to 0^+]{} \frac{\varphi(x_i)}{|f'(x_i)|} \int_{-\infty}^{\infty} \frac{\varepsilon}{u^2 + \varepsilon^2}\, du = \pi\, \frac{\varphi(x_i)}{|f'(x_i)|}. \tag{C.9}$$

Plugging into (C.6) gives

$$\int \delta\big(f(x)\big)\, \varphi(x)\, dx = \sum_i \frac{\varphi(x_i)}{|f'(x_i)|}. \tag{C.10}$$

Since this holds for all φ, we identify the distributions:

$$\delta\big(f(x)\big) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|}. \tag{C.11}$$

D More on representations

Here we will look at some more of Lie algebra representations. Given a group G its Lie algebra will be denoted by $\mathfrak{g}$.

D.1 $G = SU(2)$

$SU(2)$ is a group of 2×2 unitary complex matrices with unit determinant. Hence, if $A \in SU(2)$ then $A^\dagger A = \mathbb{1} = AA^\dagger$. Its Lie algebra will be denoted as: $\mathfrak{su}(2)$.

Let us look at close to the identity: $A = e^{i\epsilon J_a}$ where $\epsilon \in \mathbb{R}$ is an infinitesimal parameter and J_a are the generators of $\mathfrak{su}(N)$ [57]. Let us derive the properties they satisfy. Note,

$$e^{i\epsilon J_a} = \mathbb{1} + i\epsilon J_a + \dots \tag{D.1}$$

$$e^{-i\epsilon J_a^\dagger} = \mathbb{1} - i\epsilon J_a^\dagger \tag{D.2}$$

Using $A^\dagger A = \mathbb{1}$, we get, at $\mathcal{O}(\epsilon)$:

$$\begin{aligned} &\mathbb{1} + i\epsilon(J_a - J_a^\dagger) = \mathbb{1} \\ \Rightarrow\ & J_a = J_a^\dagger \end{aligned} \tag{D.3}$$

which implies J_a are $N\times N$ complex Hermitian matrices. Since $\det(A) = 1$, we get $\det(e^{i\epsilon J_a}) = e^{\mathrm{tr}(i\epsilon J_a)} = 1$, implying that $\mathrm{tr}(J_a) = 0$ [58]. Thus J_as are also traceless.

Now from the hermiticity condition (D.3), we get,

$$J_a = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} a_{11}^* & a_{21}^* & \dots \\ a_{12}^* & a_{22}^* & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} = J_a^\dagger$$

which implies that, $a_{ii} = a_{ii}^*$ implying the diagonal components are real and hence there are N independent components here. We also get from above: $a_{ji}^* = a_{ij}$, which implies that there are $N(N-1)$ off diagonal components. To see this, note that, each matrix component is a complex number and hence has two components. Now do the counting similar to the independent off-diagonal terms in a symmetric or anti-symmetric matrix and you will get $\frac{N(N-1)}{2}$ terms but since each term is complex you get total $N(N-1)$ independent terms instead. Finally, let us impose the traceless condition to get the no. of linear independent traceless complex Hermitian $N\times N$ matrices to be: $N + N(N-1) - 1 = N^2 - 1$. So, $\mathfrak{su}(N)$ has N^2-1 generators which is also the vector space dimension of it.

Now let us take $N = 2$.

$$\begin{pmatrix} a_{11} & \tilde{a}_{12} \\ \tilde{a}_{12}^* & -a_{11} \end{pmatrix}; \quad a_{11}, a_{22} \in \mathbb{R}$$

To determine the components, you can do trial and error to get:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

[57] Let us look at the general case first and later we will set $N = 2$.

[58] Here, we used the matrix identity: $\det(e^X) = e^{\mathrm{tr}(X)}$. You can readily see this identity for diagonal matrices. Note that here we are dealing with Hermitian matrices which are always diagonalisable. However, the matrix identity is true in general.

with the generators being defined as:

$$J_a \equiv \frac{\sigma_a}{2},$$

where $\frac{1}{2}$ is a convention. Now computing the commutators we get:

$$[J_i, J_j] = i\,\epsilon_{ijk} J_k, \tag{D.4}$$

which is the $\mathfrak{su}(2)$ algebra.

D.2 $G = SO(3)$

Now let us discuss the Lie algebra for the group $SO(3)$, which is a group of 3×3 orthogonal matrices with unit determinant. To begin, let us work with real matrices here and keep N arbitrary. We will later set $N = 3$.

Now expanding about the identity group element and using the definition of $SO(N)$ matrices you should be able to obtain that the generators are anti-symmetric $N \times N$ matrices. Hence, there will be $\frac{N(N-1)}{2}$ generators of the $\mathfrak{so}(3)$ Lie algebra. Let us denote the generators as: L_i.

§A1: Show that the generators of $SO(N)$ Lie algebra are $N \times N$ anti-symmetric matrices.

A convenient basis is to use the $SO(N)$ invariant tensor – the Levi-Civita tensor: $\epsilon_{i_1 \ldots i_n}$ [59]. With this, let us define the generators to be:

$$(L_k)_{ij} \equiv -\epsilon_{kij}, \tag{D.5}$$

and explicitly they are:

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad L_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{D.6}$$

They satisfy:

$$[L_i, L_j] = \epsilon_{ijk} L_k \tag{D.7}$$

If we now allow for complex generators and define:

$$J_i \equiv -i L_i \tag{D.8}$$

then from (D.7) we get the $\mathfrak{su}(2)$ algebra as given in (D.4). Thus, the two Lie algebras are *isomorphic*, $\mathfrak{su}(2) \cong \mathfrak{so}(3)$.

[59] To prove this, use the definition of determinant of a square matrix: $\det(A) = \frac{1}{n!}\epsilon^{i_1 \ldots i_n}\epsilon_{j_1 \ldots j_n} A^{j_1}_{i_1} \ldots A^{j_n}_{i_n}$ and then take $A \in SO(N)$ which then implies $\det(A) = 1$ leading to invariance of $\epsilon_{i_1 \ldots i_n}$ under $SO(N)$ transformations. You will also need the following Levi-Civita identity: $\epsilon^{i_1 \ldots i_n}\epsilon_{i_1, \ldots i_n} = n!$. Note that here upper and lower indices are the same as they are lowered and raised respectively using δ_{ij}.

D.3 $G = U(1)$

§A2: Find the generators for the Lie algebra of $U(1)$. What is the vector space described by $\mathfrak{u}(1)$ and what is its dimension? Draw a picture in the complex plane to denote $U(1)$ and its Lie algebra $\mathfrak{u}(1)$. What geometric interpretation you get for the Lie algebra?