

Willoughby Seago

Theoretical Physics

Quantum Field Theory

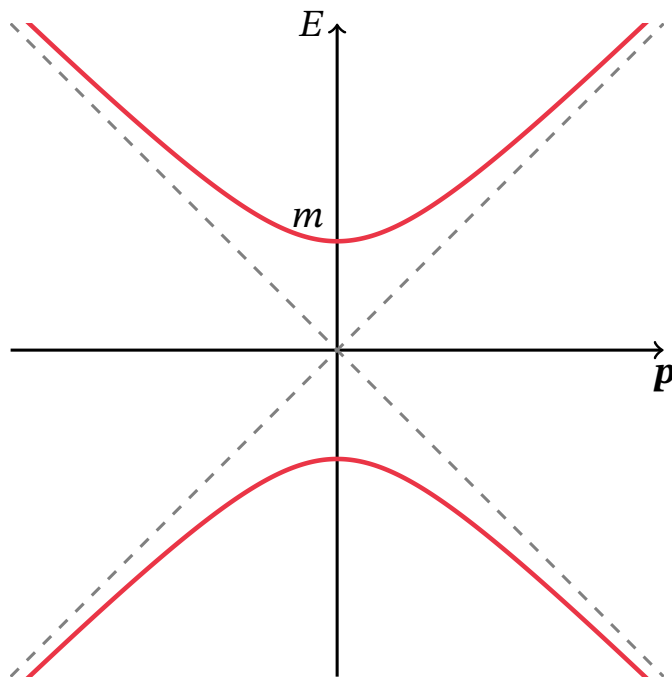
COURSE NOTES

Quantum Field Theory

Willoughby Seago

These are my notes from the course quantum field theory. I took this course as a part of the theoretical physics degree at the University of Edinburgh.

These notes were last updated at 23:23 on October 19, 2022. For notes on other topics see <https://github.com/WilloughbySeago/Uni-Notes>.



Chapters

	Page
Chapters	ii
Contents	iv
List of Figures	ix
1 Introduction	1
I Canonical Quantisation	4
2 Classical Fields	5
3 Lagrangians and Hamiltonians	11
4 Quantised Fields	24
5 Quantum Harmonic Oscillator	32
6 Mode Expansions	40
7 Covariant Commutators	50
II Interactions	55
8 Scattering	56
9 Wick's Theorem	66
10 Feynman Diagrams	76
11 Getting Measurable Results	88

<i>CHAPTERS</i>	iii
III Complex Scalar Fields	93
12 Charge	94
13 Interactions and Symmetries	101
IV Dirac Equation	109
14 Deriving the Dirac Equation	110
15 Dirac Lagrangian	119
16 Solutions to the Dirac Equation	124
V Gaussian Integrals	129
17 Gaussian Integrals	130
18 Perturbed Correlators	138
VI Path Integrals in Quantum Mechanics	143
19 Path Integrals in QM	144
20 Generating Functions	150
VII Path Integrals in Scalar Field Theory	155
21 Noninteracting Scalar Field Theory	156
22 Interacting Scalar Field Theory	163
VIII Scalar Field Correlators	169
23 Correlators	170
Appendices	177
A Pre-Course Revision	178
B Hamiltonian Mode Expansion	194
Acronyms	198
Index	199

Contents

	Page
Chapters	ii
Contents	iv
List of Figures	ix
1 Introduction	1
1.1 Course Overview	1
1.2 Conventions	2
I Canonical Quantisation	4
2 Classical Fields	5
2.1 Relativity Basics	5
2.1.1 Four-Vectors	5
2.1.2 Specific-Four Vectors	6
2.2 Nonrelativistic Particle and Wave	7
2.3 Relativistic Particle and Wave	7
2.4 Maxwell Fields	8
3 Lagrangians and Hamiltonians	11
3.1 Lagrangian Dynamics of a Particle	11
3.2 Hamiltonian Dynamics of a Particle	13
3.3 Lagrangian Dynamics of a Field	15
3.4 Hamiltonian Dynamics of a Field	16
3.5 Klein–Gordon Equation	17
3.6 Symmetries	19
3.6.1 Discrete Symmetries	19
3.6.2 Continuous Symmetries	20
4 Quantised Fields	24
4.1 Quantum Mechanics	24
4.1.1 The Basics	24
4.1.2 Time Dependence	25

4.2	Quantum Fields	28
4.2.1	Equations of Motion	29
5	Quantum Harmonic Oscillator	32
5.1	Solution	32
5.2	Interpretations	36
5.2.1	Wave Function Interpretation	36
5.2.2	Quanta	37
5.3	Harmonic Oscillator with Time Dependence	38
6	Mode Expansions	40
6.1	Expanding the Field	40
6.2	The Fourier Coefficients	42
6.3	The Hamiltonian	44
6.4	Particle Interpretation	45
6.5	Ground State Energy	47
6.5.1	Normal Ordering	47
6.6	Ground State Momentum	48
7	Covariant Commutators	50
7.1	What are They	50
7.2	Microcausality	51
7.3	Equal Time Commutation Relations	52
7.4	Contour Representation	53
7.5	Δ as Green's Functions	54
II	Interactions	55
8	Scattering	56
8.1	Interactions	56
8.1.1	φ^3 Theory	56
8.2	The S Matrix	57
8.3	Interaction Picture	58
8.4	Dyson Series	59
8.5	Interpretation	63
8.6	φ^3 Interactions	63
9	Wick's Theorem	66
9.1	Two Operator Case	66
9.2	Three Operator Case	68
9.3	General Theorem	69
9.4	Feynman Propagator	72
9.4.1	Contour Representation	73
9.4.2	Feynman Propagator as a Green's Function	74
10	Feynman Diagrams	76
10.1	Evaluating Low Order Terms	76
10.1.1	$n = 0$	76
10.1.2	$n = 1$	77
10.1.3	$n = 2$	78

10.2 Feynman Diagrams in Momentum Space	81
10.2.1 First Order	82
10.2.2 Second Order	83
10.3 Feynman Rules	87
11 Getting Measurable Results	88
11.1 Transition Rate	88
11.1.1 One to Two Decay	90
11.2 Cross Sections	90
11.2.1 Two to Two Scattering	91
 III Complex Scalar Fields	 93
12 Charge	94
12.1 Multiple Real Fields	94
12.2 Complex Scalar Field	95
12.3 Mode Expansion	96
12.4 Charge Conservation	97
12.5 Feynman Rules	99
 13 Interactions and Symmetries	 101
13.1 Types of Interaction	101
13.1.1 Four Point Interactions	101
13.1.2 Three Point Interactions	102
13.2 Symmetries	104
13.2.1 Parity	104
13.2.2 Time Reversal	106
13.2.3 Charge Conjugation	107
13.2.4 Acting on the Lagrangian	108
 IV Dirac Equation	 109
14 Deriving the Dirac Equation	110
14.1 The Dirac Equation	110
14.2 Covariant Form	112
14.3 Spinor Transformations	114
14.3.1 Boosts	116
14.3.2 Rotations	116
14.3.3 General Lorentz Transformation	117
 15 Dirac Lagrangian	 119
15.1 Dirac Adjoint	119
15.2 Dirac Bilinears	120
15.3 Dirac Lagrangian	121
 16 Solutions to the Dirac Equation	 124
16.1 Free Particle Solution	124
16.1.1 Positive Energy Solutions	125
16.1.2 Negative Energy Solutions	126

16.2 Spin, Angular Momentum and Helicity	127
16.3 Normalisation	127
V Gaussian Integrals	129
17 Gaussian Integrals	130
17.1 One Dimension	130
17.1.1 No Constant Coefficient	130
17.1.2 With Coefficient	130
17.1.3 With Linear Term	131
17.2 Arbitrary Dimension	131
17.2.1 Diagonal Case	131
17.2.2 Diagonalisable Case	132
17.2.3 With Linear Term	132
17.3 Generating Function	133
17.4 Gaussian Generating Functions	134
17.4.1 One Point Gaussian Correlator	135
17.4.2 Two Point Gaussian Correlator	135
17.4.3 General Gaussian Correlator	136
18 Perturbed Correlators	138
18.1 General Theory	138
18.2 x^4 Potential	139
18.2.1 Diagrammatic Representation	140
18.2.2 Computing $\log(Z_\lambda/Z_0)$	141
18.3 Correlators	141
VI Path Integrals in Quantum Mechanics	143
19 Path Integrals in QM	144
19.1 Preliminaries	144
19.2 Constructing the Path Integral	144
19.2.1 Quadratic Momentum	146
19.3 Correlators	147
19.3.1 One Point Correlator	147
19.3.2 Two Point Correlator	148
20 Generating Functions	150
20.1 Functional Derivatives	150
20.2 Sources	150
20.3 Projection onto the Ground State	152
20.4 Weyl Ordering	153
VII Path Integrals in Scalar Field Theory	155
21 Noninteracting Scalar Field Theory	156
21.1 Scalar Field Theory Recap	156
21.2 The Path Integral	157

21.2.1 With a Source Term	159
21.2.2 Using the Path Integral	162
22 Interacting Scalar Field Theory	163
22.1 General Theory	163
22.2 φ^3 Interactions	164
22.3 Disconnected Diagrams	167
VI Scalar Field Correlators	169
23 Correlators	170
23.1 Correlators	170
23.2 Two Point Correlator	170
23.2.1 $V = 0$	170
23.2.2 $V = 1$	171
23.2.3 $V = 2$	171
23.3 Momentum Space	173
23.3.1 $V = 0$	174
23.3.2 $V = 2$	174
Appendices	177
A Pre-Course Revision	178
A.1 Special Relativity	178
A.1.1 Four Vectors	178
A.1.2 Lorentz Transformations	179
A.1.3 Classification of Four-Vectors	181
A.1.4 Differential Operators	182
A.1.5 Momentum and Energy	182
A.2 Electromagnetic Theory	182
A.2.1 Maxwell's Equations	182
A.2.2 Four-Vector Potential	183
A.2.3 Gauge Transformations	184
A.2.4 Covariant Form of Maxwell's Equations	184
A.2.5 Interaction of Matter with EM Fields	185
A.3 Relativistic Quantum Mechanics	186
A.3.1 Klein–Gordon Equation	186
A.3.2 Continuity Equation and Probability Interpretation	187
A.3.3 Dirac Equation	188
A.3.4 Probability Density	190
A.3.5 Free Particle Solutions	191
B Hamiltonian Mode Expansion	194
Acronyms	198
Index	199

List of Figures

	Page
2.1 Nonrelativistic energy-momentum relation.	7
2.2 Relativistic energy-momentum relation	8
5.1 The harmonic potential and the evenly spaced energy levels it produces.	37
7.1 The contours used in the contour representation of Δ^\pm and Δ	53
8.1 The region of integration for the $n = 2$ case of the identity.	62
9.1 The contour used to compute the Feynman propagator.	73
9.2 The contour for calculating the Feynman propagator using the $i\epsilon$ pre- description.	74
14.1 Rotating two particles into each other.	118

One

Introduction

1.1 Course Overview

Quantum field theory (QFT) is combination of relativity and quantum mechanics. It will be our topic of study in this course. We will focus mostly on a single quantum field theory, namely **quantum electrodynamics (QED)**, which is the result of combining electrodynamics, which is inherently relativistic, and quantum mechanics.

There are various courses leading into this course, some of the important ones for which I have notes are

- *Principles of Quantum Mechanics*;
- *Quantum Theory*;
- *Classical Electrodynamics*; and
- *Symmetries of Quantum Mechanics*.

This course has a companion course, *Symmetries of Particles and Fields*, which focuses on the mathematical abstraction of symmetry. Ideas from this companion course, and the related course *Symmetries of Quantum Mechanics* will occur throughout this course. Another related course, focused more on the experimental side of this field, is *Particle Physics*. I have notes for both *Symmetries of Particles and Fields* and *Particle Physics*.

This course, and its companion, lead naturally into the course *Gauge Theories in Particle Physics*, for which I also have notes. Many ideas in this course will be expanded upon in the gauge theories course. The gauge theories course will also treat other quantum field theories, such as **electroweak** interactions (a combination of electromagnetism and the weak interactions), and **quantum chromodynamics (QCD)** (the quantum field theory of the strong force). The course also briefly touches on lattice gauge theory, a particular way of doing calculations in nonperturbative QCD.

This course is roughly divided into four sections:

- Canonical QFT: Here we deal with operators and quantisation, familiar from traditional quantum mechanics, such as in *Principles of Quantum Mechanics*.
- QED: We will use this as an example, to allow us to work in a concrete setting rather than in the abstract. This is in many ways the simplest of quantum

field theories, we study mostly interactions of photons and electrons, and since photons aren't charged the interactions are about as simple as they get. Mathematically this "simplicity" is due to the underlying $U(1)$ symmetry, which is an Abelian gauge group, and is significantly easier to work with than, say, the $SU(3)$ symmetry of QCD.

- Path integral formalism: This is a more abstract, but often more powerful, yet equivalent formulation of quantum field theory. This formalism leads itself to gauge theories, and so will be used heavily in *Gauge Theories in Particle Physics*. Nonrelativistic path integrals were treated in *Quantum Theory*.
- Renormalisation: When we do QFT calculations we often get infinite results. Renormalisation is the process of removing these infinities and extracting meaningful results. The renormalisation of QED will be treated in more detail in *Gauge Theories*.

There is some concurrent teaching of these topics, but in these notes I separate them out, so if something's not making sense maybe check to see if its been explained in more detail in a different section.

1.2 Conventions

Quantum field theory is full of conventions. One convention that almost everyone follows is that we work in natural units, where $c = \hbar = 1$. This makes the formulas look much simpler and less cluttered, and we can put c and \hbar back in by dimensional analysis.

The second convention is unfortunately much more varied, its the choice of metric. We will use the $(+---)$ metric,

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (1.2.1)$$

The most common alternative to this is $(-+++)$, where the diagonal has a single negative 1 and the rest is made of positive 1s. There are some even rarer choices to use an imaginary metric, $(i+++)$, which still gives the same relative minus sign when we square each element in the inner product.

Throughout the course, unless specified otherwise, we will use the Einstein summation convention. Specifically, if an index appears exactly twice in a term, once in a raised position and once in a lowered position, then we sum over all values of that index. We also follow the convention where Greek indices, such as μ and ν , run from 0 to the number of dimensions minus one, most commonly meaning $\mu = 0, 1, 2, 3$. On the other hand, Latin indices, such as i and j , run from 1 to the number of dimensions minus one, most commonly, $i = 1, 2, 3$. This means that Greek indices are summed over all components, whereas Latin indices are summed only over spatial components.

We follow the convention that the electromagnetic field strength tensor is

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (1.2.2)$$

An alternative convention is

$$F^{\mu\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu, \quad (1.2.3)$$

which differs from our choice by an overall minus sign. For example, this results in Maxwell's equations being written as $\partial_\mu F^{\mu\nu} = -j^\nu$, instead of $\partial_\mu F^{\mu\nu} = j^\nu$, which is what we'll be using.

Part I

Canonical Quantisation

Two

Classical Fields

2.1 Relativity Basics

In order to understand quantum field theory we will need to quantise fields. In order to quantise fields we should be sure that we understand classical fields, so this is where we start the course. We will briefly recap ideas mostly from the *Classical Electrodynamics* and *Quantum Theory* courses, so check the notes for this course for more details.

2.1.1 Four-Vectors

We can consider a generic four-vector,

$$a^\mu = (a^0, \mathbf{a}) = (a^0, a^1, a^2, a^3). \quad (2.1.1)$$

Using the metric, $\eta_{\mu\nu} := \text{diag}(1, -1, -1, -1)$, we can lower the indices to get a covariant four-vector:

$$a_\nu := \eta_{\mu\nu} a^\mu = (a^0, -\mathbf{a}) = (a^0, -a^1, -a^2, -a^3) = (a_0, a_1, a_2, a_3). \quad (2.1.2)$$

Given two four vectors, $a^\mu = (a^0, \mathbf{a})$ and $b = (b^0, \mathbf{b})$, we can take the inner product,

$$a \cdot b := a^\mu b_\mu = a_\mu b^\mu = \eta_{\mu\nu} a^\mu b^\nu = a^0 b^0 - \mathbf{a} \cdot \mathbf{b}. \quad (2.1.3)$$

Here $\mathbf{a} \cdot \mathbf{b} = a^i b^i = a^1 b^1 + a^2 b^2 + a^3 b^3$ is the standard dot product for three-vectors.

For the specific case of the inner product of a four-vector with itself we use the shorthand

$$a^2 = a \cdot a = a^0 a^0 - \mathbf{a} \cdot \mathbf{a} = (a^0)^2 - \mathbf{a}^2. \quad (2.1.4)$$

Four vectors transform under Lorentz transformations. The broadest class of such transformations is the **Lorentz group**, $O(1, 3)$. This includes all boosts and rotations, including those that invert the directions of space or time. In this course we only consider **proper orthochronous Lorentz transformations**, that is Lorentz transformations preserving the orientation of space (proper) and time (orthochronous). These form the **proper orthochronous Lorentz group** $SO^+(1, 3)$. From now on if we say Lorentz transformation we mean *proper orthochronous* Lorentz transformation.

2.1.2 Specific-Four Vectors

The position is a four-vector, $x^\mu = (t, \mathbf{x})$, where t is the time coordinate and \mathbf{x} is the spatial position. Note that in SI units this would be $x^\mu = (ct, \mathbf{x})$, since the components of a four-vector must have the same dimensions. We can also construct a covariant position four-vector, $x_\mu = (t, -\mathbf{x})$.

The momentum is a four-vector, $p^\mu = (E, \mathbf{p})$, where E is the energy, and \mathbf{p} is the relativistic three-momentum. Note that in SI units this would be $p^\mu = (E/c, \mathbf{p})$. We can also construct a covariant momentum four-vector, $p_\mu = (E, -\mathbf{p})$.

¹as in, not a virtual particle

For a (real¹) particle of mass m the four-momentum squares to the mass squared, that is

$$m^2 = p^2 = E^2 - \mathbf{p}^2. \quad (2.1.5)$$

Rearranging this gives us

$$E^2 = \mathbf{p}^2 + m^2. \quad (2.1.6)$$

This is the relativistic **energy-momentum relation**, it's perhaps more familiar if we reinstate the factors of c :

$$E^2 = \mathbf{p}^2 c^2 + m^2 c^4, \quad (2.1.7)$$

and is most famous in the case where $\mathbf{p} = \mathbf{0}$:

$$E = mc^2. \quad (2.1.8)$$

We can construct a four-vector derivative by defining the derivative with respect to x^μ :

$$\partial_\mu := \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \nabla \right). \quad (2.1.9)$$

Note that this is a covariant vector, since the contravariant quantity, x^μ , appears in the denominator, so the derivative transforms in the opposite way to how it would if x^μ were in the numerator. We can similarly define a contravariant operator,

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial t}, -\nabla \right). \quad (2.1.10)$$

The square of these operators is common enough to be given it's own name, it's called the **d'Alembert operator**,

$$\partial^2 = \partial^\mu \partial_\mu = \frac{\partial^2}{\partial t^2} - \nabla^2. \quad (2.1.11)$$

This same quantity is also denoted \square^2 , as a sort of four-dimensional (note four sides) Lorentzian-manifold analogue of the Laplacian, ∇^2 , and also confusingly sometimes denoted \square , without the superscript 2. We also call this the wave operator, since if f is a field satisfying $\partial^2 f = 0$ then expanding and rearranging this we have

$$\frac{\partial^2 f}{\partial t^2} = \frac{\partial^2 f}{\partial x^2}, \quad (2.1.12)$$

which is the equation of a wave travelling at the speed of light.

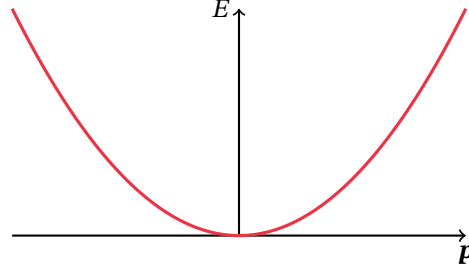


Figure 2.1: The nonrelativistic energy-momentum relation assigns a single energy to each three-momentum. Note that this is really just a slice through a four-dimensional plot, but all that really matters is the magnitude of the three-momentum.

2.2 Nonrelativistic Particle and Wave

A nonrelativistic particle of mass m has the energy-momentum relation

$$E = \frac{\mathbf{p}^2}{2m} = \frac{1}{2}m\mathbf{v}^2. \quad (2.2.1)$$

To get an equation for this particle we substitute in the usual operators,

$$E \rightarrow i\frac{\partial}{\partial t}, \quad \text{and} \quad \mathbf{p} \rightarrow -i\nabla. \quad (2.2.2)$$

Acting on an arbitrary state, ψ , we then get

$$i\frac{\partial\psi}{\partial t} = \frac{(-i\nabla)^2}{2m}\psi = -\frac{1}{2m}\nabla^2\psi. \quad (2.2.3)$$

This is exactly the **Schrödinger equation** for a free particle.

One solution to this is a plane wave,

$$\psi(x) = e^{-iEt+i\mathbf{p}\cdot\mathbf{x}} = e^{-ip\cdot x}. \quad (2.2.4)$$

This is only a solution if $E^2 = \mathbf{p}^2/2m$. Notice that while we can write the exponent with the relativistic four-vectors this is still nonrelativistic since the energy-momentum is not relativistic. In particular, we have a single energy value for each possible momentum, as seen in [Figure 2.1](#).

We know that solutions to the Schrödinger equation can be interpreted as amplitudes, which we then take the modulus square of, $|\psi|^2$, to get a probability density function. We call a function with this property a **wave function**.

2.3 Relativistic Particle and Wave

A relativistic particle of mass m has the energy-momentum relation

$$E^2 = \mathbf{p}^2 + m^2. \quad (2.3.1)$$

Making the operator substitutions and acting on an arbitrary state, φ , we get

$$\left(i\frac{\partial}{\partial t}\right)^2\varphi = (-i\nabla)^2\varphi + m^2\varphi. \quad (2.3.2)$$

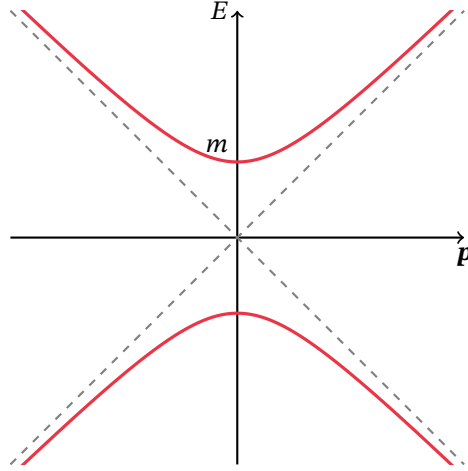


Figure 2.2: The relativistic energy momentum assigns two energies to a each three-momentum. The curve is a hyperbola. Notice that at zero three-momentum there we don't have zero energy, instead we have energy $\pm m$, since we are counting the mass towards the energy in the relativistic formulation. The hyperbola here are the "shell" referred to in "on-shell".

Rearranging this we get

$$(\partial^2 + m^2)\varphi = 0. \quad (2.3.3)$$

This is the **Klein–Gordon equation**.

One solution to this is a plane wave,

$$\varphi(x) = e^{-ip \cdot x}. \quad (2.3.4)$$

This is a solution provided that $p^2 = m^2$. If $p^2 = m^2$ then we say that p^μ is **on-shell**.

If $e^{-ip \cdot x}$ is a solution then so is

$$\varphi^* = e^{ip \cdot x}. \quad (2.3.5)$$

We see that we get pairs of solutions. These correspond to the pairs of possible energy values, $E = \pm\sqrt{\mathbf{p}^2 + m^2}$, as shown in [Figure 2.2](#).

The negative energy solutions to the Klein–Gordon equation pose a problem, in particular, it is possible for $\varphi^*\varphi$ to be negative, which means there is no probabilistic interpretation of φ , and so φ is *not* a wave function. Instead, we just call φ a field.

States of a system obeying the Klein–Gordon equation correspond to relativistic free particles (plural).

2.4 Maxwell Fields

The most familiar fields are electromagnetic fields. In this section we will briefly recap a relativistic treatment of these fields. For more details see the *Classical Electrodynamics* course.

The electromagnetic “potential” is

$$A^\mu(x) = (\varphi(x), \mathbf{A}(x)), \quad (2.4.1)$$

where φ is the electric potential and \mathbf{A} is the electromagnetic vector potential. The electric “field” is then

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (2.4.2)$$



An alternative convention,

$$F^{\mu\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu, \quad (2.4.3)$$

differs by a sign.

The electromagnetic current density is

$$j^\mu = (\rho, \mathbf{j}). \quad (2.4.4)$$

This current is conserved, mathematically this is expressed as its four-divergence vanishing,

$$\partial_\mu j^\mu = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \quad (2.4.5)$$

This is a continuity equation.

Maxwell’s equation is

$$\partial_\mu F^{\mu\nu} = j^\nu. \quad (2.4.6)$$

Writing $F^{\mu\nu}$ in terms of A^μ we get

$$\partial_\mu F^{\mu\nu} = \partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu = \partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu. \quad (2.4.7)$$

The **Lorenz gauge** condition is to choose A^μ such that its four-divergence vanishes, $\partial_\mu A^\mu = 0$. Then we have $\partial_\mu F^{\mu\nu} = \partial^2 A^\nu$, so we can write Maxwell’s equation as

$$\partial^2 A^\nu = j^\nu. \quad (2.4.8)$$

In QFT we call A^μ the **electromagnetic field** and we call $F^{\mu\nu}$ the **electromagnetic field strength**. This is partly an annoying historical artefact and partly because A^μ is the field appearing in our equations in an analogous way to φ in the Klein–Gordon equation.

In free space, that is when $j^\mu = 0$, Maxwell’s equation is

$$\partial^2 A^\mu = 0. \quad (2.4.9)$$

Notice that this is similar to the Klein–Gordon equation with massless particles, $m = 0$. This is good, because we would like electromagnetism to work with photons, which are massless. The one difference from the Klein–Gordon equation is that A^μ is a four-vector, whereas φ is a scalar. The only change in the solutions is that we, in theory, get a different solution for each component of A^μ . The way

that this manifests in the solutions is we get an extra vector out front called the **polarisation vector**, ϵ^μ :

$$A^\mu = \epsilon^\mu e^{-ip \cdot x}. \quad (2.4.10)$$

Assuming the momentum is on-shell, that is $p^2 = 0$ in this massless case. This means the momentum is light-like, more good news for electromagnetism explaining photons.

²See the *Classical Electrodynamics* course, note that there we work with $k^\mu = \hbar p^\mu$.

It can be shown that² $\epsilon_\mu p^\mu = 0$. We are also free to choose ϵ^μ such that $\epsilon^0 = 0$, then the condition $\epsilon_\mu p^\mu = 0$ becomes $\boldsymbol{\epsilon} \cdot \mathbf{p} = 0$. This means that electromagnetic waves are **transverse**, since their polarisation vector, $\boldsymbol{\epsilon}$, is perpendicular to their direction of travel, \mathbf{p} .

Since the mathematics of the vector field A^μ is so similar to the mathematics of the scalar field ϕ we will mostly deal with scalar fields in this course to develop our theory, then work with other fields once we have a good understanding off the maths. Each type of field corresponds to a different spin. Scalar fields describe spin zero particles, like the Higgs boson, and vector fields describe spin one particles, like the photon. Spin 1/2 particles are described by spinors.


Three

Lagrangians and Hamiltonians

3.1 Lagrangian Dynamics of a Particle

That's all there is to a Lagrangian dynamics course, the rest is just examples.

Richard Ball

 See the *Lagrangian Dynamics* course for more details.

Consider a particle of mass m moving in one dimension in a conservative force field. We can define its position with a **generalised coordinate**, q , which gives us a **generalised velocity**, $\dot{q} := dq/dt$. We can then define the **Lagrangian**:

$$L(q, \dot{q}) = T(\dot{q}) - V(q). \quad (3.1.1)$$

Here $T(\dot{q})$ is the kinetic energy of the particle, which we assume depends only on the velocity for simplicity, and $V(q)$ is the potential energy of the particle, which must depend only on the position. The Lagrangian is a function of the generalised position and velocity.

The **action** of the particle between the times t_1 and t_2 is defined to be

$$S[q(t)] := \int_{t_1}^{t_2} dt L(q, \dot{q}). \quad (3.1.2)$$

Hamilton's principle states that the path the particle takes, $q(t)$, is such that the action is extremised, that is the variation, δS , given by varying $q \rightarrow q + \delta q$ vanishes. We can compute the variation in a general action, S , by computing the variation in the Lagrangian:

$$\delta L = \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q}. \quad (3.1.3)$$

For a smooth variation, δq , we have

$$\delta \dot{q} = \frac{\partial}{\partial t}(\delta q), \quad (3.1.4)$$

and so

$$\delta L = \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{d}{dt}(\delta q) \quad (3.1.5)$$

Integrating this the variation in the action is

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{d}{dt}(\delta q) \right]. \quad (3.1.6)$$

The product rule tells us that

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) = \frac{\partial L}{\partial \dot{q}} \frac{d}{dt}(\delta q) + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \delta q. \quad (3.1.7)$$

Rearranging this we can rewrite the second term in δS as a total derivative minus a term:

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial q} \delta q + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \delta q \right] \quad (3.1.8)$$

$$= \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \delta q + \int_{t_1}^{t_2} dt \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) \quad (3.1.9)$$

$$= \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \delta q + \left[\frac{\partial L}{\partial \dot{q}} \delta q \right]_{t_1}^{t_2}. \quad (3.1.10)$$

Now suppose that δq vanishes at t_1 and t_2 , so $\delta q(t_1) = \delta q(t_2) = 0$, then the last term above vanishes and we have

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \delta q. \quad (3.1.11)$$

If δS is to vanish for all variations δq then we must have

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) = 0. \quad (3.1.12)$$

This is the **Euler-Lagrange equation** for a single particle in one dimension.

This generalises to n -dimensional space by replacing q with n coordinates q_i and \dot{q} with \dot{q}_i . The Lagrangian is then

$$L(q_i, \dot{q}_i) = T(\dot{q}_i) - V(q_i), \quad (3.1.13)$$

and the action is

$$S[q_i(t)] = \int_{t_1}^{t_2} dt L(q_i, \dot{q}_i). \quad (3.1.14)$$

We then get n Euler-Lagrange equations:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0. \quad (3.1.15)$$

Example 3.1.16 The Lagrangian for a simple harmonic oscillator is

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2x^2. \quad (3.1.17)$$

We have

$$\frac{\partial L}{\partial x} = -m\omega^2x, \quad \text{and} \quad \frac{\partial L}{\partial \dot{x}} = m\dot{x} \implies \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = m\ddot{x}. \quad (3.1.18)$$

Hence,

$$m\ddot{x} = -m\omega^2x \implies \ddot{x} = -\omega^2x \quad (3.1.19)$$

is the equation of motion for the simple harmonic oscillator, which is exactly what we would expect.

3.2 Hamiltonian Dynamics of a Particle

Let's go back to the one-dimensional case. We can define the **canonical momentum** of the particle as

$$p := \frac{\partial L}{\partial \dot{q}}. \quad (3.2.1)$$

Note that, in general, this is *not* the normal momentum.

We then define the **Hamiltonian**, H , as a function of the generalised position and canonical momentum:

$$H(q, p) := p\dot{q} - L(q, \dot{q}). \quad (3.2.2)$$

Note that we must eliminate any \dot{q} remaining in the expression for the Hamiltonian, this can be done by solving the defining relation $p = dL/d\dot{q}$ for \dot{q} and then substituting in the result.

Now consider what happens when we vary H . Varying the left hand side we get

$$dH = \frac{\partial H}{\partial q} dq + \frac{\partial H}{\partial p} dp. \quad (3.2.3)$$

Varying the right hand side we have

$$dH = p dq + q dp - \frac{\partial L}{\partial q} dq - \frac{\partial L}{\partial \dot{q}} d\dot{q}. \quad (3.2.4)$$

We can rewrite the last term in terms of the canonical momentum:

$$dH = p dq + q dp - \frac{\partial L}{\partial q} dq - p d\dot{q} = q dp - \frac{\partial L}{\partial q} dq. \quad (3.2.5)$$

Using the Euler–Lagrange equations the second term can be rewritten to give

$$dH = q dp - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) dq = q dp - \dot{p} dq. \quad (3.2.6)$$

Comparing this with the result for the left hand side variation we can read off

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial q}. \quad (3.2.7)$$

These are **Hamilton's equations**.

This all generalises to n dimensions. First replace q with q_i , and p with $p_i := \partial L / \partial \dot{q}_i$. Then the Hamiltonian is

$$H(q_i, p_i) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i). \quad (3.2.8)$$

Hamilton's equations are

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \quad (3.2.9)$$

It can be useful when doing Hamiltonian mechanics to define the **Poisson bracket** of two functions, A and B , depending on position, q_i , and momentum, p_i :

$$\{A, B, :=\} \sum_i \left(\frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} - \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} \right). \quad (3.2.10)$$

This has the advantage of allowing us to express Hamilton's equations in the more symmetric form

$$\dot{q}_i = \{H, q_i\}, \quad \text{and} \quad \dot{p}_i = \{H, p_i\}. \quad (3.2.11)$$

Note the similarity to Heisenberg's operator equation of motion for a time independent operator, A :

$$\dot{A}(t) = i[H, A]. \quad (3.2.12)$$

Replacing Poisson brackets with i times the commutator can get us quite a long way in quantum mechanics.

The Hamiltonian is conserved. To show this consider the time derivative:

$$\frac{dH}{dt} = \frac{\partial H}{\partial p} \dot{p} + \frac{\partial H}{\partial q} \dot{q} = \dot{q} \dot{p} - \dot{p} \dot{q} = 0. \quad (3.2.13)$$

We can often identify the Hamiltonian with the total energy when this is also conserved.

Example 3.2.14 The canonical momentum for a simple harmonic oscillator is

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}, \quad (3.2.15)$$

which is just the normal momentum in this case. The Hamiltonian is then

$$H = p\dot{q} - L = m\dot{x}^2 - \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2 = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2, \quad (3.2.16)$$

which is exactly the energy of a simple harmonic oscillator.

3.3 Lagrangian Dynamics of a Field

Lagrangian dynamics doesn't change that much when we work with fields. We start by replacing the generalised coordinate $q(t)$ with the field, $\varphi(x)$, which now depends on a position in spacetime, rather than just a time. Instead of a Lagrangian we work with a **Lagrangian density**, $\mathcal{L}(\varphi, \partial_\mu \varphi)$, replacing the time derivative of the coordinate with a four-vector derivative. We then have to integrate over a region of spacetime to get the action:

$$S[\varphi(x)] := \int d^4x \mathcal{L}(\varphi, \partial_\mu \varphi). \quad (3.3.1)$$

The integral should be taken over some arbitrary region of spacetime. Note that we can write this as

$$S[\varphi(x)] = \int dt \int d^3\mathbf{x} \mathcal{L}(\varphi, \partial_\mu \varphi) = \int dt L(\varphi, \partial_\mu \varphi) \quad (3.3.2)$$

where

$$L(\varphi, \partial_\mu \varphi) := \int d^3\mathbf{x} \mathcal{L}(\varphi, \partial_\mu \varphi). \quad (3.3.3)$$

So, integrating \mathcal{L} over space gives L , which we call the Lagrangian, which explains the interpretation of \mathcal{L} as a Lagrangian *density*. While this distinction is important the Lagrangian doesn't actually appear that much in QFT, so people often just call \mathcal{L} the Lagrangian, leaving the density part implicit.

The Euler–Lagrange equations for the Lagrangian density don't change much, and can be derived in a very similar way. We start by varying the field, $\varphi \rightarrow \varphi + \delta\varphi$, and looking at the resulting variation in the Lagrangian density:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta(\partial_\mu\varphi). \quad (3.3.4)$$

For a smooth variation we have

$$\delta(\partial_\mu\varphi) = \partial_\mu(\delta\varphi) \quad (3.3.5)$$

and so

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\partial_\mu(\delta\varphi). \quad (3.3.6)$$

We can use the product rule,

$$\partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi \right) = \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \right) \delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \partial_\mu(\delta\varphi), \quad (3.3.7)$$

to write the last term as a total derivative minus a term:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi \right) - \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \right) \delta\varphi. \quad (3.3.8)$$

Integrating over some spacetime region, Ω , to get the variation in the action we have

$$\delta S = \int_{\Omega} d^4x \left[\frac{\partial \mathcal{L}}{\partial \varphi} \delta \varphi + \partial_{\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi)} \delta \varphi \right) - \partial_{\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi)} \right) \delta \varphi \right] \quad (3.3.9)$$

$$= \int_{\Omega} d^4x \left[\frac{\partial \mathcal{L}}{\partial \varphi} \delta \varphi - \partial_{\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi)} \right) \delta \varphi \right] + \int_{\Omega} d^4x \partial_{\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi)} \delta \varphi \right). \quad (3.3.10)$$

Applying the divergence theorem to the second integral we can replace it with an integral of $(\partial \mathcal{L} / \partial (\partial_{\mu} \varphi)) \delta \varphi$ over the boundary, $\partial \Omega$. Choosing a variation, $\delta \varphi$, which vanishes on the boundary this term will vanish, and we will be left with

$$\delta S = \int_{\Omega} d^4x \left[\frac{\partial \mathcal{L}}{\partial \varphi} \delta \varphi - \partial_{\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi)} \right) \delta \varphi \right]. \quad (3.3.11)$$

For this to vanish for all variations of the field we require that

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_{\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi)} \right) = 0. \quad (3.3.12)$$

These are the **Euler-Lagrange equations** for a scalar field.

3.4 Hamiltonian Dynamics of a Field

Following the same logic as for a single particle we can define the **canonical momentum**,

$$\pi(x) := \frac{\partial \mathcal{L}}{\partial \dot{\varphi}}. \quad (3.4.1)$$

This is *not* the same as the actual momentum. Notice that this definition treats time differently to position, this will result in the Hamiltonian formulation not being Lorentz invariant, although the results we get still are, the steps between are just frame dependent. This corresponds to the non-Lorentz invariance of the energy, the energy of a particle of mass m is γm .

We then define the Hamiltonian, H , to be

$$H := \int \pi \dot{\varphi} d^4x - L = \int d^4x (\pi \dot{\varphi} - \mathcal{L}) = \int d^3x \mathcal{H} \quad (3.4.2)$$

where

$$\mathcal{H} := \pi \dot{\varphi} - \mathcal{L} \quad (3.4.3)$$

is the **Hamiltonian density**. Both the Hamiltonian and Hamiltonian density must be functions of the field, φ , and the canonical momentum, π , which is just another field.

The same derivation as before works without modification, so we won't repeat it here, the result is **Hamilton's equations** for a field:

$$\dot{\varphi} = \frac{\partial \mathcal{H}}{\partial \pi}, \quad \text{and} \quad \dot{\pi} = -\frac{\partial \mathcal{H}}{\partial \varphi}. \quad (3.4.4)$$

As before, the Hamiltonian density is conserved:

$$\frac{d\mathcal{H}}{dt} = \int d^3\mathbf{x} \left(\frac{\partial \mathcal{H}}{\partial \varphi} \dot{\varphi} + \frac{\partial \mathcal{H}}{\partial \pi} \dot{\pi} \right) \quad (3.4.5)$$

$$= \int d^3\mathbf{x} (-\dot{\pi} \dot{\varphi} + \dot{\varphi} \dot{\pi}) \quad (3.4.6)$$

$$= 0. \quad (3.4.7)$$

Again, when energy is conserved we can often identify it with the Hamiltonian.

3.5 Klein-Gordon Equation

We have to learn to run before we walk.

Richard Ball

The Lagrangian for a scalar field, φ , is

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \varphi) (\partial^\mu \varphi) - \frac{1}{2} m^2 \varphi^2. \quad (3.5.1)$$

Clearly we have

$$\frac{\partial \mathcal{L}}{\partial \varphi} = -m^2 \varphi. \quad (3.5.2)$$

We can also compute the derivative with respect to $\partial_\mu \varphi$, note that $(\partial_\mu \varphi) (\partial^\mu \varphi)$ is just $(\partial \varphi)^2$, so we can treat this just like the derivative of a quantity squared, giving

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} = \partial^\mu \varphi. \quad (3.5.3)$$

We end with an upper index, since the left hand side has a lower index in the denominator, which is an upper index in the numerator. We then have

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \varphi} \right) = \partial_\mu \partial^\mu \varphi = \partial^2 \varphi. \quad (3.5.4)$$

Using the Euler-Lagrange equations we then have

$$\partial^2 \varphi + m^2 \varphi = 0, \quad (3.5.5)$$

which is just the Klein-Gordon equation.

The canonical momentum associated with this Lagrangian is

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \frac{\partial \mathcal{L}}{\partial (\partial_0 \varphi)} = \partial^0 \varphi = \dot{\varphi}. \quad (3.5.6)$$

The Hamiltonian is then

$$\mathcal{H} = \pi \dot{\varphi} - \mathcal{L} = \frac{1}{2} (\pi^2 + (\nabla \varphi)^2 + m^2 \varphi^2). \quad (3.5.7)$$

The $\nabla\varphi$ term comes from

$$\pi^2 - \frac{1}{2}(\partial_\mu\varphi)(\partial^\mu\varphi) = \pi^2 - \frac{1}{2}[(\partial_0\varphi)(\partial^0\varphi) - (\partial_i\varphi)(\partial^i\varphi)] \quad (3.5.8)$$

$$= \pi^2 - \frac{1}{2}[\dot{\varphi}^2 - (\nabla\varphi)^2] = \frac{1}{2}[\pi^2 - (\nabla\varphi)^2] \quad (3.5.9)$$

where we've used $\dot{\varphi} = \pi$.

Compare the Hamiltonian density for the Klein–Gordon equation with the Hamiltonian for the harmonic oscillator:

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 x^2). \quad (3.5.10)$$

Up to a factor of ω^2 and $1/m$ this is pretty much the same as the Klein–Gordon Hamiltonian density, minus the $(\nabla\varphi)^2$ term. Similarly the Klein–Gordon Lagrangian density is very similar to the harmonic oscillator Lagrangian. This suggests that we should study the harmonic oscillator, and indeed we shall soon.

Example 3.5.11 — Electromagnetism The Lagrangian density for the electromagnetic field, A^μ , is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (3.5.12)$$

We can insert the definition of $F^{\mu\nu}$ and expand this out to get

$$\mathcal{L} = -\frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) \quad (3.5.13)$$

$$= -\frac{1}{4}[(\partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\nu A^\mu)(\partial_\mu A_\nu) \quad (3.5.14)$$

$$- (\partial^\mu A^\nu)(\partial_\nu A_\mu) + (\partial^\nu A^\mu)(\partial_\nu A_\mu)] \quad (3.5.15)$$

$$= -\frac{1}{2}(\partial^\mu A^\nu)(\partial_\mu A_\nu) + \frac{1}{2}(\partial^\mu A^\nu)(\partial_\nu A_\mu). \quad (3.5.16)$$

We treat each component of A_ν as an independent field. We then have

$$\frac{\partial\mathcal{L}}{\partial A_\nu} = 0, \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\partial^\mu A^\nu + \partial^\nu A^\mu = -F^{\mu\nu}. \quad (3.5.17)$$

Hence, the Euler–Lagrange equations give

$$\partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = -\partial_\mu \partial^\mu A^\nu + \partial^\nu \partial_\mu A^\mu = 0. \quad (3.5.18)$$

That is,

$$\partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu = 0 \iff -\partial_\mu F^{\mu\nu} = 0 \quad (3.5.19)$$

This is Maxwell's equation in a vacuum in an arbitrary gauge.

The canonical momentum is

$$\pi^\mu(x) = \frac{\partial\mathcal{L}}{\partial\dot{A}_\mu} = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\mu)} = -\partial^\mu A^0 + \partial^0 A^\mu = -F^{0\mu}. \quad (3.5.20)$$

Antisymmetry of $F^{\mu\nu}$ implies that $\pi^\mu(x) = 0$, and so $\pi^\mu(x) = (0, -\mathbf{E}(x))$, where $\mathbf{E}(x)$ is the electric field.

The fact that $\pi^0(x) = 0$ will cause problems later when we try to quantise the electromagnetic field.

In the presence of sources the Lagrangian density is instead

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - J^\mu A_\mu. \quad (3.5.21)$$

This doesn't change the term given by differentiating with respect to $\partial_\mu A_\nu$, all that changes is we now have

$$\frac{\partial \mathcal{L}}{\partial A_\mu} = -J^\mu, \quad (3.5.22)$$

and so the result of applying the Euler–Lagrange equations is now

$$\partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu = J^\nu, \quad (3.5.23)$$

which is Maxwell's equation in the presence of a source.

3.6 Symmetries

3.6.1 Discrete Symmetries

There are three discrete symmetries of particular interest in quantum field theory, they are

- **parity:** $\mathcal{P} : \mathbf{x} \mapsto -\mathbf{x}$, $\mathcal{P} : x^\mu = (x^0, x^i) \mapsto (x^0, -x^i) = (x_0, x_i) = x_\mu$, so parity acts to swap raised and lowered indices of positions;
- **time reversal:** $\mathcal{T} : t \mapsto t$, $\mathcal{T} : x^\mu = (x^0, x^i) \mapsto (-x^0, x^i) = -(x^0, -x^i) = -(x_0, x_i) = -x_\mu$, so time reversal acts to swap raised and lowered indices and negate positions;
- **charge conjugation:** $\mathcal{C} : e \mapsto -e$, where e is the charge of the particle.

Consider, for example, the Lagrangian of the Klein–Gordon equation:

$$\mathcal{L} = \frac{1}{2}(\partial^\mu \varphi)(\partial_\mu \varphi) - \frac{1}{2}m^2 \varphi^2. \quad (3.6.1)$$

Under parity the scalar field is unchanged, and $\partial^\mu \leftrightarrow \partial_\mu$, so the Lagrangian is unchanged under parity. Under time reversal the scalar field is unchanged, and $\partial^\mu \leftrightarrow -\partial_\mu$, so the Lagrangian is unchanged under time reversal. Since this Lagrangian has no charges involved it is trivially invariant under charge conjugation.

Now consider the Electromagnetic Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - J^\mu A_\mu. \quad (3.6.2)$$

Under parity transformations $F^{\mu\nu} \leftrightarrow F_{\mu\nu}$, $J^\mu \rightarrow J_\mu$, and $A_\mu \rightarrow A^\mu$, so

$$\mathcal{P}\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J_\mu A^\mu = \mathcal{L}, \quad (3.6.3)$$

so the Lagrangian is unchanged by parity. Under time reversal transformations $F^{\mu\nu} \leftrightarrow -F_{\mu\nu}$ since the derivatives in the definition of $F^{\mu\nu}$ transform as $\partial^\mu \rightarrow -\partial_\mu$ under time reversal. On the other hand, $J^\mu \rightarrow J^\mu$, and $A_\mu \rightarrow A_\mu$, since time reversal doesn't effect the current or potential. Hence,

$$\mathcal{T}\mathcal{L} = -\frac{1}{4}(-F_{\mu\nu})(-F^{\mu\nu}) - J^\mu A_\mu = \mathcal{L}, \quad (3.6.4)$$


so the Lagrangian is unchanged under time reversal. Under charge conjugation $F^{\mu\nu} \leftrightarrow -F_{\mu\nu}$, $J^\mu \rightarrow -J_\mu$, and $A^\mu \rightarrow -A_\mu$, so

$$\mathcal{C}\mathcal{L} = -\frac{1}{4}(-F_{\mu\nu})(-F^{\mu\nu}) - (-J_\mu)(-A^\mu) = \mathcal{L}, \quad (3.6.5)$$

so the Lagrangian is invariant under charge conjugation.

Invariance under each of these symmetries means that both of these Lagrangians are invariant under the combined \mathcal{CPT} symmetry, given by $x^\mu \mapsto -x^\mu$ and $e \mapsto -e$.

3.6.2 Continuous Symmetries

 For more details see the *Symmetries of Particles and Fields* course.

If the action is invariant under a group of continuous symmetries then we have a conserved quantity. This is the essence of **Noether's theorem**. The most common examples being

- invariance under translations in time leading to energy conservation;
- invariance under spatial translations leading to momentum conservation;
- invariance under rotation leading to angular momentum conservation.

The simplest case is, perhaps, when the Lagrangian is invariant under translations, then the action is also necessarily translation invariant. Suppose we have a translation

$$\varphi(x) \rightarrow \varphi'(x) = \varphi(x) + \delta\varphi(x). \quad (3.6.6)$$

Note that the translation is allowed to depend on x . The variation in \mathcal{L} is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta(\partial_\mu\varphi) \quad (3.6.7)$$

$$= \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \right) \delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \delta(\partial_\mu\varphi) \quad (3.6.8)$$

$$= \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \delta\varphi \right). \quad (3.6.9)$$

The second equality is simply an application of the Euler–Lagrange equations to rewrite the first term. The last equality is just recognising the product rule.

Imposing that \mathcal{L} is invariant under translations like this we have $\delta\mathcal{L} = 0$, which gives us the continuity equation

$$\partial_\mu J^\mu(x) = 0, \quad (3.6.10)$$

where

$$J^\mu(x) := \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \delta \varphi \quad (3.6.11)$$

is the conserved current, a generalisation of the normal current in electrodynamics. There is also a corresponding conserved “charge”:

$$Q = \int d^3 \mathbf{x} J^0(x). \quad (3.6.12)$$

To see that this quantity is conserved consider

$$\dot{Q} = \int d^3 \mathbf{x} \partial_0 J^0(x) = - \int d^3 \mathbf{x} \partial_i J^i(x) = 0 \quad (3.6.13)$$

where in the last step we’ve used the divergence theorem to rewrite the last term as a surface integral, then take this surface to be at infinity, and made the usual assumption that fields vanish sufficiently quickly at infinity.

Example 3.6.14 — Probability Current Consider the Lagrangian

$$\mathcal{L} = i\psi^* \partial_t \psi - \frac{1}{2m} (\nabla \psi^*) \cdot (\nabla \psi) - V(\mathbf{x}, t) \psi^* \psi. \quad (3.6.15)$$

If we treat ψ and ψ^* as separate fields then varying one of them gives the Schrödinger equation. This Lagrangian is invariant under the transformation

$$\psi \rightarrow e^{i\alpha} \psi, \quad \text{and} \quad \psi^* \rightarrow e^{-i\alpha} \psi^* \quad (3.6.16)$$

where α is a constant. Linearising we get the transformation

$$\psi \rightarrow \psi(1 + i\alpha), \quad \text{and} \quad \psi^* \rightarrow \psi^*(1 - i\alpha). \quad (3.6.17)$$

Since we have two fields, ψ and ψ^* , the conserved current is a sum over field:

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \delta \psi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} \delta \psi^*. \quad (3.6.18)$$

Computing the derivatives we have

$$\frac{\partial \mathcal{L}}{\partial(\partial_t \psi)} = i\psi^*, \quad \frac{\partial \mathcal{L}}{\partial(\partial_i \psi)} = -\frac{1}{2m} \partial^i \psi^*, \quad (3.6.19)$$

$$\frac{\partial \mathcal{L}}{\partial(\partial_t \psi^*)} = 0, \quad \frac{\partial \mathcal{L}}{\partial(\partial_i \psi^*)} = -\frac{1}{2m} \partial^i \psi. \quad (3.6.20)$$

Hence, using the variations

$$\delta \psi = i\alpha \psi, \quad \text{and} \quad \delta \psi^* = -i\alpha \psi^* \quad (3.6.21)$$

we have the time component of the conserved current:

$$J^0 = \rho = -\alpha \psi^* \psi, \quad (3.6.22)$$

and the position components:

$$J^i = \frac{i\alpha}{2m}(\psi^* \partial^i \psi - \psi \partial^i \psi^*) \quad (3.6.23)$$

$$\mathbf{J} = \frac{i\alpha}{2m}(\psi \nabla \psi^* - \psi^* \nabla \psi). \quad (3.6.24)$$

We can then identify J^μ as, up to a constant, the conserved probability current of [Section A.3.2](#).

It can be shown that the Euler–Lagrange equations are invariant under the addition of a four-divergence,

$$\mathcal{L} \rightarrow \mathcal{L} + \partial_\mu \Lambda^\mu \quad (3.6.25)$$

for some four-vector field Λ^μ with a sufficiently smooth derivative. The same derivation as before, but setting $\delta\mathcal{L} = \partial_\mu \Lambda^\mu$ instead of zero, then gives a conserved current

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \delta \varphi - \Lambda^\mu. \quad (3.6.26)$$

As well as changing the Lagrangian directly we can instead act on spacetime, for example, by translation:

$$x^\mu \rightarrow x'^\mu = x^\mu + a^\mu \quad (3.6.27)$$

for some infinitesimal four-vector a^μ . This induces a change in the scalar field, given by Taylor expanding:

$$\varphi(x) \rightarrow \varphi(x + a) = \varphi(x) + a^\mu \partial_\mu \varphi(x). \quad (3.6.28)$$

Since the Lagrangian density is also a scalar it must transform the same way:

$$\mathcal{L} \rightarrow \mathcal{L} + a^\mu \partial_\mu \mathcal{L} = \mathcal{L} + a^\nu \partial_\mu (\delta^\mu_\nu \mathcal{L}). \quad (3.6.29)$$

Comparing this to the transformation $\mathcal{L} \rightarrow \mathcal{L} + \partial_\mu \Lambda^\mu$ we can identify

$$\partial_\mu \Lambda^\mu = a^\nu \partial_\mu (\partial^\mu_\nu \mathcal{L}). \quad (3.6.30)$$

We then have four conserved currents, one for each value of ν , such as

$$J^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} a^0 \partial_0 \varphi - a^0 \delta^\mu_0 \mathcal{L}, \quad (3.6.31)$$

which comes from setting $\nu = 0$. We can combine these four currents into a single rank 2 tensor, scaling out the a^ν parameter, to get

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \mathcal{L})} \partial_\nu \varphi - \mathcal{L} \delta^\mu_\nu. \quad (3.6.32)$$

Raising the index we get

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial^\nu \varphi - \mathcal{L} \eta^{\mu\nu}. \quad (3.6.33)$$

This is the **energy-momentum tensor**, also called the **stress-energy tensor**¹.

The conserved “charge” associated with the $\nu = 0$ component is

$$\int d^3\mathbf{x} T^{00} = \int \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \varphi)} \partial^0 \varphi - \mathcal{L} \right) = \int d^3\mathbf{x} \mathcal{H} = H, \quad (3.6.34)$$

so this is an expression of energy conservation. This shouldn’t be surprising as translations by a^0 correspond to time translations.

Similarly the conserved charges associated with the $\nu = i$ components are

$$P^i = \int d^3\mathbf{x} T^{0i} = \int d^3\mathbf{x} \pi \partial^i \varphi. \quad (3.6.35)$$

Since we expect these three components to form a four-vector with time component given by H we interpret them as the actual momentum of the particle, as opposed to the conjugate momentum, π . We can then interpret $\pi \partial_i \varphi$ as the momentum density.

¹Strictly this is the energy-momentum density tensor, but very rarely does anyone bother to make that distinction.

Four

Quantised Fields

Quantum field theory is a quantum theory of fields. We've seen classical fields, now we need some quantum mechanics before we can start quantising fields. So, in this section we'll give a quick recap of quantum mechanics. For more details see the *Principles of Quantum Mechanics* and *Quantum Theory* courses.

4.1 Quantum Mechanics

It's slightly disturbing that people believe it.

Richard Ball, on quantum mechanics

4.1.1 The Basics

In quantum mechanics we start with coordinates q_i , and a Lagrangian, $L(q_i, \dot{q}_i)$. We then define canonical momenta,

$$p_i := \frac{\partial L}{\partial \dot{q}_i}. \quad (4.1.1)$$

From this we can find a Hamiltonian,

$$H = \sum_i p \dot{q}_i - L. \quad (4.1.2)$$

So far this could all be classical. The quantum mechanics begins when we interpret q_i , p_i , and H as **Hermitian operators**¹ acting on a Hilbert space of states. The Hermitian requirement is so that these operators have real eigenvalues, which we can then interpret as the results of measurements. For example, the average energy is given by

$$E = \langle \psi, t | H | \psi, t \rangle, \quad (4.1.3)$$

where $|\psi, t\rangle$ is the state of the particle, that is a vector in the Hilbert space, and $\langle \psi, t | = |\psi, t\rangle^\dagger$ is the corresponding vector in the dual space.

The next step is to realise that operators don't necessarily commute, and so we impose the **canonical commutation relation (CCR)**:

$$[q_i, p_i] := i\delta_{ij}, \quad (4.1.4)$$

¹An operator, A , is Hermitian if $A^\dagger = A$, where A^\dagger is the **Hermitian conjugate**, defined such that $\langle \psi | A | \varphi \rangle = \langle \varphi | A^\dagger | \psi \rangle^*$.

where $[A, B] := AB - BA$ is the **commutator**. As well as this we impose that the positions commute, and the momenta commute, so

$$[q_i, q_j] = [p_i, p_j] = 0. \quad (4.1.5)$$

4.1.2 Time Dependence

The time evolution of a state, $|\psi, t\rangle$, under a system with Hamiltonian H is given by the **Schrödinger equation**

$$i \frac{d}{dt} |\psi, t\rangle = H |\psi, t\rangle. \quad (4.1.6)$$

This is a fairly simple differential equation, if we ignore the fact that H is an operator and $|\psi, t\rangle$ is a vector, the solution is just

$$|\psi, t\rangle = e^{-iHt} |\psi\rangle \quad (4.1.7)$$

where $|\psi\rangle = |\psi, 0\rangle$ is the initial state of the system. Fortunately this is perfectly valid, even though H is an operator and $|\psi, t\rangle$ is vector, we just have to interpret the exponential through its power series:

$$e^{-iHt} := \sum_{n=0}^{\infty} \frac{1}{n!} (-iHt)^n. \quad (4.1.8)$$

This scheme with which we have worked so far, where operators are time independent and states are time dependent, is called the **Schrödinger picture**. It turns out that its actually easier to do QFT by interpreting operators to be time dependent and states to be time independent, called the **Heisenberg picture**. We distinguish between these two pictures by either including t or not, as appropriate in our operators and states. A state, $|\psi\rangle$, in the Heisenberg picture is related to the state $|\psi, t\rangle$ in the Schrödinger picture by

$$|\psi\rangle = e^{iHt} |\psi, t\rangle, \quad (4.1.9)$$

which is just what we get rearranging [Equation \(4.1.7\)](#). An operator, $A(t)$, in the Heisenberg picture is related to the operator A in the Schrödinger picture by

$$A(t) = e^{iHt} A e^{-iHt}. \quad (4.1.10)$$

Importantly, expectation values are the same in both pictures, and so both pictures describe the same physics:

$$\langle A \rangle_S = \langle \psi, t | A | \psi, t \rangle \quad (4.1.11)$$

$$= \langle \psi | e^{iHt} A e^{-iHt} \quad (4.1.12)$$

$$= \langle \psi | A(t) | \psi \rangle \quad (4.1.13)$$

$$= \langle A \rangle_H, \quad (4.1.14)$$

where the subscripts refer to the picture in which we interpret the expectation value.

This picture changing is the same for all operators, so in particular, the position and momentum in the Heisenberg picture are given by

$$q_i(t) = e^{iHt} q_i e^{-iHt}, \quad \text{and} \quad p_i(t) = e^{iHt} p_i e^{-iHt}. \quad (4.1.15)$$

We also have

$$H(t) = e^{iHt} H e^{-iHt} = e^{iHt} e^{-iHt} H = H, \quad (4.1.16)$$

which works since H commutes with itself, and the exponentials are just power series in H , so also commute with H . This shows that the Hamiltonian is time independent in both pictures, which is a statement of energy conservation.

Consider two operators, $A(t)$ and $B(t)$, in the Heisenberg picture. If we know their commutator, $[A, B]$, in the Schrödinger picture then we can also compute it in the Heisenberg picture:

$$[A(t), B(t)] = \langle e^{iHt} A e^{-iHt} | e^{iHt} B e^{-iHt} \rangle \quad (4.1.17)$$

$$= e^{iHt} A e^{-iHt} e^{iHt} B e^{-iHt} - e^{iHt} B e^{-iHt} e^{iHt} A e^{-iHt} \quad (4.1.18)$$

$$= e^{iHt} A B e^{-iHt} - e^{iHt} B A e^{-iHt} \quad (4.1.19)$$

$$= e^{iHt} (AB - BA) e^{-iHt} \quad (4.1.20)$$

$$= e^{iHt} [A, B] e^{-iHt}. \quad (4.1.21)$$

In particular,

$$[q_i(t), p_j(t)] = e^{iHt} [q_i, p_j] e^{-iHt} = e^{iHt} i\delta_{ij} e^{-iHt} = i\delta_{ij} e^{iHt} e^{-iHt} = i\delta_{ij}, \quad (4.1.22)$$

so the canonical commutation relations hold if both operators are evaluated at the same time. Similarly,

$$[q_i(t), q_j(t)] = [p_i(t), p_j(t)] = 0. \quad (4.1.23)$$

In the Heisenberg picture instead of the Schrödinger equation we have the **Heisenberg equation**, which we can derive by considering the time derivative of a generic operator, $A(t)$:

$$\frac{d}{dt} A(t) = \frac{d}{dt} (e^{iHt} A e^{-iHt}) \quad (4.1.24)$$

$$= iH e^{iHt} A e^{-iHt} + e^{iHt} A (-iH) e^{-iHt} \quad (4.1.25)$$

$$= i e^{iHt} H A e^{-iHt} - i e^{iHt} A H e^{-iHt} \quad (4.1.26)$$

$$= i e^{iHt} [H, A] e^{-iHt} \quad (4.1.27)$$

$$= i[H, A(t)]. \quad (4.1.28)$$

That is,

$$\frac{d}{dt} A(t) = i[H, A(t)]. \quad (4.1.29)$$

The Heisenberg equation applied to the position and momentum gives

$$\frac{d}{dt} q(t) = i[H, q(t)], \quad \text{and} \quad \frac{d}{dt} p(t) = i[H, p(t)]. \quad (4.1.30)$$

Note the similarity to [Equation \(3.2.11\)](#). In general when we work in the Heisenberg picture things can look pretty similar to classical mechanics, but with i times the commutator in place of Poisson brackets.

Example 4.1.31 Consider the Lagrangian for a particle of mass m in a position dependent potential, V , in one dimension:

$$L = \frac{1}{2}m\dot{q}^2 - V(q). \quad (4.1.32)$$

The Hamiltonian for this system is

$$H = \frac{p^2}{2m} + V(q). \quad (4.1.33)$$

We then have

$$\dot{q} = i[H, q(t)] = \frac{i}{2m}[p^2, q]. \quad (4.1.34)$$

Now, consider three operators, A , B , and C . We have

$$[AB, C] = ABC - CBA \quad (4.1.35)$$

$$= ABC - ACB + ACB - CBA \quad (4.1.36)$$

$$= A[B, C] + [A, C]B. \quad (4.1.37)$$

Using this with $A = B = p$ and $C = q$ gives

$$[p^2, q] = p[p, q] + [p, q]p = -2ip, \quad (4.1.38)$$

where the negative comes from the antisymmetry of the commutator, $[p, q] = -[q, p] = -i$. Hence,

$$\dot{q} = \frac{p}{2m}. \quad (4.1.39)$$

Classically we would have $p = mv = m\dot{q}$, and this is just the quantum analogue of the classical nonrelativistic momentum.

We also have

$$\dot{p} = i[H, p(t)] = i[V(q), p]. \quad (4.1.40)$$

We can expand $V(q)$ as a power series in q , we also have the identity

$$[q^n, p] = inq^{n-1}. \quad (4.1.41)$$

This can be proven with induction on n . First, take $n = 0$, then

$$[q^0, p] = [\mathbb{1}, p] = 0. \quad (4.1.42)$$

Now suppose that

$$[q^k, p] = ikq^{k-1} \quad (4.1.43)$$

for some nonnegative integer k . Then

$$\begin{aligned} [q^{k+1}, p] &= [qq^k, p] = q[q^k, p] + [q, p]q^k \\ &= qikq^{k-1} + iq^k = i(k+1)q^k, \end{aligned} \quad (4.1.44)$$

and so by induction the identity holds for all natural numbers, n .
Expanding V as a Taylor series we have

$$V(q) = \sum_{n=0}^{\infty} \frac{q^n}{n!} \left. \frac{\partial^n V}{\partial q^n} \right|_{q=0}. \quad (4.1.45)$$

Then we have

$$[V(q), p] = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{\partial^n V}{\partial q^n} \right|_{q=0} [q^n, p] \quad (4.1.46)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{\partial^n V}{\partial q^n} \right|_{q=0} i n q^{n-1} \quad (4.1.47)$$

$$= \sum_{n=1}^{\infty} \frac{i q^{n-1}}{(n-1)!} \left. \frac{\partial^n V}{\partial q^n} \right|_{q=0}. \quad (4.1.48)$$

To first order we then have

$$[V(q), p] = i \left. \frac{\partial V}{\partial q} \right|_{q=0} \quad (4.1.49)$$

and so

$$\dot{p} = i[V(q), p] = -\frac{\partial V}{\partial q}. \quad (4.1.50)$$

Compare this to the classical nonrelativistic equation

$$F = -\frac{\partial V}{\partial x}, \quad (4.1.51)$$

expressing the force due to a potential.

Both of these two analogies, known as **Ehrenfest's theorem**, between classical mechanics and quantum mechanics are really due to the similarity between the classical and quantum equations of motion in the Heisenberg picture.

4.2 Quantum Fields

Fundamentally silly

Richard Ball

The procedure for quantising fields is almost identical to the procedure for quantising position and momentum. Instead of the position and momentum we have the fields φ and π . We interpret these as operators in the Heisenberg picture. Since in field theory we deal with densities instead to be integrated over, such as \mathcal{L} instead of L , we replace δ_{ij} in the canonical commutation relations with $\delta^3(\mathbf{x}-\mathbf{x}')$, note that this treats time differently to space. We postulate that the at some time,

t , we have

$$[\varphi(t, \mathbf{x}), \pi(t, \mathbf{x}')] = i\delta^3(\mathbf{x} - \mathbf{x}') \quad (4.2.1)$$

and the fields commute with themselves at all positions:

$$[\varphi(t, \mathbf{x}), \varphi(t, \mathbf{x}')] = [\pi(t, \mathbf{x}), \pi(t, \mathbf{x}')] = 0. \quad (4.2.2)$$

This is called the **equal time commutation relation (ETCR)**.

In quantum mechanics as done in the previous section we treat t as a parameter and \mathbf{x} as an operator. This is clearly not Lorentz covariant. There are also other problems, such as having only a single \mathbf{x} operator which only allows us to talk of a single particle. In quantum field theory we interpret both t and \mathbf{x} as parameters and φ as an operator. We can then make this Lorentz covariant and treat systems with multiple particles.

4.2.1 Equations of Motion

The Heisenberg equation,

$$\dot{A}(t, \mathbf{x}) = i[H, A(t, \mathbf{x})], \quad (4.2.3)$$

applies to any operator, A , and hence applies to φ . We then have

$$\dot{\varphi}(t, \mathbf{x}) = i[H, \varphi(t, \mathbf{x})] = i \int d^3\mathbf{x}' [\mathcal{H}(t, \mathbf{x}'), \varphi(t, \mathbf{x})] \quad (4.2.4)$$

We've replaced the Hamiltonian with an integral over the Hamiltonian density. We'll now introduce a shorthand notation where primes denote quantities evaluated at (t, \mathbf{x}') and an absence of a prime is a quantity evaluated at (t, \mathbf{x}) , so $\varphi' = \varphi(t, \mathbf{x}')$ and $\varphi = \varphi(t, \mathbf{x})$. The Hamiltonian density is

$$\mathcal{H}' = \frac{1}{2}(\pi'^2 + (\nabla' \varphi')^2 + m^2 \varphi'^2). \quad (4.2.5)$$

The commutator then becomes

$$[\mathcal{H}', \varphi] = \frac{1}{2}([\pi'^2, \varphi] + [(\nabla' \varphi')^2, \varphi] + m^2[\varphi'^2, \varphi]). \quad (4.2.6)$$

The last two terms vanish, since the fields commute with themselves at all positions. The first term isn't too bad if we use

$$[AB, C] = A[B, C] + [A, C]B, \quad (4.2.7)$$

we have

$$\frac{1}{2}[\pi'^2, \varphi] = \frac{1}{2}\pi'[\pi', \varphi] + \frac{1}{2}[\pi', \varphi]\pi' \quad (4.2.8)$$

$$= -\frac{1}{2}\pi'[\varphi, \pi'] - \frac{1}{2}[\varphi, \pi']\pi' \quad (4.2.9)$$

$$= -\frac{1}{2}\pi' i\delta^3(\mathbf{x} - \mathbf{x}') - \frac{1}{2}i\delta^3(\mathbf{x} - \mathbf{x}')\pi' \quad (4.2.10)$$

$$= -i\pi'\delta^3(\mathbf{x} - \mathbf{x}'). \quad (4.2.11)$$

Hence, we have the equation of motion

$$\dot{\phi}(t, \mathbf{x}) = i \int d^3 \mathbf{x}' [\mathcal{H}(t, \mathbf{x}'), \phi(t, \mathbf{x})] \quad (4.2.12)$$

$$= i \int d^3 \mathbf{x}' (-i\pi(t, \mathbf{x}') \delta^3(\mathbf{x} - \mathbf{x}')) \quad (4.2.13)$$

$$= \pi(t, \mathbf{x}). \quad (4.2.14)$$

So, $\dot{\phi} = \pi$. This should be no surprise since we found this same relationship for ϕ satisfying the Klein–Gordon equation before we started quantising.

Similarly we can find the equation of motion for π :

$$\dot{\pi}(t, \mathbf{x}) = i[H, \pi(t, \mathbf{x})] = i \int d^3 \mathbf{x}' [\mathcal{H}(t, \mathbf{x}'), \pi(t, \mathbf{x})]. \quad (4.2.15)$$

The commutator in this case is

$$[\mathcal{H}', \pi] = \frac{1}{2}([\pi'^2, \pi] + [(\nabla' \phi')^2, \pi] + m^2[\phi'^2, \pi]). \quad (4.2.16)$$

The first term vanishes this time. The last term can be computed easily:

$$\frac{1}{2}m^2[\phi'^2, \pi] = \frac{1}{2}m^2\phi'^2[\phi', \pi] + \frac{1}{2}m^2[\phi', \pi]\phi' \quad (4.2.17)$$

$$= \frac{1}{2}m^2 i \delta^3(\mathbf{x} - \mathbf{x}') + \frac{1}{2}m^2 \phi' i \delta^3(\mathbf{x} - \mathbf{x}') \quad (4.2.18)$$

$$= i m^2 \delta^3(\mathbf{x} - \mathbf{x}'). \quad (4.2.19)$$

The second term isn't too bad, we just have to notice that ∇' acts on \mathbf{x}' and not on \mathbf{x} , allowing us to pull it outside of commutators like $[\nabla' \phi', \pi] = \nabla'[\phi', \pi]$. Proceeding as before we then have

$$\frac{1}{2}[(\nabla' \phi')^2, \pi] = \frac{1}{2}(\nabla' \phi') \cdot [\nabla' \phi', \pi] + \frac{1}{2}[\nabla' \phi', \pi] \cdot (\nabla' \phi') \quad (4.2.20)$$

$$= \frac{1}{2}(\nabla' \phi') \cdot (\nabla'[\phi', \pi]) + \frac{1}{2}(\nabla'[\phi', \pi]) \cdot (\nabla' \phi') \quad (4.2.21)$$

$$= \frac{1}{2}(\nabla' \phi') \cdot (\nabla' i \delta^3(\mathbf{x} - \mathbf{x}')) + \frac{1}{2}(\nabla' i \delta^3(\mathbf{x} - \mathbf{x}')) \cdot (\nabla' \phi') \\ = i(\nabla' \phi') \cdot (\nabla' \delta^3(\mathbf{x} - \mathbf{x}')). \quad (4.2.22)$$

We can deal with the gradient of the delta distribution by integrating by parts, and assuming that fields vanish sufficiently quickly at infinity:

$$\int d^3 \mathbf{x}' (\nabla' \phi) \cdot (\nabla' \delta^3(\mathbf{x} - \mathbf{x}')) = \int d^3 \mathbf{x}' \nabla' \cdot [\delta^3(\mathbf{x} - \mathbf{x}') (\nabla' \phi')] \quad (4.2.23)$$

$$= - \int d^3 \mathbf{x}' (\nabla'^2 \phi') \delta^3(\mathbf{x} - \mathbf{x}') \quad (4.2.24)$$

$$= \int_{\partial V} [\delta^3(\mathbf{x} - \mathbf{x}') \nabla' \phi'] \quad (4.2.25)$$

$$= - \int d^3 \mathbf{x}' (\nabla'^2 \phi') \delta^3(\mathbf{x} - \mathbf{x}') \quad (4.2.26)$$

$$= - \int d^3 \mathbf{x}' (\nabla'^2 \phi') \delta^3(\mathbf{x} - \mathbf{x}') \quad (4.2.27)$$

where as usual the boundary term vanishes. This means that, as distributions,

$$i(\nabla' \varphi') \cdot (\nabla' \delta^3(\mathbf{x} - \mathbf{x}')) = -i(\nabla'^2 \varphi') \delta^3(\mathbf{x} - \mathbf{x}'). \quad (4.2.28)$$

Hence, we have

$$\dot{\pi}(t, \mathbf{x}) = i \int d^3 \mathbf{x}' [\mathcal{H}(t, \mathbf{x}'), \varphi(t, \mathbf{x})] \quad (4.2.29)$$

$$= i \int d^3 \mathbf{x}' [-i(\nabla'^2 \varphi') \delta^3(\mathbf{x} - \mathbf{x}') + im^2 \delta^3(\mathbf{x} - \mathbf{x}')] \quad (4.2.30)$$

$$= \nabla^2 \varphi - m^2 \varphi. \quad (4.2.31)$$

Now, we have $\pi = \dot{\varphi}$, so $\dot{\pi} = \ddot{\varphi} = \partial_0 \partial^0 \varphi$. Hence,

$$\partial_0 \partial^0 \varphi = \nabla^2 \varphi - m^2 \varphi = \partial_i \partial^i \varphi - m^2 \varphi. \quad (4.2.32)$$

Rearranging this we get

$$0 = (\partial_0 \partial^0 - \partial_i \partial^i) \varphi + m^2 \varphi = (\partial^2 + m^2) \varphi, \quad (4.2.33)$$

so the quantised fields satisfy the Klein–Gordon equation. This means that the Heisenberg equations for field operators look a lot like classical field equations with operators replacing the fields and i times the commutator replacing Poisson brackets.

Even though the Hamiltonian formulation and the equal time commutation relations aren't Lorentz covariant the final result is, and that's all that we care about. The steps in the derivation, however, are frame dependent. This is an inescapable problem with canonical quantisation.

Five

Quantum Harmonic Oscillator

By understanding the harmonic oscillator you understand everything! Well, except from all the things we can't solve.

Richard ball

Surely the hydrogen atom isn't a harmonic oscillator? Well it is!

Richard Ball

The quantum harmonic oscillator is one of the most important systems in quantum mechanics. It's everywhere, we use it to model atoms and phonons, molecules, and pretty much anything else. Even other systems we study in quantum mechanics, such as the hydrogen atom or pairs of particles, can be interpreted as a harmonic oscillator by a change of coordinates. After all, any potential in physics can always be Taylor expanded to second order at a minimum to get a harmonic potential. Part of the reason why we study the harmonic oscillator is that it's very nice. We can solve it exactly. The results are also nice, evenly spaced energy levels, bounded below.

The harmonic oscillator is the system with the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \quad (5.0.1)$$

We have previously remarked on the similarity of this and the Klein–Gordon Hamiltonian density, and suggested that study of the harmonic oscillator could be enlightening for the Klein–Gordon equation. In this section we'll study the harmonic oscillator, computing the spectrum of the Hamiltonian, and discussing various interpretations. In the next chapter we'll go back through the same steps but for fields obeying the Klein–Gordon and the Hamiltonian density.

5.1 Solution

There are two ways to solve the harmonic oscillator: the analytic way, involving Hermite polynomials, which is frankly, horrible, and the algebraic way, which is

maybe more mysterious but is much more inline with the quantum field theory interpretation, so this is the way we'll proceed.

Start by defining a new operators from the position and momentum operators:

$$a := \sqrt{\frac{m\omega}{2}}q + \frac{i}{\sqrt{2m\omega}}p. \quad (5.1.1)$$

The factors are chosen so that our final statements have a simple form, note that a is dimensionless. The Hermitian conjugate of this is

$$a^\dagger = \sqrt{\frac{m\omega}{2}}q - \frac{i}{\sqrt{2m\omega}}p. \quad (5.1.2)$$

Note that a is not Hermitian, it's not an observable, for now we can just think of it as a mathematical trick. We'll see later that combinations of a and a^\dagger can be treated as observables with meaningful values. The motivation for this definition is that if a and a^\dagger commuted (they don't, as we'll see soon) we could factorise the Hamiltonian as

$$\omega aa^\dagger. \quad (5.1.3)$$

A bit of algebra allows us to rewrite q and p in terms of a and a^\dagger :

$$q = \frac{1}{2m\omega}(a + a^\dagger), \quad \text{and} \quad p = -i\sqrt{\frac{m\omega}{2}}(a - a^\dagger). \quad (5.1.4)$$

We can now compute the commutation relations for a and a^\dagger from the canonical commutation relation:

$$i = [q, p] \quad (5.1.5)$$

$$= -\frac{i}{2}[a + a^\dagger, a - a^\dagger] \quad (5.1.6)$$

$$= -\frac{i}{2}([a, a] - [a, a^\dagger] + [a^\dagger, a] - [a^\dagger, a^\dagger]) \quad (5.1.7)$$

$$= i[a, a^\dagger]. \quad (5.1.8)$$

Hence, we have that

$$[a, a^\dagger] = 1. \quad (5.1.9)$$

An alternative approach to quantum mechanics is to start with a and a^\dagger , motivating them starting with the interpretation we'll see later, and then show that the canonical commutation relations are what they are as a consequence of this choice.

We can write the Hamiltonian in terms of a and a^\dagger :

$$H = -\frac{\omega}{4}(a - a^\dagger)^2 + \frac{\omega}{4}(a + a^\dagger)^2 \quad (5.1.10)$$

$$= -\frac{\omega}{4}(a^2 - aa^\dagger - a^\dagger a + (a^\dagger)^2) + \frac{\omega}{4}(a^2 + aa^\dagger + a^\dagger a + (a^\dagger)^2) \quad (5.1.11)$$

$$= \frac{\omega}{2}(aa^\dagger + a^\dagger a). \quad (5.1.12)$$

Now use

$$1 = [a, a^\dagger] = aa^\dagger - a^\dagger a \implies aa^\dagger = a^\dagger a + 1 \quad (5.1.13)$$

and we get

$$H = \frac{\omega}{2}(aa^\dagger + a^\dagger a) = \frac{\omega}{2}(a^\dagger a + 1 + a^\dagger a) = \omega\left(a^\dagger a + \frac{1}{2}\right). \quad (5.1.14)$$

To compute the spectrum of the Hamiltonian we work with the operator $N := a^\dagger a$. This is a Hermitian operator, $N^\dagger = (a^\dagger a)^\dagger = a^\dagger (a^\dagger)^\dagger = a^\dagger a = N$. We can then write the Hamiltonian as

$$H = \omega\left(N + \frac{1}{2}\right), \quad (5.1.15)$$

if you remember what the energy eigenvalues of the harmonic oscillator are you may see where this is going.

Let $|n\rangle$ be an eigenstate of N with eigenvalue n , so $N|n\rangle = n|n\rangle$. For now all we know is that $n \in \mathbb{R}$, since N has real eigenvalues as a Hermitian operator. We choose to normalise these states so that

$$\langle n|m\rangle = \delta_{nm} \quad (5.1.16)$$

The use of n and m as labels, as well as the Dirac delta hints at something we will show later, that n is an integer, but for now we just treat n as a real number.

The expectation value of N in the state $|n\rangle$ is

$$\langle N \rangle = \langle n|N|n\rangle = n\langle n|n\rangle = n. \quad (5.1.17)$$

Using $N = a^\dagger a$ we can calculate this expectation value a different way:

$$n = \langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle = \|a|n\rangle\|^2, \quad (5.1.18)$$

and since $\|a|n\rangle\|^2 \geq 0$ we have that $n \geq 0$.

Now consider commutators of N with a^\dagger and a :

$$[N, a^\dagger] = [a^\dagger a, a^\dagger] = a^\dagger [a, a^\dagger] + [a^\dagger, a^\dagger]a = a^\dagger. \quad (5.1.19)$$

Similarly,

$$[N, a] = [a^\dagger a, a] = a^\dagger [a, a] + [a^\dagger, a]a = -a. \quad (5.1.20)$$

Suppose we have an eigenstate, $|n\rangle$. What happens when we act on this with N ? Well, it turns out that we get another eigenstate:

$$Na^\dagger|n\rangle = ([N, a^\dagger] + a^\dagger N)|n\rangle \quad (5.1.21)$$

$$= (a^\dagger + a^\dagger N)|n\rangle \quad (5.1.22)$$

$$= a^\dagger|n\rangle + a^\dagger N|n\rangle \quad (5.1.23)$$

$$= a^\dagger|n\rangle + a^\dagger n|n\rangle \quad (5.1.24)$$

$$= (n+1)a^\dagger|n\rangle. \quad (5.1.25)$$

So, $a^\dagger|n\rangle$ is an eigenstate of N with eigenvalue $n + 1$. That is, $a^\dagger|n\rangle \propto |n + 1\rangle$. We can fairly easily figure out the constant of proportionality by considering $\|a^\dagger|n\rangle\|^2$:

$$\|a^\dagger|n\rangle\|^2 = \langle n|aa^\dagger|n\rangle \quad (5.1.26)$$

$$= \langle n|(a^\dagger a + [a, a^\dagger])|n\rangle \quad (5.1.27)$$

$$= \langle n|(a^\dagger a + 1)|n\rangle \quad (5.1.28)$$

$$= \langle n|a^\dagger a|n\rangle + \langle n|n\rangle \quad (5.1.29)$$

$$= \langle n|N|n\rangle + 1 \quad (5.1.30)$$

$$= n + 1. \quad (5.1.31)$$

Since the overall phase of a state has no physical significance we are free to choose the phase of the proportionality constant, so we make the sensible choice that its real and positive. That is, we choose the positive square root, $\|a^\dagger|n\rangle\| = \sqrt{n + 1}$, and we then have

$$a^\dagger|n\rangle = \sqrt{n + 1}|n + 1\rangle. \quad (5.1.32)$$

We can proceed similarly for $a|n\rangle$, we have

$$Na|n\rangle = ([N, a] + aN)|n\rangle \quad (5.1.33)$$

$$= (-a + aN)|n\rangle \quad (5.1.34)$$

$$= (n - 1)a|n\rangle. \quad (5.1.35)$$

So $a|n\rangle$ is an eigenstate of N with eigenvalue $n - 1$, so $a|n\rangle \propto |n - 1\rangle$. We can, again, work out the constant of proportionality:

$$\|a|n\rangle\|^2 = \langle n|a^\dagger a|n\rangle = \langle n|N|n\rangle = n. \quad (5.1.36)$$

So, again choosing a real, positive, solution we have

$$a|n\rangle = \begin{cases} \sqrt{n}|n - 1\rangle & n > 0, \\ 0 & n = 0, \end{cases} \quad (5.1.37)$$

the reason for the careful distinction when $n = 0$ is that the state $|0 - 1\rangle = |-1\rangle$ is not defined, as the eigenvalues of N are nonnegative. The reason we don't have to exclude, say, $n = 0.5$, which would have $|0.5 - 1\rangle = |-0.5\rangle$, is that, as we will now show, n is a nonnegative integer, that is $n \in \mathbb{N} = \{0, 1, 2, \dots\}$.

Suppose that $n \notin \mathbb{N}$. Consider some $m \in \mathbb{N} \setminus \{0\}$. Then we can apply a m times to $|n\rangle$:

$$\begin{aligned} a^m|n\rangle &= \sqrt{n}a^{m-1}|n\rangle = \sqrt{n}\sqrt{n-1}a^{m-2}|n\rangle \\ &= \dots = \sqrt{n}\sqrt{n-1}\dots\sqrt{n-m+1}|n-m\rangle. \end{aligned} \quad (5.1.38)$$

For sufficiently large m we will have $n - m < 0$. However, $n - m$ is an eigenvalue and all eigenvalues of N are nonnegative. This is a contradiction and so we must have $n \in \mathbb{N}$. In this case one of the square roots will vanish before we get to negative eigenvalues and we'll avoid the contradiction.

Since $H = \omega(N + 1/2)$ we can see that the eigenvalues of the Hamiltonian are

$$\left(n + \frac{1}{2}\right)\omega \quad \text{where} \quad n \in \mathbb{N}. \quad (5.1.39)$$

The eigenvectors of the Hamiltonian can all be written as

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle \quad (5.1.40)$$

where $|0\rangle$ is such that when a acts on it we get zero:

$$a|0\rangle := 0. \quad (5.1.41)$$

Note that $|0\rangle$ is an eigenstate with eigenvalue 0, and 0 is the zero vector, these are not the same. The factor of $1/\sqrt{n!}$ is just normalisation, if $|0\rangle$ is normalised then $a^\dagger|0\rangle = |1\rangle = \sqrt{1}|1\rangle$, then $(a^\dagger)^2|0\rangle = \sqrt{1}a^\dagger|1\rangle = \sqrt{1}\sqrt{2}|2\rangle$, and so on, so we can get rid of all these factors with $1/\sqrt{n!}$, and we're left with the normalised states $|n\rangle$.

The most important thing about the harmonic oscillator is that the energy eigenvalues are equally spaced. This makes harmonic oscillators very nice. Another nice property is that the energy eigenvalues are nonnegative. This means that the energy is bounded below, which is good as if it wasn't then systems would quickly decay to having negative infinite energies. The equally spaced eigenvalues is what makes all of quantum field theory possible.

5.2 Interpretations

It's fine, but in a sense, it's just wrong.

Richard Ball

There are two ways to interpret the harmonic oscillator. Historically the first we discuss here was Schrödinger's interpretation, and actually came a year after the second interpretation we'll discuss, which is due to Heisenberg. However, for a long time Schrödinger's interpretation was preferred, mostly because it uses more familiar mathematics, like differential equations, and also because it aligns more similarly with the notion of a classical harmonic oscillator.

5.2.1 Wave Function Interpretation

This is the way you were taught the harmonic oscillator in kindergarten I guess.

Richard Ball

The quantum harmonic oscillator is a quantised classical harmonic oscillator. We take the normal harmonic oscillator Hamiltonian, replace the position and momentum with operators, and solve the equations of motion. The picture that people then have in their head is a harmonic potential with equally spaced energy levels, as seen in [Figure 5.1](#).

The physical variables are q and p , with $q = x$ the position operator, and p chosen to satisfy the canonical commutation relations, which it can be shown is the case if $p = -id/dx$. We solve the equation $a|0\rangle = 0$ to find the vacuum state $|0\rangle$, and it's wave function $\psi_0(x) = \langle x|0\rangle$:

$$\psi_0(x) = \sqrt{\frac{m\omega}{2\pi}} \exp\left[-\frac{1}{2}m\omega x^2\right]. \quad (5.2.1)$$

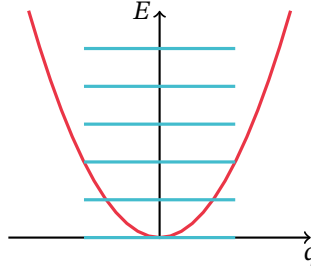


Figure 5.1: The harmonic potential and the evenly spaced energy levels it produces.

We then find the wave equation for the n th excited state, $\sqrt{n!}\psi_n(x) = \langle x|(a^\dagger)^n|0\rangle$:

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \left(0 \frac{i}{\sqrt{m\omega}} \frac{d}{dx} + i\sqrt{\frac{m\omega}{2}} x \right)^n \psi_0(x) \quad (5.2.2)$$

where the expression in brackets is just $a^\dagger = i/\sqrt{2m\omega}p + \sqrt{m\omega/2}q$ with the operators substituted in. After working through lots of derivatives of Gaussians we will find that the solutions are Gaussians times a Hermite polynomial, H_n :

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi} \right)^{1/4} \exp \left[-\frac{1}{2} m\omega x^2 \right] H_n(\sqrt{m\omega}x). \quad (5.2.3)$$

This interpretation is called **first quantisation**, presumably because it is the first way that most people are taught to interpret the harmonic oscillator.

5.2.2 Quanta

The energy eigenvalues of the harmonic oscillator are evenly spaced. This allows us to interpret the state $|n\rangle$ as consisting of n identical quanta, all of the same energy, ω . The state $|0\rangle$ is then the vacuum state, where there are no quanta, and we have the vacuum energy $\omega/2$.

Rather than try to interpret what a harmonic oscillator is we focus on what it does. Think of it as a black box which can absorb a quanta of energy, so it has $n+1$ quanta, or it can emit a quanta, so it has $n-1$ quanta. Clearly these occurrences align with the action of $|a\rangle^\dagger$ and $|a\rangle$ respectively. This leads to us interpreting a^\dagger as a **creation operator**, producing a quanta of energy in the black box, and a as a **annihilation operator**, destroying a quanta of energy in the black box. Of course this energy must come from somewhere and go somewhere, but just within the black box it may as well be appearing and vanishing.

This interpretation, while a bit odd at first, is actually closer to what we do in experiments. We scatter particles, essentially adding them to the black box, and see what comes out, that is, what leaves the black box. This interpretation also works well when considering atomic spectra, which occur in much the same way.

This interpretation is intrinsically multiparticle, which is good for quantum field theory. This interpretation is called **second quantisation**, presumably because it is the second way that most people are taught to interpret the harmonic oscillator.

5.3 Harmonic Oscillator with Time Dependence

If we include time dependence in the Harmonic oscillator the, working in the Heisenberg picture, this corresponds to replacing q and p with time dependent operators, which in term means introducing a time dependence to a and a^\dagger . This time dependence doesn't change the first step of solving the Harmonic oscillator, which is finding q and p in terms of a and a^\dagger , which gives us

$$q(t) = \frac{1}{\sqrt{2m\omega}}[a(t) + a^\dagger(t)], \quad (5.3.1)$$

$$p(t) = i\sqrt{\frac{m\omega}{2}}[a(t) - a^\dagger(t)]. \quad (5.3.2)$$

The commutation relations for $a(t)$, $a^\dagger(t)$, and H are $[a(t), a^\dagger(t)] = 1$, $[H, a(t)] = -\omega a(t)$, and $[H, a^\dagger(t)] = \omega a^\dagger(t)$, which are the same as if we replace H with N , since $H = \omega(N + 1/2)$. The time dependence doesn't change the commutators, which we can check quite easily:

$$[a(t), a^\dagger(t)] = [e^{-iHt} a e^{iHt}, e^{-iHt} a^\dagger e^{iHt}] \quad (5.3.3)$$

$$= e^{-iHt} a e^{iHt} e^{-iHt} a^\dagger e^{iHt} - e^{-iHt} a^\dagger e^{iHt} e^{-iHt} a e^{iHt} \quad (5.3.4)$$

$$= e^{-iHt} a a^\dagger e^{iHt} - e^{-iHt} a^\dagger a e^{iHt} \quad (5.3.5)$$

$$= e^{-iHt} [a, a^\dagger] e^{iHt} \quad (5.3.6)$$

$$= e^{-iHt} 1 e^{iHt} \quad (5.3.7)$$

$$= 1. \quad (5.3.8)$$

Likewise, the commutators with the Hamiltonian are unchanged:

$$[H, a^\dagger(t)] = [H, e^{-iHt} a^\dagger e^{iHt}] \quad (5.3.9)$$

$$= H e^{-iHt} a^\dagger e^{iHt} - e^{-iHt} a^\dagger e^{iHt} H \quad (5.3.10)$$

$$= e^{-iHt} H a^\dagger e^{iHt} - e^{-iHt} a^\dagger H e^{iHt} \quad (5.3.11)$$

$$= e^{-iHt} [H, a^\dagger] e^{iHt} \quad (5.3.12)$$

$$= e^{-iHt} \omega a^\dagger e^{iHt} \quad (5.3.13)$$

$$= \omega a^\dagger(t). \quad (5.3.14)$$

An similarly for $[H, a(t)]$.

We can then solve the Heisenberg equation for the creation and annihilation operators:

$$\dot{a}(t) = i[H, a(t)] = -i\omega a(t). \quad (5.3.15)$$

This very simple differential equation, $\dot{a}(t) = -i\omega a(t)$, has a plane wave solution,

$$a(t) = a e^{-i\omega t}, \quad (5.3.16)$$

where $a = a(0)$ is the annihilation operator in the Schrödinger picture. Taking the Hermitian conjugate of this we get

$$\dot{a}^\dagger(t) = a^\dagger e^{i\omega t}. \quad (5.3.17)$$

The position and momentum operators are then

$$q(t) = \frac{1}{\sqrt{2m\omega}}[ae^{-i\omega t} + a^\dagger e^{i\omega t}], \quad (5.3.18)$$

$$p(t) = i\sqrt{\frac{m\omega}{2}}[ae^{-i\omega t} - a^\dagger e^{i\omega t}]. \quad (5.3.19)$$

Six

Mode Expansions

6.1 Expanding the Field

The field $\varphi(x)$ satisfies the Klein–Gordon equations. We’ve seen that this allows plane wave solutions, $e^{\pm ip \cdot x}$. This suggests that we can write a general solution as a superposition of wave solutions, which is just a Fourier transform. Since φ is a function of x this is really an inverse Fourier transform. For comparison consider the inverse transform of a single variable, real, one-dimensional function, f :

$$f(x) = \int \frac{dp}{2\pi} (\tilde{f}(p)e^{-ipx} + \tilde{f}^*(p)e^{ipx}). \quad (6.1.1)$$

Here $\tilde{f}(p)$ is the Fourier transformed function, which acts as a Fourier coefficient here. We need both the e^{ipx} and conjugate e^{-ipx} terms to get a real function. The factor of 2π is part of the normalisation, but exactly where it appears between the forward and inverse is just convention.

We want our field to have on-shell momentum, that is $p^2 = m^2$. To enforce this we include a factor of $2\pi\delta(p^2 - m^2)$, where the 2π is factor is from the integral representation of the Dirac delta, as the Fourier transform of one:

$$\int dx e^{-ip \cdot x} = 2\pi\delta(p). \quad (6.1.2)$$

We also want to restrict ourselves to positive energy solutions, so $p_0 > 0$, we can enforce this with a factor of $\theta(p_0)$, where θ is the **Heaviside step function** defined by

$$\theta(u) = \begin{cases} 1 & u > 0, \\ 1/2 & u = 0, \\ 0 & u < 0. \end{cases} \quad (6.1.3)$$

The **mode expansion** of φ is then

$$\varphi(x) = \int \frac{d^4 p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\theta(p_0)[a(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (6.1.4)$$

Here $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$ are Fourier coefficients, we’ll see later that they are the quantum field theory analogues of the creation and annihilation operators for the harmonic oscillator. They depend only on \mathbf{p} , rather than the four-vector p , since we can take $\delta(p^2 - m^2)$ as fixing the value of p_0 .

Define $\omega(\mathbf{p}) := +\sqrt{\mathbf{p}^2 + m^2}$, which is always nonnegative. This allows us to rewrite the argument of the Dirac delta:

$$(p_0 - \omega(\mathbf{p}))(p_0 + \omega(\mathbf{p})) = p_0 p_0 - \omega(\mathbf{p})^2 = p_0 p_0 - \mathbf{p}^2 - m^2 = p^2 - m^2. \quad (6.1.5)$$

So, now we have $\delta(\omega(\mathbf{p}))$ in our integral. Recall the distribution identity

$$\delta(f(x)) = \sum_i \frac{\delta(x - a_i)}{|f'(a_i)|} \quad (6.1.6)$$

where a_i are all the zeros of f lying in the range of integration. Since in the integral we are enforcing $p_0 \geq 0$ the only time $p^2 - m^2 = (p_0 - \omega(\mathbf{p}))(p_0 + \omega(\mathbf{p}))$ vanishes is when $p_0 = \omega(\mathbf{p})$, and

$$\left. \frac{\partial}{\partial p} (p^2 - m^2) \right|_{p_0=\omega(\mathbf{p})} = 2\omega(\mathbf{p}) \quad (6.1.7)$$

so, as distributions, we have

$$\delta(p^2 - m^2)\theta(p_0) = \frac{1}{2\omega(\mathbf{p})}\delta(p^0 - \omega(\mathbf{p})). \quad (6.1.8)$$

This lets us write $\varphi(x)$ as

$$\varphi(x) = \int \frac{d^4 p}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} \delta(p^0 - \omega(\mathbf{p})) [a(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (6.1.9)$$

Performing the p^0 integral using the sifting property of the Dirac delta we have

$$\varphi(x) = \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} [a(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}] \Big|_{p_0=\omega(\mathbf{p})}. \quad (6.1.10)$$

At this point we make some remarks on this expansion:

- The expansion is manifestly Lorentz invariant, while the presence of p^0 may be concerning orthochronous Lorentz transformations (the only type we consider in this course) don't change the sign of p^0 , and so $\theta(p^0)$ is Lorentz invariant.
- Since $p^0 = \omega(\mathbf{p})$ the energy is always positive, and so is bounded below.
- The integral over $\delta(p^0 - \omega(\mathbf{p}))$ fixes $p^0 = \omega(\mathbf{p})$, and from now on this will be implicit in the integrals, instead of writing $|_{p_0=\omega(\mathbf{p})}$.
- The operator $\varphi(x)$ is in the Heisenberg picture, but the time dependence comes from the exponential factors, both $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$ are time independent operators.
- We need both the a and a^\dagger terms for φ to be Hermitian.
- The measure

$$\frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} = \frac{d^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2) \theta(p^0) \quad (6.1.11)$$

is Lorentz invariant, but takes a lot of writing, we will use the shorthand notation

$$d\mathbf{p} = \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} = \int \frac{d^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2) \theta(p^0). \quad (6.1.12)$$

Now that we've expanded the field φ it is easy to expand the conjugate field, π , using $\pi = \dot{\varphi}$:

$$\pi(x) = \dot{\varphi}(x) = -\frac{i}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} [a(\mathbf{p})e^{-ip \cdot x} - a^\dagger(\mathbf{p})e^{ip \cdot x}] \quad (6.1.13)$$

$$= -\frac{i}{2} \int d^3 p [a(\mathbf{p})e^{-ip \cdot x} - a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (6.1.14)$$

Here we've used

$$\begin{aligned} \frac{\partial}{\partial x^0} e^{\pm ip \cdot x} &= \frac{\partial}{\partial x^0} \exp[\pm i p^0 x_0 \mp p^i x_i] \\ &= \pm i p^0 \exp[\pm i p^0 x_0 \mp p^i x_i] = \pm i p^0 \exp[\pm i p \cdot x] \end{aligned} \quad (6.1.15)$$

and then set $p^0 = \omega(\mathbf{p})$.

6.2 The Fourier Coefficients

As with the Harmonic oscillator we now want to invert these relations to find expressions for $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$. Since these are inverse Fourier transformations the inversion happens through Fourier transformation. For a single variable, real, one-dimensional function, f , the Fourier transform is

$$\tilde{f}(p) = \int dx f(x) e^{ip \cdot x}. \quad (6.2.1)$$

Using this, but in three dimensions, we have

$$\begin{aligned} \int d^3 \mathbf{x} e^{ip' \cdot x} \varphi(x) &= \int d^3 p \frac{1}{2\omega(\mathbf{p})} \int d^3 \mathbf{x} [a(\mathbf{p})e^{i(p' - \mathbf{p}) \cdot x} + a^\dagger(\mathbf{p})e^{i(\mathbf{p}' + \mathbf{p}) \cdot x}] \\ &= \int d^3 p (2\pi)^3 [a(\mathbf{p})\delta^3(\mathbf{p}' - \mathbf{p})e^{i(p'^0 - p^0)t} \\ &\quad + a^\dagger(\mathbf{p})\delta^3(\mathbf{p}' + \mathbf{p})e^{i(p'^0 + p^0)t}] \\ &= \int \frac{d^3 \mathbf{p}}{2\omega(\mathbf{p})} [a(\mathbf{p})\delta^3(\mathbf{p}' - \mathbf{p})e^{i(p'^0 - p^0)t} \\ &\quad + a^\dagger(\mathbf{p})\delta^3(\mathbf{p}' + \mathbf{p})e^{i(p'^0 + p^0)t}]. \end{aligned}$$

Here we've recognised the integral representation of the Dirac delta:

$$\int d^3 \mathbf{x} e^{-ip \cdot x} = (2\pi)^3 \delta^3(\mathbf{x}) e^{ip^0 t}, \quad (6.2.2)$$

and made the usual assumption that everything is sufficiently convergent for us to be able to swap the order of integrals. Performing the integrals now using the Dirac deltas, and the fact that $\omega(-\mathbf{p}) = \omega(\mathbf{p})$, we have

$$\int d^3 \mathbf{x} e^{ip' \cdot x} \varphi(x) = \frac{1}{2\omega(\mathbf{p}')} a(\mathbf{p}') + \frac{1}{2\omega(-\mathbf{p}')} a^\dagger(-\mathbf{p}') e^{2ip'^0 t} \quad (6.2.3)$$

$$= \frac{1}{2\omega(\mathbf{p}')} [a(\mathbf{p}') + a^\dagger(-\mathbf{p}') e^{2i\omega(\mathbf{p}')t}]. \quad (6.2.4)$$

Here we've used the Dirac delta's to set $\mathbf{p} = \pm \mathbf{p}'$ in each term as appropriate. Restricting ourselves to on-shell, positive energy, solutions this then fixes the value of p^0 to $\sqrt{\mathbf{p}'^2 + m^2} = \omega(\mathbf{p}')$.

We can do exactly the same for π , the only difference is signs and a factor of $1/\omega(\mathbf{p})$, so we get

$$\int d^3\mathbf{x} e^{ip'\cdot x} \pi(x) = -\frac{i}{2}[a(\mathbf{p}') - a(-\mathbf{p}')^\dagger e^{2i\omega(\mathbf{p}')t}]. \quad (6.2.5)$$

Adding i times this to $\omega(\mathbf{p})$ times the result for φ we get

$$a(\mathbf{p}) = \int d^3\mathbf{x} e^{ip\cdot x} [\omega(\mathbf{p})\varphi(t, \mathbf{x}) + i\pi(t, \mathbf{x})]. \quad (6.2.6)$$

Taking the conjugate we get

$$a^\dagger(\mathbf{p}) = \int d^3\mathbf{x} e^{-ip\cdot x} [\omega(\mathbf{p})\varphi(t, \mathbf{x}) - i\pi(t, \mathbf{x})]. \quad (6.2.7)$$

Now we have expressions for $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$ we can compute their commutator:

$$[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = \int d^3\mathbf{x} \int d^3\mathbf{x}' e^{ip\cdot x - ip'\cdot x'} [\omega(\mathbf{p})\varphi(t, \mathbf{x}) + i\pi(t, \mathbf{x}), \omega(\mathbf{p}')\varphi(t, \mathbf{x}') - i\pi(t, \mathbf{x}')]. \quad (6.2.8)$$

Using the equal time commutation relations only the cross terms don't vanish, and so this becomes

$$[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = \int d^3\mathbf{x} \int d^3\mathbf{x}' e^{ip\cdot x - ip'\cdot x'} (-i\omega(\mathbf{p})[\varphi(t, \mathbf{x}), \pi(t, \mathbf{x}')] + i\omega(\mathbf{p}')[\pi(t, \mathbf{x}), \varphi(t, \mathbf{x}')]) \quad (6.2.9)$$

$$= \int d^3\mathbf{x} \int d^3\mathbf{x}' e^{ip\cdot x - ip'\cdot x'} (\omega(\mathbf{p})\delta^3(\mathbf{x} - \mathbf{x}') + \omega(\mathbf{p}')\delta^3(\mathbf{x} - \mathbf{x}')) \quad (6.2.10)$$

$$= \int d^3\mathbf{x} e^{i(p-p')\cdot x} (\omega(\mathbf{p}) + \omega(\mathbf{p}')) \quad (6.2.11)$$

$$= (2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') 2\omega(\mathbf{p}) \quad (6.2.12)$$

where again we've recognised the integral representation of the Dirac delta, and made use of the resulting Dirac delta to set $\mathbf{p} = \pm \mathbf{p}'$, and hence $p^0 = p'^0$ choosing to stay on-shell and positive energy. We introduce the shorthand notation,

$$(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') 2\omega(\mathbf{p}) = \tilde{\delta}(\mathbf{p} - \mathbf{p}'). \quad (6.2.13)$$

This allows us to write

$$[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = \tilde{\delta}(\mathbf{p} - \mathbf{p}'). \quad (6.2.14)$$

Compare this to $[a, a^\dagger] = 1$ for the harmonic oscillator. We also have

$$[a(\mathbf{p}), a(\mathbf{p}')] = [a^\dagger(\mathbf{p}'), a^\dagger(\mathbf{p}')] = 0. \quad (6.2.15)$$

Note that $\tilde{\delta}(\mathbf{p} - \mathbf{p}')$ is just a number, not an operator. This notation is nice, since if we combine the invariant measure, $d\mathbf{p}$, and $\tilde{\delta}(\mathbf{p} - \mathbf{p}')$ then we can just pretend we have $d\mathbf{p}$ and $\delta^3(\mathbf{p} - \mathbf{p}')$:

$$\int d\mathbf{p} f(\mathbf{p}) \tilde{\delta}(\mathbf{p} - \mathbf{p}') = f(\mathbf{p}'). \quad (6.2.16)$$

6.3 The Hamiltonian

To complete the analogy with the harmonic oscillator we need to find the Hamiltonian in terms of $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$. Starting with the hamiltonian

$$H = \frac{1}{2} \int d^3\mathbf{x} [\pi(x)^2 + (\nabla\varphi(x))^2 + m^2\varphi(x)^2] \quad (6.3.1)$$

we can split this into three components,

$$H_1 = \frac{1}{2} \int d^3\mathbf{x} \pi(x)^2, \quad (6.3.2)$$

$$H_2 = \frac{1}{2} \int d^3\mathbf{x} (\nabla\varphi(x))^2, \quad (6.3.3)$$

$$H_3 = \frac{1}{2} \int d^3\mathbf{x} m^2\varphi(x)^2, \quad (6.3.4)$$

$$(6.3.5)$$

We'll then treat each of these separately.

Start with the first term. Substituting in the expansion for $\pi(x)$, being careful to use different integration variables for each factor of π , we get

$$H_1 = -\frac{1}{8} \int d^3\mathbf{x} \int \frac{d^3\mathbf{p}}{(2\pi)^3} [a(\mathbf{p})e^{-ip \cdot x} - a^\dagger(\mathbf{p})e^{ip \cdot x}] \times \int \frac{d^3\mathbf{p}'}{(2\pi)^3} [a(\mathbf{p}')e^{-ip' \cdot x} - a^\dagger(\mathbf{p}')e^{ip' \cdot x}]. \quad (6.3.6)$$

Expanding this all the integrand becomes

$$a(\mathbf{p})a(\mathbf{p}')e^{-i(p+p') \cdot x} - a(\mathbf{p})a^\dagger(\mathbf{p}')e^{-i(p-p') \cdot x} - a^\dagger(\mathbf{p})a(\mathbf{p}')e^{i(p-p') \cdot x} + a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')e^{i(p+p') \cdot x}. \quad (6.3.7)$$

We can perform the integration over position first, and, recalling that $p^0 = \omega(\mathbf{p})$, we get two types of terms, time dependent and time independent, these come from recognising the integral representation of the Dirac delta:

$$\int d^3\mathbf{x} e^{\pm i(p+p') \cdot x} = (2\pi)^3 \delta(\mathbf{p} + \mathbf{p}') e^{\pm 2i\omega(\mathbf{p})t}, \quad (6.3.8)$$

$$\int d^3\mathbf{x} e^{\pm i(p-p') \cdot x} = (2\pi)^3 \delta(\mathbf{p} - \mathbf{p}'). \quad (6.3.9)$$

We should expect that the time dependent terms in H will all cancel, since H is time independent.

After performing the integral over positions we are left with the following integrand

$$(2\pi)^3 a(\mathbf{p})a(\mathbf{p}')\delta(\mathbf{p} - \mathbf{p}')e^{-2i\omega(\mathbf{p})t} - (2\pi)^3 a(\mathbf{p})a^\dagger(\mathbf{p}')\delta(\mathbf{p} - \mathbf{p}') - (2\pi)^3 a^\dagger(\mathbf{p})a(\mathbf{p}')\delta(\mathbf{p} - \mathbf{p}') + (2\pi)^3 a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')\delta(\mathbf{p} + \mathbf{p}')e^{2i\omega(\mathbf{p})t}.$$

Performing the integral over \mathbf{p}' , and cancelling one factor of $(2\pi)^3$ we are left with

$$H_1 = -\frac{1}{8} \int \frac{d^3\mathbf{p}}{(2\pi)^3} [a(\mathbf{p})a(\mathbf{p})e^{-2i\omega(\mathbf{p})t} - a(\mathbf{p})a^\dagger(\mathbf{p}) - a^\dagger(\mathbf{p})a(\mathbf{p}) + a^\dagger(\mathbf{p})a^\dagger(\mathbf{p})e^{2i\omega(\mathbf{p})t}]. \quad (6.3.10)$$

Keeping only the time independent terms, since the time independent terms will cancel with terms in H_2 and H_3 , we get the contribution from the first term:

$$H_1 \rightarrow \frac{1}{8} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} [a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})]. \quad (6.3.11)$$

Doing the same for the other two terms we get their time independent contributions:

$$H_2 \rightarrow \frac{1}{8} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{\mathbf{p}^2}{\omega(\mathbf{p})^2} [a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})], \quad (6.3.12)$$

$$H_3 \rightarrow \frac{1}{8} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{m^2}{\omega(\mathbf{p})^2} [a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})]. \quad (6.3.13)$$

See [Chapter B](#) for the full calculation. Noticing that $\mathbf{p}^2 + m^2$ cancels with $1/\omega(\mathbf{p})^2$ in the full expression we get

$$H = \frac{1}{4} \int \frac{d^3 p}{(2\pi)^3} [a(\mathbf{p})a^\dagger(\mathbf{p}) + a(\mathbf{p})a^\dagger(\mathbf{p})]. \quad (6.3.14)$$

Compare this to the Harmonic oscillator result,

$$H = \frac{1}{2} \omega (aa^\dagger + a^\dagger a). \quad (6.3.15)$$

We can now compute the commutators of H with $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$:

$$[H, a^\dagger(\mathbf{p})] = \frac{1}{4} \int \frac{d^3 p'}{(2\pi)^3} [a(\mathbf{p}')a^\dagger(\mathbf{p}') + a^\dagger(\mathbf{p}')a(\mathbf{p}'), a^\dagger(\mathbf{p})] \quad (6.3.16)$$

$$\begin{aligned} &= \frac{1}{4} \int \frac{d^3 p'}{(2\pi)^3} ([a(\mathbf{p}')a^\dagger(\mathbf{p}'), a^\dagger(\mathbf{p})] + [a^\dagger(\mathbf{p}')a(\mathbf{p}'), a^\dagger(\mathbf{p})]) \\ &= \frac{1}{4} \int \frac{d^3 p'}{(2\pi)^3} (a(\mathbf{p}')[a^\dagger(\mathbf{p}'), a^\dagger(\mathbf{p})] + [a(\mathbf{p}'), a^\dagger(\mathbf{p})]a^\dagger(\mathbf{p}') \\ &\quad + a^\dagger(\mathbf{p}')[a(\mathbf{p}'), a^\dagger(\mathbf{p})] + [a^\dagger(\mathbf{p}'), a^\dagger(\mathbf{p})]) \end{aligned} \quad (6.3.17)$$

$$= \frac{1}{4} \int d^3 p' 2\omega(\mathbf{p}) [\delta(\mathbf{p} - \mathbf{p}')a^\dagger(\mathbf{p}') + a^\dagger(\mathbf{p}')\delta(\mathbf{p}' - \mathbf{p})] \quad (6.3.18)$$

$$= \omega(\mathbf{p})a^\dagger(\mathbf{p}). \quad (6.3.19)$$

Similarly, with $a(\mathbf{p})$ we get

$$[H, a(\mathbf{p})] = -\omega(\mathbf{p})a(\mathbf{p}). \quad (6.3.20)$$

Compare these to the harmonic oscillator results, $[H, a^\dagger] = \omega a^\dagger$ and $[H, a] = -\omega a$.

We can view the quantum fields as being a quantum harmonic oscillator of energy $\omega(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2}$ for every Fourier mode, \mathbf{p} .

6.4 Particle Interpretation

Consider an energy eigenstate, $|E\rangle$, with eigenvalue E , so $H|E\rangle = E|E\rangle$. Then the state $a^\dagger(\mathbf{p})|E\rangle$ is also an energy eigenstate, with energy $E + \omega(\mathbf{p})$:

$$Ha^\dagger(\mathbf{p})|E\rangle = ([H, a^\dagger(\mathbf{p})] + a^\dagger(\mathbf{p})H)|E\rangle \quad (6.4.1)$$

$$= (\omega(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})E)|E\rangle \quad (6.4.2)$$

$$= (\omega(\mathbf{p}) + E)a^\dagger(\mathbf{p})|E\rangle. \quad (6.4.3)$$

Similarly, $a(\mathbf{p})|E\rangle$ is an energy eigenstate with energy $E - \omega(\mathbf{p})$.

This suggests that, following the interpretation of the harmonic oscillator, $a^\dagger(\mathbf{p})$ is a **creation operator**, creating a quantum with energy $\omega(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2}$, and $a(\mathbf{p})$ is an **annihilation operator**, destroying a quantum with energy $\omega(\mathbf{p})$. Note that all quanta in this interpretation have positive energy. There is another interpretation, which we'll discuss later, where $a^\dagger(\mathbf{p})$ creates a quantum with energy $\omega(\mathbf{p})$ and $a(\mathbf{p})$ creates a quantum with energy $-\omega(\mathbf{p})$.

Since the Hamiltonian is nonnegative, its defined as a sum of squares of real quantities, there must be some lowest energy state, $|0\rangle$, which we call the **vacuum state**. This state is such that any annihilation operator acting on it gives the zero vector:

$$a(\mathbf{p})|0\rangle = 0 \quad \forall \mathbf{p}. \quad (6.4.4)$$

Starting with $|0\rangle$ we can act with creation operators to make new states. Each creation operator, $a^\dagger(\mathbf{p})$, makes a new particle with energy $\omega(\mathbf{p})$ and momentum \mathbf{p} . For example,

$$|\psi\rangle = a^\dagger(\mathbf{p})a^\dagger(\mathbf{q})a^\dagger(\mathbf{r})|0\rangle \quad (6.4.5)$$

is a state with three particles with total energy $\omega(\mathbf{p}) + \omega(\mathbf{q}) + \omega(\mathbf{r})$, and total momentum $\mathbf{p} + \mathbf{q} + \mathbf{r}$.

Since all $a^\dagger(\mathbf{p})$ commute the state

$$a^\dagger(\mathbf{p})a^\dagger(\mathbf{q})a^\dagger(\mathbf{r})|0\rangle \quad (6.4.6)$$

is the same as $|\psi\rangle$. We interpret this as meaning that the particles are bosons, swapping any two particles leaves the state unchanged. This means the particles follow Bose–Einstein statistics.

The space of all states which can be constructed in this way, called **Fock space**, is the space given by all complex-linear combinations of

$$a^\dagger(\mathbf{p}_1) \cdots a^\dagger(\mathbf{p}_n)|0\rangle \quad (6.4.7)$$

for any values of $n \in \mathbb{N}$.

Formally, the Fock space (for bosons) is defined to be

$$F(\mathbb{H}) = \overline{\bigoplus_{n=0}^{\infty} S\mathbb{H}^{\otimes n}} = \overline{\mathbb{C} \oplus \mathbb{H} \oplus S(\mathbb{H} \otimes \mathbb{H}) \oplus S(\mathbb{H} \otimes \mathbb{H} \otimes \mathbb{H}) \oplus \cdots} \quad (6.4.8)$$

where \mathbb{H} is the single particle Hilbert space,

$$\mathbb{H}^{\otimes n} := \bigotimes_{i=1}^n \mathbb{H} = \underbrace{\mathbb{H} \otimes \cdots \otimes \mathbb{H}}_{n \text{ times}}, \quad (6.4.9)$$

the operator S symmetrises all tensors, the direct sum, \oplus , runs from $n = 0$, for states with no particles, to infinity, with each term representing an n -particle state space, and the line represents that we complete the space to form a Hilbert space, this entails adding in the values of all absolutely convergent series.

We make the normalisation choice that $\langle 0|0\rangle = 1$, then, if $|\mathbf{p}\rangle = a^\dagger(\mathbf{p})|0\rangle$ is a single particle state we have

$$\langle \mathbf{p}|\mathbf{q}\rangle = \langle 0|a(\mathbf{p})a^\dagger(\mathbf{q})|0\rangle \quad (6.4.10)$$

$$= \langle 0|[a(\mathbf{p}), a^\dagger(\mathbf{q})]|0\rangle + \langle 0|a^\dagger(\mathbf{q})a(\mathbf{p})|0\rangle \quad (6.4.11)$$

$$= \delta(\mathbf{p} - \mathbf{q}). \quad (6.4.12)$$

Note that $a(\mathbf{p})|0\rangle$ and $\langle 0|a^\dagger(\mathbf{q}) = [a(\mathbf{q})|0\rangle]^\dagger$ both vanish, since they are annihilators acting on the vacuum state.

As with the harmonic oscillator we can define a **number density operator**, $N(\mathbf{p}) := a^\dagger(\mathbf{p})a(\mathbf{p})$, the eigenvalues of which integrate to give the number of particles in the state with momentum \mathbf{p} .

6.5 Ground State Energy

We can rewrite the Hamiltonian in terms of the number density operator:

$$H = \frac{1}{4} \int \mathfrak{d}p \, \omega(\mathbf{p}) (a^\dagger(\mathbf{p})a(\mathbf{p}) + a(\mathbf{p})a^\dagger(\mathbf{p})) \quad (6.5.1)$$

$$= \frac{1}{2} \int \mathfrak{d}p \, \omega(\mathbf{p}) \left(a^\dagger(\mathbf{p})a(\mathbf{p}) + \frac{1}{2} [a(\mathbf{p}), a^\dagger(\mathbf{p})] \right) \quad (6.5.2)$$

$$= \frac{1}{2} \int \mathfrak{d}p \, \omega(\mathbf{p}) (N(\mathbf{p}) + \omega(\mathbf{p})(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p})) \quad (6.5.3)$$

$$= \frac{1}{2} \int \mathfrak{d}p \, \omega(\mathbf{p}) (N(\mathbf{p}) + \omega(\mathbf{p})(2\pi)^3 \delta^3(\mathbf{0})). \quad (6.5.4)$$

Clearly we have $\langle 0|N(\mathbf{p})|0\rangle = \langle 0|a^\dagger(\mathbf{p})a(\mathbf{p})|0\rangle = 0$, and so

$$\langle 0|H|0\rangle = \frac{1}{2} \int \mathfrak{d}^3 \mathbf{p} \, \omega(\mathbf{p}) \delta^3(\mathbf{0}). \quad (6.5.5)$$

This is infinite. In fact, its *very* infinite. Each $\delta(0)$ is infinite, and we have three of them multiplied together, then we integrate over all momentum space, which is infinite, to make matters worse $\omega(\mathbf{p})$ grows arbitrarily large with $|\mathbf{p}|$.

The solution, as with so many similar problems in physics, is just not to think about it. All that we can measure is the energy of particles absorbed and emitted by the system. We cannot measure the ground state energy, so it doesn't matter that our theory gives a nonsensical answer when we try to compute the expected ground state energy.

6.5.1 Normal Ordering

One solution to the infinity arising in the ground state energy is to be careful about how we order $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$. If we keep all $a^\dagger(\mathbf{p})$ on the left of all $a(\mathbf{p})$ then, when we take expectation values, one of our $a(\mathbf{p})$ will act on $|0\rangle$ giving the zero vector, and one of our $a^\dagger(\mathbf{p})$ will act on $\langle 0|$, also giving the zero vector.

Mathematically what happens is that when we change to order of operators to have all $a^\dagger(\mathbf{p})$ on the left we'll get commutators of $a(\mathbf{p})$ and $a^\dagger(\mathbf{p}')$, which give rise to Dirac deltas which cancel with the existing ones, to leave a finite result.

This idea that careful ordering of operators can avoid infinities goes beyond the expectation value of the ground state energy. Consider our field, $\varphi(x)$. We can split this into parts,

$$\varphi(x) = \varphi^+(x) + \varphi^-(x) \quad (6.5.6)$$

where

$$\varphi^+(x) = \int \mathfrak{d}p \, a(\mathbf{p}) e^{-ip \cdot x}, \quad (6.5.7)$$

and

$$\varphi^-(x) = \int \mathrm{d}^3p \, a^\dagger(\mathbf{p}) e^{ip \cdot x} = (\varphi^+)^\dagger. \quad (6.5.8)$$

If these were single particle states in quantum mechanics then they would correspond to positive and negative energy states respectively, but they aren't, although we'll see an interpretation along these lines later.

Now suppose we want to consider the product $\varphi(x)\varphi(y)$ for two arbitrary space-time points, x and y . We can expand this as follows:

$$\varphi(x)\varphi(y) = [\varphi^+(x) + \varphi^-(x)][\varphi^+(y) + \varphi^-(y)] \quad (6.5.9)$$

$$= \varphi^+(x)\varphi^+(y) + \varphi^+(x)\varphi^-(y) + \varphi^-(x)\varphi^+(y) + \varphi^-(x)\varphi^-(y). \quad (6.5.10)$$

The **normal ordering** of $\varphi(x)\varphi(y)$, denoted $:\varphi(x)\varphi(y):$, is defined as the result of taking this product, but swapping terms so that all $a^\dagger(\mathbf{p})$ appear on the left of all $a(\mathbf{p}')$, that is all φ^+ appear on the left of all φ^- , so

$$:\varphi(x)\varphi(y): = \varphi^+(x)\varphi^+(y) + \varphi^+(x)\varphi^-(y) + \varphi^-(x)\varphi^+(y) + \varphi^-(x)\varphi^-(y). \quad (6.5.11)$$

Then we have

$$\langle 0 | : \varphi(x)\varphi(y) : | 0 \rangle = 0. \quad (6.5.12)$$

¹That is, X is not just a scalar, in which case $\langle 0 | : X : | 0 \rangle = X \langle 0 | 0 \rangle = X$.

In general, if X is any nontrivial¹ polynomial in $a^\dagger(\mathbf{p})$ and $a(\mathbf{p})$ we'll have

$$\langle 0 | : X : | 0 \rangle = 0. \quad (6.5.13)$$

Subtracting the zero point energy of H from H is equivalent to normal ordering

$$:H: = H - \langle 0 | H | 0 \rangle, \quad (6.5.14)$$

since by normal ordering H we swap exactly those terms which give infinity, and all others give zero, so we can think of normal ordering as subtracting them off.

From now on we'll assume that H and \mathbf{p} are normal ordered, often without writing it.

6.6 Ground State Momentum

The physical momentum operator is given by

$$\mathbf{P} = \int \mathrm{d}^3\mathbf{x} \, \pi(\mathbf{x}) \nabla \varphi(\mathbf{x}). \quad (6.6.1)$$

Recall that this comes from the T^{i0} component of the energy-momentum tensor. Expressing the fields in terms of $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$ we get

$$\mathbf{P} = \frac{1}{2} \int \mathrm{d}^3p \, \mathbf{p} (a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})) \quad (6.6.2)$$

where the factor of \mathbf{p} comes from computing $\nabla \varphi$. Rewriting this using the commutator of $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$ we get

$$\mathbf{P} = \int \mathrm{d}^3p \, \mathbf{p} (N(\mathbf{p}) + \omega(\mathbf{p})(2\pi)^3 \delta(\mathbf{0})). \quad (6.6.3)$$

Unlike with the ground state energy the $\delta(\mathbf{0})$ factor doesn't cause a problem. Integrating over all modes in the modes with momenta \mathbf{p} and $-\mathbf{p}$ cancel, since $\omega(\mathbf{p}) = \omega(-\mathbf{p})$. So, we have

$$\langle 0|\mathbf{P}|0\rangle = \langle 0|:\mathbf{P}:|0\rangle = 0, \quad (6.6.4)$$

that is, the vacuum state has no momentum, which really ought to be the case.

Seven

Covariant Commutators

7.1 What are They

The equal time commutation relations treat time specially. This is generally undesirable in relativity. So, in this section we derive a commutator, $[\varphi(x), \varphi(y)]$ at two arbitrary spacetime points, x and y . We start by writing

$$\varphi(x) = \varphi^+(x) + \varphi^-(x), \quad (7.1.1)$$

where φ^+ and φ^- are as in [Equations \(6.5.7\) and \(6.5.8\)](#). Since φ^+ only consists of annihilation operators and φ^- only consists of creation operators we have

$$[\varphi^+(x), \varphi^+(y)] = [\varphi^-(x), \varphi^-(y)] = 0. \quad (7.1.2)$$

Hence, expanding out the commutator we have

$$[\varphi(x), \varphi(y)] = [\varphi^-(x), \varphi^+(y)] + [\varphi^-(x), \varphi^+(y)]. \quad (7.1.3)$$

Consider the first of these,

$$[\varphi^+(x), \varphi^-(y)] = \int \tilde{d}p \int \tilde{d}p' [a(\mathbf{p}), a^\dagger(\mathbf{p}')] e^{-ip \cdot x + ip' \cdot y} \quad (7.1.4)$$

$$= \int \tilde{d}p \int \tilde{d}p' \delta(\mathbf{p} - \mathbf{p}') e^{-ip \cdot x + ip' \cdot y} \quad (7.1.5)$$

$$= \int \tilde{d}p e^{-ip \cdot (x-y)} \quad (7.1.6)$$

$$=: i\Delta^+(x-y), \quad (7.1.7)$$

where the last line defines Δ^+ :

$$\Delta^+(x) := i \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} e^{ip \cdot x} \quad (7.1.8)$$

The factor of i is a convention to make Δ^+ a real function.

The second commutator can now easily be worked out using antisymmetry:

$$[\varphi^-(x), \varphi^+(y)] = -[\varphi^+(y), \varphi^-(x)] = -i\Delta^+(y-x) =: i\Delta^-(x-y), \quad (7.1.9)$$

where the last equality defines Δ^- .

Combining these results we have

$$[\varphi(x), \varphi(y)] = i\Delta^+(x-y) + i\Delta^-(x-y) =: i\Delta(x-y) \quad (7.1.10)$$

where

$$\Delta(x) := \Delta^+(x) + \Delta^-(x) \quad (7.1.11)$$

$$= i \int \mathfrak{d}p (e^{ip \cdot x} - e^{-ip \cdot x}) \quad (7.1.12)$$

$$= -2 \int \mathfrak{d}p \sin(p \cdot x). \quad (7.1.13)$$

Note that Δ is a real ($\Delta(x) = \Delta(x)^*$), odd ($\Delta(-x) = -\Delta(x)$) function. This is what we would expect since the field is real and the commutator is antisymmetric.

Consider the full form of Δ^+ :

$$i\Delta^+(x) = \int \frac{d^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2) \theta(p_0) e^{-ip \cdot x}. \quad (7.1.14)$$

We can also write Δ^- similarly:

$$i\Delta^-(x) = - \int \frac{d^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2) \theta(p_0) e^{ip \cdot x}. \quad (7.1.15)$$

If we make a change of variables, $p \rightarrow -p$, then $d^4 p \rightarrow d^4 p$, and

$$i\Delta^-(x) = - \int \frac{d^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2) \theta(-p_0) e^{-ip \cdot x}. \quad (7.1.16)$$

This allows us to write Δ as

$$i\Delta(x) = \int \frac{d^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2) \varepsilon(p_0) e^{-ip \cdot x} \quad (7.1.17)$$

where

$$\varepsilon(p_0) := \theta(p_0) - \theta(-p_0) = \begin{cases} +1 & p_0 > 0, \\ 0 & p_0 = 0, \\ -1 & p_0 < 0, \end{cases} \quad (7.1.18)$$

is the sign function.

Note that for proper orthochronous Lorentz transformations Δ is Lorentz invariant, since the sign of p_0 doesn't change and all other terms are manifestly Lorentz invariant.

7.2 Microcausality

Suppose x is space-like, that is $x^2 < 0$. Then there exists some proper orthochronous Lorentz transformation, Λ , taking x to $-x$. Space-like vectors lie outside the light cone, and we can view this transformation as rotating around the light cone. Since $\Delta^\pm(x)$ are individually Lorentz invariant we can apply a Lorentz transformation to one of them and not the other when computing Δ and get the same result, so

$$\Delta(x) = \Delta^+(x) + \Delta^-(x) \quad (7.2.1)$$

$$= \Delta^+(x) + \Delta^-(\Lambda x) \quad (7.2.2)$$

$$= \Delta^+(x) + \Delta^-(-x) \quad (7.2.3)$$

$$= \Delta^+(x) - \Delta^+(x) \quad (7.2.4)$$

$$= 0, \quad (7.2.5)$$

so Δ vanishes for all space-like points. Hence, $[\varphi(x), \varphi(y)]$ vanishes for all space-like separations, $x - y$.

On the other hand no such transformation exists for time-like or light-like points, since it would have to map from one half of the light cone to the other, which is a discontinuous transformation.

We can interpret this as a requirement for causality, specifically for a nonzero commutator between fields at space-like separated points which happen to have the same time coordinate we'd need some form of communication at this time between these points, and that can't happen between space-like points at the same time without violating causality.

This analysis tells us that Δ is not analytic, it has a sudden jump from zero to nonzero at light-like points.

7.3 Equal Time Commutation Relations

The equal time commutation relations can now be viewed as a special case of the covariant commutation relations. Since we have $\pi = \dot{\varphi}$ we want to compute $[\varphi(t, \mathbf{x}), \pi(t, \mathbf{x}')] = [\varphi(t, \mathbf{x}), \dot{\varphi}(t, \mathbf{x}')]$. To do this take the time derivative of Δ :

$$\frac{\partial}{\partial x^0} \Delta(x - y) = i \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} \frac{\partial}{\partial x^0} [e^{ip \cdot (x-y)} - e^{-ip \cdot (x-y)}] \quad (7.3.1)$$

$$= i \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} i\omega(\mathbf{p}) [e^{ip \cdot (x-y)} + e^{-ip \cdot (x-y)}] \quad (7.3.2)$$

$$= -\frac{1}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} [e^{ip \cdot (x-y)} + e^{-ip \cdot (x-y)}]. \quad (7.3.3)$$

For the equal time commutation relations we have $x^0 = y^0$ and so

$$p \cdot (x - y) = \mathbf{p} \cdot (\mathbf{x} - \mathbf{y}). \quad (7.3.4)$$

Hence,

$$\frac{\partial}{\partial x^0} \Delta(x - y) = -\frac{1}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} [e^{i\mathbf{p} \cdot (\mathbf{x} - \mathbf{y})} + e^{-i\mathbf{p} \cdot (\mathbf{x} - \mathbf{y})}] \quad (7.3.5)$$

$$= -\delta^3(\mathbf{x} - \mathbf{y}), \quad (7.3.6)$$

where we've recognised the integral representation of the Dirac delta:

$$\int \frac{d^3 \mathbf{p}}{(2\pi)^3} e^{\pm i\mathbf{p} \cdot (\mathbf{x} - \mathbf{y})} = \delta^3(\mathbf{x} - \mathbf{y}). \quad (7.3.7)$$

We then have

$$\left. \frac{\partial}{\partial x^0} [i\Delta(x - y)] \right|_{x^0=y^0=t} = \frac{\partial}{\partial x^0} [\varphi(t, \mathbf{x}), \varphi(t, \mathbf{y})] \quad (7.3.8)$$

$$= [\dot{\varphi}(t, \mathbf{x}), \varphi(t, \mathbf{y})] \quad (7.3.9)$$

$$= [\pi(t, \mathbf{x}), \varphi(t, \mathbf{y})] \quad (7.3.10)$$

$$= -i\delta^3(\mathbf{x} - \mathbf{y}) \quad (7.3.11)$$

Hence,

$$[\varphi(t, \mathbf{x}), \pi(t, \mathbf{y})] = i\delta^3(\mathbf{x} - \mathbf{y}). \quad (7.3.12)$$

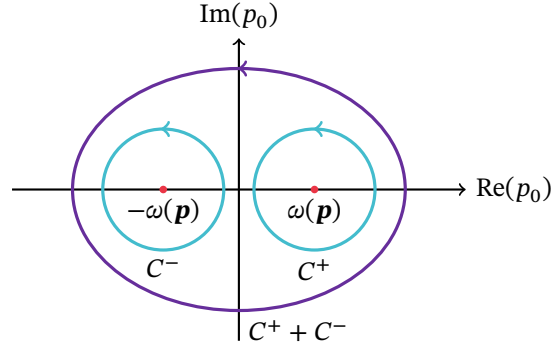


Figure 7.1: The contours used in the contour representation of Δ^\pm and Δ .

So, we can view the equal time commutation relations as a specific case of the covariant commutation relations. This is why the equal time commutation relations can still produce Lorentz invariant physics.

7.4 Contour Representation

R This section makes use of some complex analysis, in particular contour integrals and the residue theorem. See either *Methods of Theoretical Physics* or *Methods of Mathematical Physics* for details.

The contour representation of Δ^\pm is

$$\Delta^\pm(x) = - \int_{C^\pm} \frac{d^4 p}{(2\pi)^4} \frac{e^{-ip \cdot x}}{p^2 - m^2}. \quad (7.4.1)$$

The contours C^\pm are anticlockwise circles containing the poles at $\pm\omega(\mathbf{p})$. This is shown in Figure 7.1.

To see why this works start with

$$p^2 - m^2 = (p_0 + \omega(\mathbf{p}))(p_0 - \omega(\mathbf{p})), \quad (7.4.2)$$

which shows there are poles at $p_0 = \pm\omega(\mathbf{p})$. We can then apply the residue theorem, which says that

$$\oint_\gamma f(z) dz = 2\pi i \sum_i \text{Res}(f, a_i) \quad (7.4.3)$$

where γ is a closed contour, f is analytic on and inside γ , a_i are the poles of f inside γ , and $\text{Res}(f, a_i)$ is the residue of f at a_i .

We have a particularly easy case of two simple poles. In the case of a simple pole the residue at a_i is given by

$$\text{Res}(f, a_i) = (z - a_i)f(z)|_{z=a_i}, \quad (7.4.4)$$

so

$$\text{Res}\left(\frac{e^{-ip \cdot x}}{(p_0 + \omega(\mathbf{p}))(p_0 - \omega(\mathbf{p}))}, \pm\omega(\mathbf{p})\right) = \frac{e^{-ip \cdot x}}{p_0 \pm \omega(\mathbf{p})} \Big|_{p_0 = \mp\omega(\mathbf{p})} = \mp \frac{e^{-ip \cdot x}}{2\omega(\mathbf{p})}. \quad (7.4.5)$$

Hence,

$$-\int_{C^+} \frac{dp_0}{2\pi} \frac{e^{-ip \cdot x}}{p^2 - m^2} = -i \frac{e^{-ip \cdot x}}{2\omega(\mathbf{p})} \Big|_{p_0=\omega(\mathbf{p})}, \quad (7.4.6)$$

and

$$-\int_{C^-} \frac{dp_0}{2\pi} \frac{e^{-ip \cdot x}}{p^2 - m^2} = i \frac{e^{-ip \cdot x}}{2\omega(\mathbf{p})} \Big|_{p_0=-\omega(\mathbf{p})}. \quad (7.4.7)$$

That is,

$$-\int_{C^\pm} \frac{dp_0}{2\pi} \frac{e^{-ip \cdot x}}{p^2 - m^2} = \mp i \frac{e^{\mp ip \cdot x}}{2\omega(\mathbf{p})} \Big|_{p_0=\omega(\mathbf{p})}. \quad (7.4.8)$$

Having performed the integral over p_0 we are left with an integral over \mathbf{p} so

$$-\int_{C^\pm} \frac{d^4 p}{(2\pi)^4} \frac{e^{-ip \cdot x}}{p^2 - m^2} = \mp i \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{e^{mpip \cdot x}}{2\omega(\mathbf{p})} = \Delta^\pm(x). \quad (7.4.9)$$

The contour representation of Δ is then simply given by

$$\Delta(x) = - \int_C \frac{d^4 p}{(2\pi)^4} \frac{e^{-ip \cdot x}}{p^2 - m^2} \quad (7.4.10)$$

where $C = C^+ + C^-$ is the combined contour, with the middle parts cancelling out, to give a contour containing both poles as shown in [Figure 7.1](#).

7.5 Δ as Green's Functions

To interpret Δ and Δ^\pm consider the Klein–Gordon equation,

$$0 = (\partial^2 + m^2)\varphi. \quad (7.5.1)$$

Taking the Fourier transform and using the rule that $\partial_x \rightarrow -ip$ we have

$$0 = ((-ip)^2 + m^2)\tilde{\varphi} \implies (p^2 - m^2)\tilde{\varphi} = 0. \quad (7.5.2)$$

Transforming back then gives us the contour representation of Δ or Δ^\pm , depending on the boundary conditions.

This means we can interpret Δ and Δ^\pm as classical Green's functions for the Klein–Gordon equation, so

$$(\partial^2 + m^2)\Delta(x) = \delta^4(x). \quad (7.5.3)$$

We call Δ and Δ^\pm propagators, since they tell us how a quantum propagates between states.

Part II

Interactions

Eight

Scattering

8.1 Interactions

Quantum things are generally more difficult than classical things.

Richard Ball

So far, we have only considered Lagrangians which are quadratic in the field, we call these **free field Lagrangians**, since quadratic Lagrangians give linear equations of motion with plane wave solutions, which we interpret as many independent harmonic oscillators, each corresponding to a noninteracting particle. In order to have our particles do anything, such as be measured, we need interactions. This means introducing a nonquadratic term to our Lagrangian.

8.1.1 φ^3 Theory

The simplest nonquadratic term that we can add to our Lagrangian is a cubic term, giving the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2 \varphi^2 - \frac{1}{6}g\varphi^3. \quad (8.1.1)$$

The first two terms are just the usual Klein–Gordon Lagrangian, call it \mathcal{L} . We can treat this as the Lagrangian for a free particle. The extra term, $-g\varphi^3/6$, is our interaction term, call it \mathcal{L}_I . We call a theory with a cubic term like this a **φ^3 theory**. The factor of $1/6$ in this factor is just conventional, its there to make equations simpler later. A φ^n theory would equivalently have a factor of $1/n!$, each time we differentiate this factor becomes closer to 1. The factor of g is called the **coupling constant**. It measures how strong the interaction is.

We can easily apply the Euler–Lagrange equations to this modified Lagrangian, without much changing. We have

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} = \partial^2 \varphi, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \varphi} = -m^2 \varphi - \frac{1}{3}g\varphi^2. \quad (8.1.2)$$

Hence, we have the equations of motion

$$(\partial^2 + m^2)\varphi = -\frac{1}{2}g\varphi^2. \quad (8.1.3)$$

This is just the Klein–Gordon equation with some sort of source term.

This is now a nonlinear differential equation. These are, in general, hard to solve. There is no general theory and we are usually restricted to specific examples with solutions. Because of this we assume that the interaction isn't very strong, that is that $g \ll 1$, and then we solve by expanding in g and truncating at some appropriate point in our calculations.

8.2 The S Matrix

Most processes in physics are, when it comes down to it, scattering. We see light after it scatters off an object, two objects collide in classical mechanics, two protons collide in the LHC, a whole bunch of particles collide all over the place and a statistical study of this gives statistical mechanics, and so on. Scattering is a very simple process, there are three steps:

- We start in some initial state with free particles of definite momentum, spin, etc.
- The particles undergo some scattering process, which we treat as a black box.
- We finish in some final state with free particles of definite momentum, spin, etc.

In order to allow us to consider free particles we assume that the initial state happens a long long time before the scattering, taking $t \rightarrow -\infty$. Similarly we assume the final state is a long long time after the scattering, taking $t \rightarrow +\infty$. Call the initial state $|\Psi, -\infty\rangle$ and the final state $|\Psi, +\infty\rangle$. Then an in between state at time t is $|\Psi, t\rangle$.

The initial and final state must somehow be related if we are to pass from one to the other. We take this relation to be of the form

$$|\Psi, +\infty\rangle = S|\Psi, -\infty\rangle \quad (8.2.1)$$

where S , known as the **S matrix**, is the operator describing the scattering process.

Now suppose that we start with some fixed, known initial state, $|i\rangle$, so $|\Psi, -\infty\rangle = |i\rangle$. We can choose some complete orthonormal basis of possible final states, $\{|f\rangle\}$, which allows us to expand the final state as

$$|\Psi, +\infty\rangle = \sum_f |f\rangle \langle f|\Psi, +\infty\rangle. \quad (8.2.2)$$

We interpret

$$|\langle f|\Psi, +\infty\rangle|^2 = \mathbb{P}(i \rightarrow f) \quad (8.2.3)$$

as the probability that when we start in state $|i\rangle$ our scattering process finishes in state $|f\rangle$.

Using the definition of S we can write the amplitude as

$$\langle f|\Psi, +\infty\rangle = \langle f|S|\Psi, -\infty\rangle = \langle f|S|i\rangle =: S_{fi}, \quad (8.2.4)$$

where $S_{fi} := \langle f|S|i \rangle$ is a matrix element of the scattering operator. So the matrix elements of the scattering operator tell us the amplitude for a particular scattering process.

We can expand the final state as

$$|\Psi, +\infty\rangle = \sum_f |f\rangle S_{fi}, \quad (8.2.5)$$

and take the conjugate to get

$$\langle\Psi, +\infty| = \sum_f S_{if}^* \langle f|. \quad (8.2.6)$$

Combining this, and working with normalised states, we get

$$1 = \langle\Psi, +\infty|\Psi, +\infty\rangle = \sum_{f,f'} S_{fi} S_{if'}^* \langle f|f'\rangle = \sum_{f,f'} S_{fi} S_{if'}^* \delta_{ff'} = \sum_f S_{if}^* S_{fi} \quad (8.2.7)$$

or in terms of operators,

$$1 = S^\dagger S, \quad (8.2.8)$$

so the scattering operator is unitary.

We can interpret this as conservation of probability, since $\{|f\rangle\}$ is a complete set of states we must have

$$\sum_f \mathbb{P}(i \rightarrow f) = 1, \quad (8.2.9)$$

and hence if we have a definite start state that is also a complete set, so the evolution must be unitary.

Note that the notion of the S matrix as discussed here is more general than in relativistic quantum mechanics since in QFT we can create and destroy particles, so our scattering can include processes like annihilation or pair production.

8.3 Interaction Picture

We now introduce the third, and final, picture of quantum mechanics, called the **interaction picture**, or **Dirac picture**. Suppose we have a Hamiltonian

$$H = H_0 + H_I \quad (8.3.1)$$

where H_0 describes a free field and H_I an interaction.

In the Dirac picture both operators and states have time dependence. The states evolve as if they were in the Schrödinger picture, but under only the free Hamiltonian, so

$$|\psi, t\rangle_D := e^{iH_0 t} |\psi, t\rangle_S = e^{iH_0 t} e^{-iH t} |\psi\rangle_H \quad (8.3.2)$$

where subscripts D, S, and H denote the Dirac, Schrödinger, and interaction pictures respectively. The first equality is the definition of a state in the Dirac picture, and the second comes from inverting the definition of a state in the Heisenberg picture:

$$|\psi\rangle_H := e^{iH t} |\psi, t\rangle_S. \quad (8.3.3)$$

Operators in the Dirac picture transform as if they were in the Heisenberg picture, but only under the free Hamiltonian, so

$$A_D(t) = e^{iH_0 t} A_S e^{-iH_0 t}. \quad (8.3.4)$$

Take the derivative of the equation defining a state in the Dirac picture, this gives

$$i \frac{\partial}{\partial t} |\psi, t\rangle_D = -H_0 e^{iH_0 t} |\psi, t\rangle_S + e^{iH_0 t} i \frac{\partial}{\partial t} |\psi, t\rangle_S. \quad (8.3.5)$$

Now recognise the final term as $e^{iH_0 t}$ times one side of the Schrödinger equation:

$$i \frac{\partial}{\partial t} |\psi, t\rangle_S = H |\psi, t\rangle_S = H_0 |\psi, t\rangle_S + H_1 |\psi, t\rangle_S. \quad (8.3.6)$$

Substituting this into the previous equation the first term cancels leaving us with

$$i \frac{\partial}{\partial t} |\psi, t\rangle_D = e^{iH_0 t} H_1 |\psi, t\rangle_S = e^{iH_0 t} H_1 e^{-iH_0 t} |\psi, t\rangle_D = H_1^D |\psi, t\rangle_D. \quad (8.3.7)$$

This shows that states evolve in time according to the free Hamiltonian, H_0 . The time evolution due to the interaction is exhibited by the operators.

We derived the Heisenberg equation by differentiating an arbitrary operator in both pictures. Doing the same here, between the Dirac and Schrödinger picture gives rise to the equation of motion

$$\frac{\partial}{\partial t} A_D(t) = i[H_0, A_D(t)]. \quad (8.3.8)$$

Note that $H_0^S = H_0^D = H_0^H$, but in general $H_1^D \neq H_1^S$, which comes from the general fact that $[H_0, H_1] \neq 0$.

From now on we will work entirely in the Dirac picture. Accordingly we will drop the labels D, S, and H.

8.4 Dyson Series

The **Dyson series** is a formal solution to the evolution equation

$$i \frac{\partial}{\partial t} |\Psi, t\rangle = H_I(t) |\Psi, t\rangle. \quad (8.4.1)$$

Now we make the time dependence of the interaction Hamiltonian explicit. The solution to this is

$$|\Psi, t\rangle = |\Psi, -\infty\rangle - i \int_{-\infty}^t dt_1 H_I(t_1) |\Psi, t_1\rangle. \quad (8.4.2)$$

To see this just take the derivative:

$$i \frac{\partial}{\partial t} |\Psi, t\rangle = i \frac{\partial}{\partial t} |\Psi, -\infty\rangle + \frac{\partial}{\partial t} \int_{-\infty}^t dt_1 H_I(t_1) |\Psi, t_1\rangle = H_I(t) |\Psi, t\rangle. \quad (8.4.3)$$

Here the first term vanishes as it is constant and the second term is easily computed by an application of the fundamental theorem of calculus.

The problem is that this solution for $|\Psi, t\rangle$ has $|\Psi, t_1\rangle$ in it, in particular we integrate up to $t_1 = t$. This means that we need to know the state at time t in order to evaluate the state at time t . This seems like a problem. The solution is that we are assuming weak interactions. This means that the operator $H_I(t)$ is, in some sense, “small”, and the operator $H_I(t)H_I(t_1)$ is “smaller”. Of course, operators don’t have sizes in this way, but we can think about specific matrix elements and it all works out. This suggests that we can solve this problem recursively with each level of recursion being a smaller and smaller correction. The first level of recursion is to substitute $|\Psi, t_1\rangle$ with the entire right hand side of Equation (8.4.2), being careful to avoid reusing integration variables, giving

$$|\Psi, t\rangle = |\Psi, -\infty\rangle - i \int_{-\infty}^t dt_1 H_I(t_1) \left[|\Psi, -\infty\rangle - i \int_{-\infty}^{t_1} dt_2 H_I(t_2) |\Psi, t_2\rangle \right]. \quad (8.4.4)$$

Note that the second integral is quadratic in the Hamiltonian, so is “smaller” than the rest of the terms. We can repeat this process, writing $|i\rangle = |\Psi, -\infty\rangle$ for compactness, and we get

$$|\Psi, t\rangle = |i\rangle - i \int_{-\infty}^t dt_1 H_I(t_1) \left[|i\rangle - i \int_{-\infty}^{t_1} dt_2 H_I(t_2) \left\{ |i\rangle - i \int_{-\infty}^{t_2} dt_3 H_I(t_3) |\Psi, t_3\rangle \right\} \right].$$

Continuing on like this we have a series

$$\begin{aligned} |\Psi, t\rangle = & |i\rangle + (-i) \int_{-\infty}^t dt_1 H_I(t_1) |i\rangle + (-i)^2 \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1) H_I(t_2) |i\rangle \\ & + (-i)^3 \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3 H_I(t_1) H_I(t_2) H_I(t_3) |i\rangle + \dots \end{aligned} \quad (8.4.5)$$

Each term in this series is “smaller” than the previous term and so we can truncate at some point once we’ve achieved the required level of accuracy.

Taking t to infinity we get

$$\begin{aligned} |\Psi, +\infty\rangle = & |i\rangle + (-i) \int_{-\infty}^{\infty} dt_1 H_I(t_1) |i\rangle + (-i)^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1) H_I(t_2) |i\rangle \\ & + (-i)^3 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3 H_I(t_1) H_I(t_2) H_I(t_3) |i\rangle + \dots \end{aligned} \quad (8.4.6)$$

Recognising that this has the form of an operator, albeit an infinite series of integrals, acting on the initial state to give the final state we can identify this operator as the S matrix:

$$\begin{aligned} S = & \mathbb{1} + (-i) \int_{-\infty}^{\infty} dt_1 H_I(t_1) + (-i)^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1) H_I(t_2) \\ & + (-i)^3 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3 H_I(t_1) H_I(t_2) H_I(t_3) + \dots \end{aligned} \quad (8.4.7)$$

Or, more compactly,

$$S = \sum_{n=0}^{\infty} (-i)^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \dots \int_{-\infty}^{t_{n-1}} dt_n H_I(t_1) H_I(t_2) \dots H_I(t_n). \quad (8.4.8)$$

Notice that the times, t_1 , t_2 , and so on appear in increasing order, that is in the integral $t_1 < t_2 < \dots < t_n$, we can neglect the endpoints of the integration ranges where we may have equality since individual points don't contribute to an integral.

To proceed we make use of the following identity:

$$\begin{aligned} \int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_1} dt_2 \int_{t_a}^{t_2} \dots \int_{t_a}^{t_{n-1}} dt_n A(t_1)A(t_2) \dots A(t_n) \\ = \frac{1}{n!} \int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_b} dt_2 \dots \int_{t_a}^{t_b} dt_n T[A(t_1)A(t_2) \dots A(t_n)]. \end{aligned} \quad (8.4.9)$$

Here t_a and t_b are arbitrary initial and final times with $t_a < t_b$, and A is an arbitrary operator. The operator T is called the **time ordering**. We take the arguments, which are operators, and order according to evaluation time in order of increasing time from left to right, so that the first operator to act on some state is the operator evaluated at the earliest time. That is, for two operators

$$T[A(t_1), B(t_2)] := \begin{cases} A(t_1)B(t_2) & t_1 > t_2 \\ B(t_2)A(t_1) & t_2 > t_1 \end{cases} \quad (8.4.10)$$

$$= \theta(t_1 - t_2)A(t_1)B(t_2) + \theta(t_2 - t_1)B(t_2)A(t_1). \quad (8.4.11)$$

This can then be extended to n operators in the obvious way, although it becomes harder to write in terms of Heaviside step functions.

We'll show how this works for the $n = 2$ case, and then the general case follows by induction. Start with the right hand side of the identity. For $n = 2$ we have

$$I = \int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_b} dt_2 T[A(t_1)A(t_2)]. \quad (8.4.12)$$

We can split the integral over t_2 into two parts, one running from t_a to t_1 , and the other from t_1 to t_b :

$$I = \int_{t_a}^{t_b} dt_1 \left[\int_{t_a}^{t_1} dt_2 T[A(t_1)A(t_2)] + \int_{t_1}^{t_b} dt_2 T[A(t_1)A(t_2)] \right]. \quad (8.4.13)$$

Consider the second term,

$$\int_{t_a}^{t_b} dt_1 \int_{t_2}^{t_b} dt_2 T[A(t_1)A(t_2)]. \quad (8.4.14)$$

This is a double integral. This corresponds to integrating over the triangle in the plane bounded by $t_1 = t_2$, $t_1 = t_a$, and $t_2 = t_b$. We can integrate over the same region in a different way. Instead integrate along t_2 from t_a to t_1 and then along t_1 from t_a to t_b . This corresponds to the integral

$$\int_{t_a}^{t_b} dt_2 \int_{t_a}^{t_2} dt_1 T[A(t_1)A(t_2)]. \quad (8.4.15)$$

Now, t_1 and t_2 are just integration variables, and the integrand is symmetric in t_1 and t_2 , as $T[A(t_1)A(t_2)] = T[A(t_2)A(t_1)]$, the ordering of operators at different times within a time ordering is not important, since we're going to time order

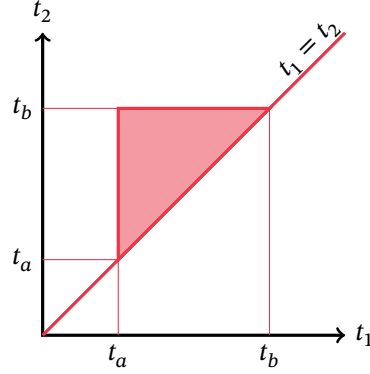


Figure 8.1: The region of integration for the $n = 2$ case of the identity.

them. We are then free to rename t_1 and t_2 , and in particular swap them without changing anything else,

$$\int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_1} dt_2 T[A(t_1)A(t_2)]. \quad (8.4.16)$$

Combining this back with the first term we get

$$I = 2 \int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_1} dt_2 T[A(t_1)A(t_2)]. \quad (8.4.17)$$

Noticing with these integration ranges the operators are already time ordered, so we can drop the explicit time ordering, and that $2 = 2!$ and dividing through by 2 we have proven the identity for $n = 2$.

Now, suppose that the identity holds for k integrals, and we have $k+1$ integrals. We can use the induction hypothesis to swap k integrals, and then we pairwise swap integrals as above before the unswapped integral has also been swapped.

We can use this identity to write

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \cdots \int_{-\infty}^{\infty} dt_n T[H_I(t_1)H_I(t_2) \cdots H_I(t_n)]. \quad (8.4.18)$$

We can now replace H_I with an integral over the Hamiltonian density giving

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int dx_1 \int dx_2 \cdots \int dx_n T[\mathcal{H}_I(x_1)\mathcal{H}_I(x_2) \cdots \mathcal{H}_I(x_n)]. \quad (8.4.19)$$

This is called the **Dyson series**.

Note that the interaction Hamiltonian density is given by $\mathcal{H}_I = -\mathcal{L}_I$. This follows directly from the definition

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_I = \pi\dot{\phi} - \mathcal{L} = \pi\dot{\phi} - \mathcal{L}_0 - \mathcal{L}_I \quad (8.4.20)$$

where the $\pi\dot{\phi} - \mathcal{L}_0$ term is the unperturbed Hamiltonian density. Using this we can write

$$S = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int dx_1 \int dx_2 \cdots \int dx_n T[\mathcal{L}_I(x_1)\mathcal{L}_I(x_2)\cdots\mathcal{L}_I(x_n)] \quad (8.4.21)$$

$$= \exp \left[i \int d^4x \mathcal{L}_I \right]. \quad (8.4.22)$$

So the S matrix can be viewed as the phase factor given by the interaction action. This last step is simply a formal rewriting recognising the exponential series. When expanding make sure to have each integral in each power be over a different variable, and to time order the integrand. From this form we can see that, assuming the action is Lorentz invariant, then so is the S matrix. This works because time ordering is Lorentz invariant, because it can be done using only terms like $\theta(t - t')$, and each such term is invariant under proper orthochronous Lorentz transformations.

8.5 Interpretation

For a given initial state, $|i\rangle$, and final state, $|f\rangle$, we need to find terms in the interaction Lagrangian, \mathcal{L}_I which give a nonzero transition amplitude for $|i\rangle$ to $|f\rangle$. The entire effect of the interaction is in \mathcal{L}_I , which is an operator given by a polynomial in the fields. This means that really it's just a collection of creation and annihilation operators. Therefore the entire effect of the interaction is to create and destroy particles. We then look for terms with

- annihilation operators destroying the initial state,
- creation operators creating the final state,
- paired creation/annihilation operators creating and then destroying intermediate states.

In terms of particles, we look for terms which destroy the incoming particles, potentially create and then destroy one or more new particles, then create some particles which form the final state. The initial and final particles must be on-shell for us to observe them. The intermediate particles however needn't be on-shell, since we cannot observe them. These are called **virtual particles**. Don't put too much stock in the existence and off-shell nature of these virtual particles, they're really just a mathematical construct with different combinations of intermediate particles corresponding to different terms in the Dyson series.

In the interaction picture the fields satisfy the free field equation, and therefore have the usual creation and annihilation operators. We can therefore use normal ordering, placing all creation operators on the left and all annihilation operators on the right, in order to ensure that incoming particles are destroyed and then new particles are created, rather than the other way round.

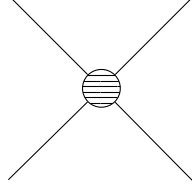
8.6 φ^3 Interactions

In φ^3 theory the interaction Hamiltonian is

$$\mathcal{H}_I = -\mathcal{L}_I = \frac{1}{6} g : \varphi^3 :, \quad (8.6.1)$$

where we now normal order the φ^3 factor in order to destroy existing particles and then create new particles.

We will consider 2–2 scattering, that is two particles come in, interact in some way, and two particles leave. This can be represented by a diagram like



(8.6.2)

Here the blob represents some unknown black box type interaction. Time flows, in our convention, from left to right, so the left two legs represent the incoming particles, and the right two the outgoing particles. In order to destroy the incoming particles we start with a $\varphi^+\varphi^+$ term. Then to create two new particles we need a $\varphi^-\varphi^-$ term. Recall that we order operators from right to left, as we want the $\varphi^+\varphi^+$ term to act on the initial before the $\varphi^-\varphi^-$ term. We could therefore look for a

$$\varphi^-\varphi^-\varphi^+\varphi^+ \quad (8.6.3)$$

term in the series. However, we won't find such a term since in φ^3 theory we can only have terms with a multiple of 3 factors of φ^\pm . This means that there are no direct 2–2 scattering events in φ^3 theory, such events must always go via at least one intermediate virtual particle state. So, next we look at the term in the series which is quadratic in the Hamiltonian, and hence has six factors of φ^\pm . There are two possible orderings of terms with the required starting point of $\varphi^+\varphi^+$ to annihilate the initial state and end point of $\varphi^-\varphi^-$ to create the final state. They are

$$\varphi^-\varphi^-\varphi^+\varphi^+\varphi^+, \quad \text{and} \quad \varphi^-\varphi^-\varphi^+\varphi^-\varphi^+. \quad (8.6.4)$$

When acting on the initial state the first vanishes, since after annihilating the initial state we are left with the vacuum state, which we then act on with the annihilator φ^+ . So, the only nonzero second term quadratic in the Hamiltonian is

$$\varphi^-\varphi^-\varphi^+\varphi^-\varphi^+. \quad (8.6.5)$$

We interpret this as annihilating the initial state, creating a new particle, annihilating it, and then destroying it again. The amplitude for this particular term is proportional to g^2 , since each Hamiltonian introduces a factor of g . Further, by assuming that energy is conserved and everything is local when we annihilate the two initial particles we must immediately create the virtual particle, and all of this must happen at the same position, so this is a single spacetime event, x_2 . Then when the virtual particle is annihilated we must immediately, and at the same location, create the new particles, so this is a second spacetime event, x_1 . These correspond to the two points at which the interaction Hamiltonians are evaluated so we'll have a nonzero term

$$-\frac{g^2}{36} \int_{-\infty}^{\infty} d^4x_1 \int_{-\infty}^{\infty} d^4x_2 \varphi^-(x_2)\varphi^-(x_2)\varphi^+(x_1)\varphi^-(x_2)\varphi^+(x_2)\varphi^+(x_2). \quad (8.6.6)$$

Pictorially we can represent this as


(8.6.7)

Here the left and right legs represent the incoming and outgoing particles, and the single line in the middle is the virtual particle.

Next, we would consider terms cubic in the Hamiltonian. However, things quickly get out of hand if we proceed as we did here. The cubic term will have nine factors of φ . So, we want a way to do this all more systematically. This is the theory we will develop in the next chapter.

Nine

Wick's Theorem

To proceed we need to be able to deal with time orderings of operators. Our interpretation, of particles being created and destroyed, works best if we normal order things, placing all creation operators on the left and annihilation operators on the right. So, we look for a way to rewrite a time ordering in terms of normal orderings, and other quantities which can be calculated more easily. We proceed with the case of two operators, then we introduce a general theorem central to much of quantum field theory.

9.1 Two Operator Case

Suppose we have two operator valued fields, A and B , which we can write as

$$A = A^+ + A^-, \quad \text{and} \quad B = B^+ + B^-, \quad (9.1.1)$$

where A^+ and B^+ are expressions in the annihilation operators and A^- and B^- are expressions in the creation operators. Consider the product $:A(x)B(x'):$. Notice that the normal ordering of a sum is the sum of the normal ordering of the individual terms, that is $:C + D: = :C: + :D:$. Expanding in terms of A^\pm and B^\pm we have

$$\begin{aligned} :A(x)B(x'): &= :[A^+(x) + A^-(x)][B^+(x') + B^-(x')]: \\ &= :A^+(x)B^+(x') + A^+(x)B^-(x') + A^-(x)B^+(x') + A^-(x)B^-(x') : \\ &= A^+(x)B^+(x') + B^-(x')A^+(x) + A^-(x)B^+(x') + A^-(x)B^-(x'). \end{aligned}$$

Now, consider what happens when we subtract this result from $A(x)B(x')$ without the normal ordering, that is

$$A(x)B(x') = A^+(x)B^+(x') + A^+(x)B^-(x') + A^-(x)B^+(x') + A^-(x)B^-(x'). \quad (9.1.2)$$

The only term that doesn't cancel is the second term where we have to swap the ordering, so

$$A(x)B(x') - :A(x)B(x'): = A^+(x)B^-(x') - B^-(x')A^+(x) = [A^+(x), B^-(x')]. \quad (9.1.3)$$

The commutator of creation and annihilation operators is just a number, or more accurately a Dirac delta, and so it doesn't act on states, in particular

$$\langle 0|[A^+(x), B^-(x')]|0\rangle = \langle 0|0|[A^+(x), B^-(x')] = [A^+(x), B^-(x')], \quad (9.1.4)$$

since we choose to normalise the vacuum state. We also have

$$\langle 0|[A^+(x), B^-(x')]|0\rangle = \langle 0|A^+(x)B^-(x')|0\rangle - \langle 0|B^-(x')A^+(x)|0\rangle = \langle 0|A^+(x)B^-(x')|0\rangle, \quad (9.1.5)$$

where the second term vanishes since the annihilation operator $A^+(x)$ acts on the vacuum.

We then have

$$A(x)B(x') = :A(x)B(x'): + \langle 0|A(x)B(x')|0\rangle. \quad (9.1.6)$$

We're nearly there, recall that the definition of time ordering is

$$T[A(x)B(x')] = \theta(t - t')A(x)B(x') + \theta(t' - t)B(x')A(x), \quad (9.1.7)$$

where t is the time component of x and t' is the time component of x' .

At this point we make the observation that it doesn't matter what order we write operators in if we are normal ordering them, since we're going to change the order anyway, so

$$:A(x)B(x'): = :B(x')A(x):. \quad (9.1.8)$$

Hence, if we normal order the time ordering we have

$$:T[A(x)B(x')]: = \theta(t - t')A(x)B(x') + \theta(t' - t):B(x')A(x): \quad (9.1.9)$$

$$= \theta(t - t')A(x)B(x') + \theta(t' - t):A(x)B(x'): \quad (9.1.10)$$

$$= :A(x)B(x'): \quad (9.1.11)$$

So, the normal ordering of a time ordering is just the normal ordering.

Consider [Equation \(9.1.6\)](#). This holds for any product of two operators, in particular if we time order the operators it still works, since we can just separate into the $t < t'$ and $t > t'$ cases and then proceed without the time ordering. Hence,

$$T[A(x)B(x')] = :A(x)B(x'): + \langle 0|T[A(x)B(x')]|0\rangle. \quad (9.1.12)$$

Here we use $:T[A(x)B(x')]: = :A(x)B(x'):$ to drop the time ordering in the first term on the right.

This equation, and others like it, are very important in quantum field theory, so important that we introduce a notation to simplify it. We define the **Wick contraction** of two operators, A and B to be

$$\overline{A(x)B(x')} := \langle 0|T[A(x)B(x')]|0\rangle. \quad (9.1.13)$$

Note that this is just a number, not an operator. This allows us to state Wick's theorem for two operators.

Lemma 9.1.14 For two operators, A and B , such that $A = A^+ + A^-$ and $B = B^+ + B^-$ where A^+ and B^+ are expressions in the annihilation operator

and A^- and B^- are expressions in the creation operators we have

$$T[A(x)B(x')] = :A(x)B(x') : + \overline{A(x)B(x')}. \quad (9.1.15)$$

Proof. See above work. \square

9.2 Three Operator Case

The proof of Wick's theorem, the generalisation of the previous lemma to an arbitrary number of operators, proceeds inductively. It will be simpler to prove if we first prove the three operator case to see how the inductive logic works in detail. As before our goal is to express a time ordered product of operators in terms of their normal ordering and contractions. We now have an extra operator, C , which we again assume can be written as $C = C^+ + C^-$ where C^+ is an expression in the annihilation operators and C^- an expression in the creation operators.

We want to evaluate $T[A(x_1)B(x_2)C(x_2)]$ where x_i has time component t_i . For simplicity, and without loss of generality, we assume that $t_1 > t_2 > t_3$. If this isn't the case we can reorder freely within the time ordering and then relabel our operators so that it is. We don't consider here the case where the operators are evaluated at the same time, we'll do this later. From now on we suppress the arguments to our operators, but be aware that they are not all evaluated at the same point.

Since, by assumption, A is the first operator evaluated we can move it outside of the time ordering, since it will always appear on the left of any time ordered product of operators:

$$T[ABC] = A T[BC]. \quad (9.2.1)$$

Now we can apply [Lemma 9.1.14](#) to the time ordering of B and C , giving

$$T[ABC] = A :BC : + \overline{ABC}. \quad (9.2.2)$$

Now express A as $A^+ + A^-$, so

$$T[ABC] = (A^+ + A^-) :BC : + \overline{BC} \quad (9.2.3)$$

$$= A^+ :BC : + A^+ \overline{BC} + A^- :BC : + A^- \overline{BC}. \quad (9.2.4)$$

Notice that since A^- is a creation operator and is on the left of $:BC :$ we can move it into the normal ordering without changing anything, so

$$T[ABC] = A^+ :BC : + A^+ \overline{BC} + :A^- BC : + A^- \overline{BC}. \quad (9.2.5)$$

Things aren't quite so simple for $A^+ :BC :$. In order to bring A^+ , which is an annihilation operator, into the normal ordering we need it to be on the right. We can achieve this at the cost of a commutator:

$$A^+ :BC : = :BC : A^+ + [A^+, :BC :] = :BC A^+ : + [A^+, :BC :]. \quad (9.2.6)$$

Now that A^+ has been brought into the normal ordering we can move it around within the ordering, so

$$A^+ :BC: = :A^+BC: + [A^+, :BC:]. \quad (9.2.7)$$

The commutator can be evaluated with the identity

$$[X, YZ] = XYZ - YZX = XYZ - YXZ + YXZ - YZX = [X, Y]Z + Y[X, Z]. \quad (9.2.8)$$

Since the commutator is just a number the normal ordering doesn't effect it, and so we have

$$A^+ :BC: = :A^+BC: + :[A^+, B^-]C: + :B[A^+, C^-]:. \quad (9.2.9)$$

We've also used here that $[X^+, Y] = [X^+, Y^-]$, since X^+ commutes with Y^+ , as both are just formed from annihilation operators.

Exactly as in the two operator case we can replace the commutator with its vacuum expectation value, and then drop the commutator, these are just numbers, so can be brought out of the normal ordering.

$$A^+ :BC: = :A^+BC: + [A^+, B^-]:C: + [A^+, C^-]:B: \quad (9.2.10)$$

$$= :A^+BC: + \langle 0|[A^+, B^-]|0\rangle:C: + \langle 0|[A^+, C^-]|0\rangle:B: \quad (9.2.11)$$

$$= :A^+BC: + \langle 0|A^+B^-|0\rangle:C: + \langle 0|A^+C^-|0\rangle:B:. \quad (9.2.12)$$

Notice that since we're normal ordering a single term which decomposes as the sum of creation and annihilation operators we don't really need the normal orderings any more, but we keep them so our result fits the general result.

Going back to our starting point we now have

$$T[ABC] = A^+ :BC: + A^+ \overline{BC} + :A^-BC: + A^- \overline{BC} \quad (9.2.13)$$

$$= :ABC: + \langle 0|T[A^+B^-]|0\rangle:C: + \langle 0|T[A^+C^-]|0\rangle:B: + \overline{ABC}$$

$$= :ABC: + \overline{AB}:C: + \overline{AC}:B:. \quad (9.2.14)$$

Here we use

$$A^+ \overline{BC} + A^- \overline{BC} = (A^+ + A^-) \overline{BC} = \overline{ABC}, \quad (9.2.15)$$

and

$$:A^+BC: + :A^-BC: = :(A^+ + A^-)BC: = :ABC: \quad (9.2.16)$$

The result here is that the time ordering is given by the normal ordering plus all pairs of contractions, normal ordering anything that isn't contracted.

9.3 General Theorem

We can now prove the general theorem relating time orderings of an arbitrary number of operators and their normal orderings and contractions.

Theorem 9.3.1 — Wick's Theorem. Given operators, A, B, C, D, \dots , which can all be written as a sum of an annihilator and creation operators. Suppose A is evaluated at x_1 , B at x_2 , C at x_3 , D at x_4 , and so on, and t_i is the time component of x_i . Then if none of the t_i are equal we have

$$T[ABCD \dots] = :ABCD \dots: \quad (9.3.2)$$

$$+ \overline{AB}:CD \dots: + \overline{AC}:BD \dots: + \overline{AD}:BC \dots: + \dots \quad (9.3.3)$$

$$+ \overline{BC}:AD \dots: + \overline{BD}:AC \dots: + \dots \quad (9.3.4)$$

$$\vdots \quad (9.3.5)$$

$$+ \overline{ABCD} \dots: + \overline{ACBD} \dots: + \overline{ADBC} \dots: + \dots \quad (9.3.6)$$

$$+ \text{all other possible contractions.} \quad (9.3.7)$$

That is, the time ordering is given by the normal ordering, plus the normal ordering contracting each pair of operators, plus the normal ordering contracting every possible pair of pairs of operators, and so on, doing every single contraction possible and normal ordering anything not contracted.

Proof. We proceed by strong induction on the number of operators. The one operator case is trivial and acts as the basis case, equivalently the two operator case can be the basis case and we then use [Lemma 9.1.14](#) as the proof of our basis case.

Now, suppose that The hypothesis holds for $k - 1$ operators for some positive integer k . Take k operators $A, B, C, D \dots$, evaluated at $x_1, x_2, x_3, x_4, \dots$ respectively, with t_i the time component of x_i . Without loss of generality we assume $t_1 > t_2 > \dots > t_k$, if this isn't the case we can reorder within the time ordering and relabel such that it is the case.

We can pull A outside of the time ordering, since it is evaluated last:

$$T[ABCD \dots] = A T[BCD \dots]. \quad (9.3.8)$$

Writing $A = A^+ + A^-$ where A^+ and A^- are expressions in the annihilation and creation operators respectively we have

$$T[ABCD \dots] = (A^+ + A^-) T[BCD \dots]. \quad (9.3.9)$$

The time ordering on the right has $k - 1$ operators, so the induction hypothesis allows us to write this as

$$T[ABCD \dots] = (A^+ + A^-)(:BCD \dots: + \overline{BC}:B \dots: \quad (9.3.10)$$

$$+ \text{all contractions not involving } A). \quad (9.3.11)$$

We can trivially put A^- into all of the normal orderings, since it is a creation operator and is on the left. Therefore consider the A^+ terms. We have

$$A^+ :BCD \dots: = :BCD \dots: A^+ + [A^+, :BCD \dots:] \quad (9.3.12)$$

$$= :BCD \dots A^+: + [A^+, :BCD \dots:] \quad (9.3.13)$$

$$= :A^+ BCD \dots: + [A^+, :BCD \dots:], \quad (9.3.14)$$

Iteratively applying the identity in Equation (9.2.8) we have

$$\begin{aligned} A^+ :BCD \cdots : &= :A^+BCD \cdots : + [A^+, B^-] :CD \cdots : \\ &+ [A^+, C^-] :BD \cdots : + [A^+, D^-] :BC \cdots : + \cdots . \end{aligned} \quad (9.3.15)$$

As with the two and three operator cases we can replace the commutators with contractions, giving

$$\begin{aligned} A^+ :BCD \cdots : &= :A^+BCD \cdots : + \overline{AB} :CD \cdots : \\ &+ \overline{AC} :BD \cdots : + \overline{AD} :BC \cdots : + \cdots . \end{aligned} \quad (9.3.16)$$

So we end up with all contractions containing A and some other operator, normal ordering all other operators.

The exact same logic works for other terms, like

$$A^+ \overline{BC} :D \cdots : = \overline{AD} \overline{BC} : \cdots : + \cdots , \quad (9.3.17)$$

where we have all pairs of pairs of contractions with one pair being B and C and the other containing A . Considering all such terms we see that we get all possible contractions involving A . Going back to our result for $T[ABCD \cdots]$ as containing all contractions not involving A , and recombining A^+ and A^- to get A we prove the theorem. \square

Corollary 9.3.18 — Wick's Theorem for Mixed Time Products Suppose that some of the operators in the statement of Theorem 9.3.1 are evaluated at the same time. Then the time ordering is instead given by all contractions not involving contractions of equal time operators and normal ordering everything left over.

Proof. This follows by applying Wick's theorem at unequal times, and then setting some times equal and seeing that contractions of equal time terms vanish, since

$$T(:A(t, \mathbf{x})B(t, \mathbf{x}') :) = :A(t, \mathbf{x})B(t, \mathbf{x}') : \quad (9.3.19)$$

so

$$\overline{AB} = T[:AB:] - :AB: = 0. \quad (9.3.20)$$

\square

9.4 Feynman Propagator

Consider the contraction

$$\overline{\varphi(x)\varphi(x')} = \langle 0 | T[\varphi(x)\varphi(x')] | 0 \rangle \quad (9.4.1)$$

$$= \theta(t - t') \langle 0 | \varphi(x)\varphi(x') | 0 \rangle + \theta(t' - t) \langle 0 | \varphi(x')\varphi(x) | 0 \rangle \quad (9.4.2)$$

where t is the time component of x and t' is the time component of x' .

We've seen that

$$i\Delta^+(x - x') = [\varphi^+(x), \varphi^-(x')] \quad (9.4.3)$$

is just a number, and so

$$\langle 0 | [\varphi^+(x), \varphi^-(x')] | 0 \rangle = [\varphi^+(x), \varphi^-(x')] \langle 0 | 0 \rangle = [\varphi^+(x), \varphi^-(x')]. \quad (9.4.4)$$

Hence, we have

$$i\Delta^+(x - x') = \langle 0 | [\varphi^+(x), \varphi^-(x')] | 0 \rangle \quad (9.4.5)$$

$$= \langle 0 | \varphi^+(x)\varphi^-(x') | 0 \rangle \quad (9.4.6)$$

$$= \langle 0 | \varphi(x)\varphi(x') | 0 \rangle, \quad (9.4.7)$$

where expanding the commutator the $\varphi^-(x')\varphi^+(x)$ term annihilates the vacuum, and so vanishes, and in the last part expanding φ using φ^\pm shows that all terms apart from the $\varphi^+(x)\varphi^-(x')$ term vanish as they involve annihilating the vacuum:

$$\varphi(x)\varphi(x') = (\varphi^+(x) + \varphi^-(x))(\varphi^+(x') + \varphi^-(x')) \quad (9.4.8)$$

$$= \varphi^+(x)\varphi^+(x') + \varphi^+(x)\varphi^-(x') + \varphi^-(x)\varphi^+(x') + \varphi^-(x)\varphi^-(x')$$

Similarly, we have

$$i\Delta^-(x - x') = -\langle 0 | [\varphi(x'), \varphi(x)] | 0 \rangle. \quad (9.4.9)$$

This just follows from $\Delta^+(x) = -\Delta^-(-x)$. We can then define

$$i\Delta_F(x - x') := \overline{\varphi(x)\varphi(x')} \quad (9.4.10)$$

$$= \langle 0 | T[\varphi(x)\varphi(x')] | 0 \rangle \quad (9.4.11)$$

$$= \theta(t - t')i\Delta^+(x - x') - \theta(t' - t)i\Delta^-(x - x'). \quad (9.4.12)$$

This is called the **Feynman propagator**. Without the factor of i it's defined as

$$\Delta_F(x) = \theta(t)\Delta^+(x) - \theta(-t)\Delta^-(x). \quad (9.4.13)$$

The interpretation is as follow:

- For $t > t'$ we have a positive energy, that is Δ^+ , mode propagating from x' to x .
- For $t < t'$ we have a negative energy, that is Δ^- , mode propagating from x' to x .

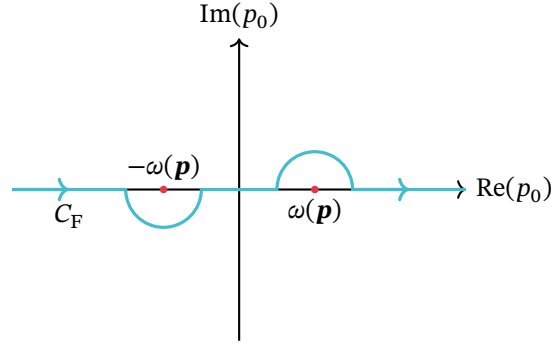


Figure 9.1: The contour used to compute the Feynman propagator.

Notice that we can rewrite the Feynman propagator as

$$\Delta_F(x - x') = \theta(t - t')\Delta^+(x - x') + \theta(t' - t)\Delta^+(x' - x). \quad (9.4.14)$$

This has the same value but a different interpretation:

- For $t > t'$ we have a positive energy, that is Δ^+ , mode propagating from x' to x .
- For $t < t'$ we have a positive energy, that is Δ^+ , mode propagating from x to x' .

Comparing these interpretations we see that a negative energy mode (which we'll later associate with antiparticles) is equivalent to a positive energy mode travelling in the opposite direction.

9.4.1 Contour Representation

As well as this interpretation the Feynman propagator has the nice property of automatically encoding the time ordering. This isn't that useful in the current form, since we have explicit Heaviside step functions doing the encoding. However, the Feynman propagator has a contour representation which makes the time ordering less obvious.

The contour representation of Δ_F is

$$\Delta_F(x) = \int_{C_F} \frac{d^4 p}{(2\pi)^4} \frac{e^{-ip \cdot x}}{p^2 - m^2} \quad (9.4.15)$$

where C_F is the contour given by the real axis in the p_0 plane with a slight dip below the pole at $-\omega(\mathbf{p})$ and a slight dip above the pole at $\omega(\mathbf{p})$. This is shown in [Figure 9.1](#).

To see why this works we consider two cases, first, suppose that $x_0 > 0$. Then close the contour in the lower half plane, so that p_0 has a large, negative, imaginary component and overall $-ip \cdot x$ has a large, negative, real component. Closing with a semicircle Jordan's lemma tells us that the integral along this arc vanishes. Hence, the value of the integral along the C_F contour is equal to the value of the integral around this closed loop. We can apply the residue theorem to the contour about

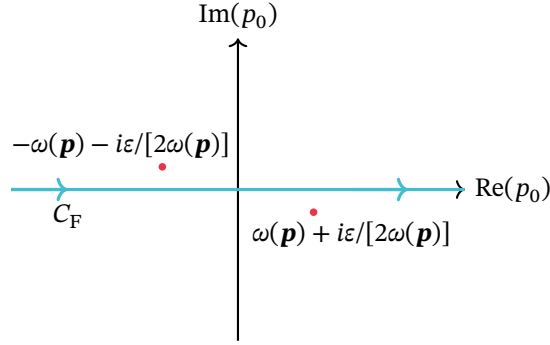


Figure 9.2: The contour for calculating the Feynman propagator using the $i\varepsilon$ prescription.

the loop. There is a single pole enclosed in this loop, the one at $\omega(\mathbf{p})$. However, we've already computed the integral of this function around a contour containing the pole $\omega(\mathbf{p})$ in Section 7.4, the only difference is a minus sign not in the Feynman propagator contour representation, so we can conclude that the result is $\Delta^+(x)$.

Similarly, if $x_0 < 0$ we close the contour in the upper half plane. Again the integral vanishes along this arc and we can then treat this as a contour containing only the pole at $-\omega(\mathbf{p})$, the result of which tells us that this integral gives $-\Delta^-(x)$. So, combining these results we see that we get the expected result of $\pm\Delta^\pm(x)$ depending on the sign of x_0 .

9.4.1.1 The $i\varepsilon$ Prescription

This contour integral is not particular nice to evaluate, we have to awkwardly detour around the poles on the real axis. Instead we can slightly modify the integrand to give the representation

$$\Delta_F(x) = \int \frac{d^4 p}{(2\pi)^4} \frac{e^{-ip \cdot x}}{p^2 - m^2 + i\varepsilon}. \quad (9.4.16)$$

This moves the two poles slightly off the real axis. Let $\varepsilon > 0$ be a small positive real number, then

$$p^2 - m^2 + i\varepsilon = p_0^2 - \left(\omega(\mathbf{p}) - \frac{i\varepsilon}{2\omega(\mathbf{p})} \right)^2, \quad (9.4.17)$$

just expand this out to check, and drop terms of order ε^2 . The poles are then shifted by $i\varepsilon/[2\omega(\mathbf{p})]$, so that the pole which was at $-\omega(\mathbf{p})$ is now at $-\omega(\mathbf{p}) + i\varepsilon/[2\omega(\mathbf{p})]$, and the pole which was at $\omega(\mathbf{p})$ is now at $\omega(\mathbf{p}) - i\varepsilon/[2\omega(\mathbf{p})]$. This is shown in Figure 9.2. We can then integrate along the real axis as normal, and then take $\varepsilon \rightarrow 0$ after our computations have finished.

9.4.2 Feynman Propagator as a Green's Function

Like the other propagators the Feynman propagator is a Green's function of the Klein-Gordon equation, that is

$$(\partial^2 + m^2)\Delta_F(x) = \delta^4(x). \quad (9.4.18)$$

The difference is just in the boundary conditions.

The classical retarded and advanced Green's functions, as well as the Δ^\pm propagators are Green's functions for an equation with Neumann boundary conditions, that is, the boundary conditions fix initial conditions on a function and its derivative, in our case we have an initial condition on φ and $\pi = \dot{\varphi}$. This is a classical way of thinking, we have some initial state and we let it evolve according to some equation. The problem is that the assumption is we can know both φ and π at the same time, this is not possible due to the uncertainty principle.

The Feynman propagator on the other hand corresponds to Dirichlet boundary conditions, where we set initial and final conditions on some function, in this case φ . This is much closer to how we think in quantum mechanics, where we consider an initial and final state and transitions between them.

Ten

Feynman Diagrams

This is a great course because there are loads and loads of integrals to do but they're all really easy, so you don't feel any pain but you get to feel proud after.

Richard Ball

In this section we develop the theory of Feynman diagrams. A diagrammatic representation of interactions which have the benefit of greatly simplifying the maths and also seeming quite physical. We stress however, that any one diagram is not a true representation of a physical process, rather it corresponds to a single term in an expansion describing a physical process.

We start from the Dyson expansion,

$$S = \sum_{n=0}^{\infty} S^{(n)}, \quad (10.0.1)$$

where

$$S^{(n)} = \frac{(-i)^n}{n!} \int d^4x_1 \cdots \int d^4x_n T[\mathcal{H}_I(x_1) \cdots \mathcal{H}_I(x_n)]. \quad (10.0.2)$$

We will use ϕ^3 theory as an example, but the process is basically the same, but usually with harder integrals, for other theories. This means that

$$\mathcal{H}_I(x) = \frac{g}{3!} : \phi(x)^3 : = \frac{g}{3!} : (\phi^+(x) + \phi^-(x))^3 :. \quad (10.0.3)$$

We consider low order terms in this series, and develop Feynman diagrams as we go along.

10.1 Evaluating Low Order Terms

10.1.1 $n = 0$

The $n = 0$ term, $S^{(0)}$, is just 1. This represents the case where nothing happens, our initial state just remains and is equal to the final state.

10.1.2 $n = 1$

The $n = 1$ term is

$$S^{(1)} = \frac{-ig}{3!} \int d^4x (\varphi^{+3} + 3\varphi^{-2}\varphi^+ + 3\varphi^-\varphi^{+2} + \varphi^{-3}). \quad (10.1.1)$$

We can consider each term in the integrand separately.

The first term represents three annihilators, therefore this corresponds to three particles in the initial state entering, being destroyed, and then a vacuum final state. As a diagram we represent this as



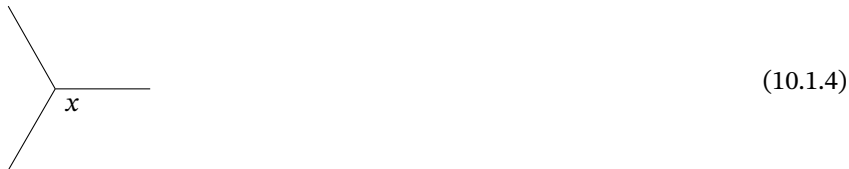
This interaction doesn't conserve momentum, so it must give zero contribution to the overall value of $S^{(1)}$.

The second term represents an annihilator and two creators, so this corresponds to a single particle entering, being destroyed, and then two new particles being created. Since all three factors of φ are evaluated at the same spacetime point, x , this occurs all at once, which we draw as



It's not immediately obvious, but this term cannot actually conserve momentum in φ^3 theory, since all three particles have the same mass, and so it gives zero, which we'll show later.

The third term is two annihilators and one creation operator, it corresponds to two particles being annihilated and then one new particle being created. As a diagram it is



Like the previous term this vanishes in φ^3 theory.

The fourth term is three creation operators, so represents three particles being

created from an initial vacuum state. Diagrammatically it is


(10.1.5)

Like the first term this one doesn't conserve momentum and so must have no net contribution to $S^{(1)}$.

10.1.3 $n = 2$

The $n = 2$ term is

$$S^{(2)} = \frac{g^2}{2!(3!)^2} \int d^4x \int d^4y T(:\varphi(x)^3 : : \varphi(y)^3 :). \quad (10.1.6)$$

We can use Wick's theorem to evaluate this normal ordered product. We can then split $S^{(2)}$ further into four terms based on the number of contractions:

$$S^{(2)} = S_0^{(2)} + S_1^{(2)} + S_2^{(2)} + S_3^{(2)}, \quad (10.1.7)$$

Where $S_i^{(2)}$ is the term in $S^{(2)}$ corresponding to terms with i contractions. Note that since we have six factors of φ we have a maximum of three contractions.

10.1.3.1 No Contractions

The terms with no contractions correspond to the first term in Wick's theorem,

$$:\varphi(x)^3 \varphi(y)^3 :. \quad (10.1.8)$$

Since there are no contractions between the field evaluated at the two spacetime points, x and y , we get disconnected diagrams, such as


(10.1.9)

This is a single diagram corresponding to two separate processes occurring at x and y . This diagram corresponds to the term

$$\varphi^-(x)\varphi^-(y)\varphi^-(y)\varphi^+(x)\varphi^+(x)\varphi^+(y), \quad (10.1.10)$$

and other terms give similar diagrams.

It can be shown that all disconnected diagrams vanish in φ^3 theory. Further, any disconnected diagram is simply a combination of connected diagrams, so we don't every need to consider disconnected diagrams. In terms of physics disconnected diagrams correspond to multiple unrelated scattering events occurring, and we can simply treat them all separately.

10.1.3.2 One Contraction

Wick's theorem, generalised for equal times, doesn't have contractions between operators at the same time. Hence, we don't need to consider contractions of $\varphi(x)$ with itself, or $\varphi(y)$ with itself. This means that we have the term

$$3^2 \overline{\varphi(x)\varphi(y)} : \varphi(x)^2 \varphi(y)^2 :. \quad (10.1.11)$$

The factor of 3^2 is combinatoric, it comes from there being three ways to select which $\varphi(x)$ field to contract and three ways to select which $\varphi(y)$ field to contract, and then since these are all really identical we can combine them into a single term with weighting 3^2 .

The contraction corresponds to a virtual particle propagating from x to y . The four factors of φ in the normal ordering tell us that we'll have four external particles. There are three ways to do this, we can have 1-3, 2-2, or 3-1 interactions, where these are the number of initial particles to final particles.

An example of a 1-3 term is given by the diagram



This corresponds to the term

$$\overline{\varphi(x)\varphi(y)} \varphi^+(x) \varphi^-(x) \varphi^-(y) \varphi^-(y). \quad (10.1.13)$$

This term, in φ^3 theory, doesn't conserve momentum since all particles have the same mass, so must give no overall contribution to $S^{(2)}$.

Similarly any 3-1 terms must vanish.

More interesting are the 2-2 terms. There are three such terms. The first is



This corresponds to the term

$$\overline{\varphi(x)\varphi(y)} \varphi^+(x) \varphi^+(x) \varphi^-(y) \varphi^-(y). \quad (10.1.15)$$

We call this the **s-channel** diagram.

The second is

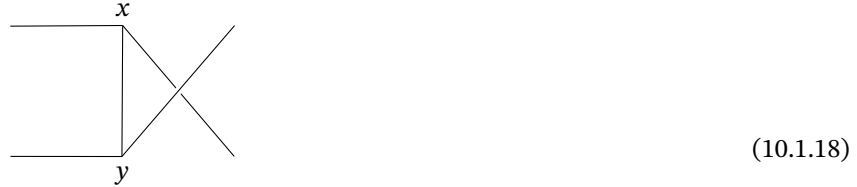


This corresponds to the term

$$\overline{\varphi(x)\varphi(y)}\varphi^+(x)\varphi^+(y)\varphi^-(x)\varphi^-(y). \quad (10.1.17)$$

We call this the ***t*-channel** diagram.

The third is



This corresponds to the term

$$\overline{\varphi(x)\varphi(y)}\varphi^+(x)\varphi^+(y)\varphi^-(y)\varphi^-(x) \quad (10.1.19)$$

If we take any of these three diagrams and swap x and y then nothing changes, so this gives us a factor of 2 to include, which nicely cancels the factor of $1/2!$ in the full expression for $S^{(2)}$.

10.1.3.3 Two Contractions

Terms with two contractions correspond to a term proportional to

$$\overline{\varphi(x)\varphi(y)}\overline{\varphi(x)\varphi(y)}:\varphi(x)\varphi(y):. \quad (10.1.20)$$

The two contractions mean we have two internally propagating virtual particles, and the two normal ordered terms mean we have two external particles.

There is a single such diagram, up to swapping x and y :



It can be shown that this term actually diverges, giving an infinite result. We call this a one **loop diagram**, since it has a single loop. In contrast the diagrams we have considered so far without loops are **tree diagrams**.

10.1.3.4 Three Contractions

Terms with three contractions correspond to a term proportional to

$$\overline{\varphi(x)\varphi(y)}\overline{\varphi(x)\varphi(y)}\overline{\varphi(x)\varphi(y)}. \quad (10.1.22)$$

This has no external particles and three internal virtual particles propagating from x to y . It corresponds to the diagram



This is called a **vacuum bubble** since it is formed entirely of virtual particles, and both the initial and final state are vacuums. Since there are no external particles there are no physically observable contributions from terms like this. It can be shown that bubble diagrams factor out to all levels in all theories, so we don't need to consider them. They correspond to some virtual particles coming into existence and then all annihilating.

10.2 Feynman Diagrams in Momentum Space

As always in QFT you have to get everything right.

Richard Ball

So far with these diagrams we've been working in position space, evaluating the fields at x . External states are momentum eigenstates, this suggests that we should instead work in momentum space instead, which means working with $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$.

We are interested in computing the matrix element

$$S_{fi} = \langle f | S | i \rangle. \quad (10.2.1)$$

We can express $|i\rangle$ and $|f\rangle$ as free particle states created by $a^\dagger(\mathbf{p})$ from the vacuum state. These states have definite momentum and all of the particles are on-shell.

Consider the one particle state with momentum \mathbf{p} . This is created with a single creation operator acting on the vacuum:

$$|\mathbf{p}\rangle := a^\dagger(\mathbf{p})|0\rangle. \quad (10.2.2)$$

Now act on this with $\varphi^+(x)$. We can expand $\varphi^+(x)$ in momentum space using Equation (6.5.7):

$$\varphi^+(x)|\mathbf{p}\rangle = \int \mathrm{d}\mathbf{p}' e^{-i\mathbf{p}' \cdot \mathbf{x}} a(\mathbf{p}') a^\dagger(\mathbf{p}) |0\rangle \quad (10.2.3)$$

$$= \int \mathrm{d}\mathbf{p}' e^{-i\mathbf{p}' \cdot \mathbf{x}} [a(\mathbf{p}'), a^\dagger(\mathbf{p})] |0\rangle \quad (10.2.4)$$

$$= \int \mathrm{d}\mathbf{p}' e^{-i\mathbf{p}' \cdot \mathbf{x}} \delta(\mathbf{p} - \mathbf{p}') |0\rangle \quad (10.2.5)$$

$$= e^{-i\mathbf{p} \cdot \mathbf{x}} |0\rangle. \quad (10.2.6)$$

Here we've used

$$[a(\mathbf{p}'), a^\dagger(\mathbf{p})] |0\rangle = a(\mathbf{p}') a^\dagger(\mathbf{p}) |0\rangle - a^\dagger(\mathbf{p}) a(\mathbf{p}') |0\rangle = a(\mathbf{p}') a^\dagger(\mathbf{p}) |0\rangle \quad (10.2.7)$$

with the second term vanishing as we annihilate the vacuum.

Taking the conjugate of this result, and using $(\varphi^+)^\dagger = \varphi^-$ we have

$$\langle \mathbf{p} | \varphi^-(x) = \langle 0 | e^{i\mathbf{p} \cdot \mathbf{x}}. \quad (10.2.8)$$

Since we're going to be working in momentum space it's useful to have an expression for the contraction in momentum space, we get this through the contour representation of the Feynman propagator:

$$\overline{\varphi(x)} \varphi(y) = i \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{e^{-ik \cdot (x-y)}}{k^2 - m^2 + i\epsilon} = i\Delta_F(x-y), \quad (10.2.9)$$

where the k^0 integral is performed along the real axis, and we set ϵ to zero after our calculations. Note that we've written this here using k , the wavenumber, which is the same as the momentum when $\hbar = 1$. This is mostly to free up the letter p for the momentum of the particles.

We are now in a position to start computing terms in the Dyson series, which we can view as an expansion in g . We won't do all of the ones discussed above, just enough to get the idea.

10.2.1 First Order

Consider the first order term



$$(10.2.10)$$

Here we've labelled the particles with their momenta. The initial state contains just a single particle, so its $|i\rangle = |\mathbf{p}\rangle$. The final states contain two particles, so the final state is a tensor product of single particle states:

$$|f\rangle = |\mathbf{q}, \mathbf{q}'\rangle = |\mathbf{q}\rangle |\mathbf{q}'\rangle = |\mathbf{q}\rangle \otimes |\mathbf{q}'\rangle. \quad (10.2.11)$$

This works because we assume no interaction between the two final particles, at least not at first order in g .

The matrix element we wish to compute is

$$\langle f | S_{1,2}^{(1)} | i \rangle = \langle \mathbf{q}, \mathbf{q}' | \left[\frac{-ig}{3!} \int d^4x \varphi^-(x)^2 \varphi^+(x) \right] | \mathbf{p} \rangle. \quad (10.2.12)$$

Here $S_{1,2}^{(1)}$ is the term in $S^{(1)}$ corresponding to the term we wish to compute. The factor of 3 in the integral comes from the number of ways we can choose one of the factors of φ to be φ^+ . We can show that

$$\langle \mathbf{q}, \mathbf{q}' | \varphi^-(x)^2 = 2e^{iq \cdot x} e^{iq' \cdot x} \langle 0 |, \quad (10.2.13)$$

where the factor of two comes from the number of ways we can act on $\langle \mathbf{q}, \mathbf{q}' |$ with $\varphi^-(x)^2$. We can either act on $\langle \mathbf{q} |$ and then $\langle \mathbf{q}' |$, or the other way round. We also have

$$\varphi^+(x) | \mathbf{p} \rangle = e^{-ip \cdot x} | 0 \rangle. \quad (10.2.14)$$

Putting this into our calculation, and noticing that the factors of 2 and 3 cancel with the 3! we have

$$\langle f | S_{1,2}^{(1)} | i \rangle = -ig \int d^4x e^{iq \cdot x} e^{iq' \cdot x} e^{-ip \cdot x} \quad (10.2.15)$$

$$= -ig(2\pi)^4 \delta^4(q + q' - p). \quad (10.2.16)$$

The Dirac delta enforces conservation of momentum, the momentum in, p , must equal the momentum out, $q + q'$, in order to get a nonzero result.

All external particles are on-shell, and in φ^3 theory all particles have the same mass. This means that $p^2 = q^2 = q'^2 = m^2$. Now, consider $(p - q)^2$. Conservation of momentum tells us that $p - q = q'$, and so $(p - q)^2 = q'^2 = m^2$. We can also expand $(p - q)^2$ to get

$$m^2 = (p - q)^2 \quad (10.2.17)$$

$$= p^2 + q^2 - 2p \cdot q \quad (10.2.18)$$

$$= m^2 + m^2 + 2p \cdot q. \quad (10.2.19)$$

Hence, we have $2p \cdot q = m^2$. In the rest frame of the incoming particle we have $p^\mu = (m, \mathbf{0})$, and we can write $q^\mu = (E, \mathbf{q})$ without loss of generality. We therefore have

$$2p \cdot q = 2mE = m^2 \implies E = \frac{m}{2}. \quad (10.2.20)$$

This is what we would expect, since the two particles are produced each must get exactly half of the energy available from the first particle, and that energy is just its mass. We also know that $E^2 = m^2 + \mathbf{q}^2$, and so $E \geq m$. This means that $E = m/2$ can't satisfy $p = q + q'$. Hence, the contribution from this term vanishes as the condition imposed by the Dirac delta can never be fulfilled.

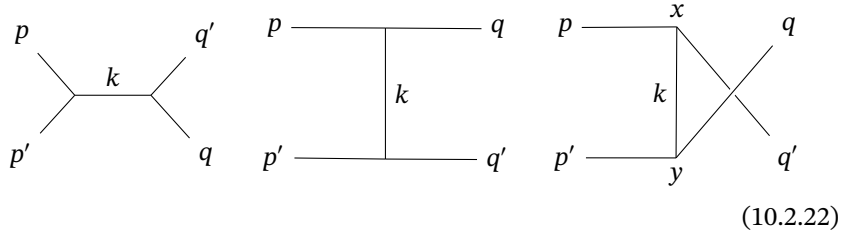
Note that in other theories with multiple particles with different masses it is possible for this term to be nonzero. It's also not true that all vertices like this vanish, the key difference is that if one of the particles is a virtual particle then it needn't satisfy the on-shell condition and so it can have arbitrary momentum, and we can conserve momentum with particles of the same mass.

The key result is that, in φ^3 theory,

$$\langle f | S_{1-2}^{(1)} | i \rangle = 0. \quad (10.2.21)$$

10.2.2 Second Order

For now we restrict ourselves to tree diagrams, we'll come to loops later. There are three diagrams to consider then:



Note that while we assume the same initial and final states the momenta of the virtual particles need not, and in fact aren't, the same, even though we've labelled them all k .

10.2.2.1 s-Channel

Let's start with the s-channel process. We want to compute

$$\langle f | S_s^{(2)} | i \rangle = \langle \mathbf{q}, \mathbf{q}' | \left[\frac{-g^2}{(3!)^2} \int d^4x \int d^4y \varphi^-(y)^2 \overline{\varphi(x)} \varphi(y) \varphi^+(x)^2 \right] | \mathbf{p}, \mathbf{p}' \rangle$$

We can compute the action of $(\varphi^\pm)^2$ on our initial and final states fairly easily:

$$\varphi^+(x)^2 | \mathbf{p}, \mathbf{p}' \rangle = 2e^{-ip \cdot x} e^{-ip' \cdot x} | 0 \rangle, \quad (10.2.23)$$

$$\langle \mathbf{q}, \mathbf{q}' | \varphi^-(y)^2 = 2e^{iq \cdot y} e^{iq' \cdot y} \langle 0 |. \quad (10.2.24)$$

Again, the factors of 2 come from there being two orders in which we can act on the paired states.

Noticing that the factor of 3^2 in the integrand, as well as the factors of 2 above, cancel with the $1/(3!)^2$ term, and using $\langle 0|0\rangle = 1$, we have

$$\langle f|S_s^{(2)}|i\rangle = -g^2 \int d^4x \int d^4y e^{iq\cdot y} e^{iq'\cdot y} \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot(x-y)}}{k^2 - m^2 + i\varepsilon} e^{ip\cdot x} e^{ip'\cdot x}.$$

Performing the x and y integrals we get more Dirac deltas:

$$\langle f|S_s^{(2)}|i\rangle = -ig^2 \int \frac{d^4k}{(2\pi)^4} (2\pi)^4 \delta^4(q + q' - k) (2\pi)^4 \delta^4(k - p - p') \frac{1}{k^2 - m^2 + i\varepsilon}.$$

The two Dirac deltas ensure that momentum is conserved at each vertex, and hence overall, since we must have $q + q' = k$ and $k = p + p'$ to have a nonzero contribution, and so $p + p' = q + q'$. We can then perform the integral over k using one the second Dirac delta and we get

$$\langle f|S_s^{(2)}|i\rangle = (2\pi)^4 \delta^4(q + q' - p - p') \frac{-ig^2}{(p + p')^2 - m^2}. \quad (10.2.25)$$

Note that we can drop the $i\varepsilon$ term since $k^2 = (p + p')^2 \geq 4m^2$, using $p^2 = p'^2 = m^2$, and so the denominator never vanishes. It turns out that this is always the case for tree level diagrams, the $i\varepsilon$ prescription is only needed in loop diagrams. If instead we had used the first Dirac delta to perform the integral we would have gotten a factor of $(q + q')^2$ in the denominator instead, but this is equal to $(p + p')^2$ by momentum conservation.

The Dirac delta in this result is again conserving momentum. The interesting thing is the rest, which we call the **Feynman amplitude** for this process:

$$\mathcal{M}_s = \frac{-ig^2}{(p + p')^2 - m^2}. \quad (10.2.26)$$

10.2.2.2 t -Channel

Now consider the t -channel diagram. Like the s -channel term the t -channel term comes from a term which is sixth order in φ . In particular, it is $:\varphi(x)^3\varphi(y)^3:$, where we call the vertex shared by p and q x , and the vertex shared by p' and q' y . Since we have four external particles we must contract a pair of fields, and since contractions at the same point give zero we must contract $\varphi(x)$ and $\varphi(y)$. There are 3 ways to choose which $\varphi(x)$ and 3 ways to choose which $\varphi(y)$. We then have two factors of $\varphi(x)$ left and two factors of $\varphi(y)$. Looking back at the diagrams we see that a particle is created and destroyed at each vertex, as well as the internal line corresponding to the contraction. This means one of our $\varphi(x)$ values contributes a $\varphi^+(x)$ to this term, destroying an incoming particle, and the other contributes a $\varphi^-(x)$, creating one of the final particles. There are two ways to select which $\varphi(x)$ contributes $\varphi^+(x)$ and which contributes $\varphi^-(x)$. The exact same is true for $\varphi(y)$, so on top of the two factors of 3 from choosing the contraction we have two factors of 2. This gives an overall combinatorial factor of $3 \cdot 3 \cdot 2 \cdot 2 = (3!)^2$, cancelling the $(3!)^2$ which appears in the interaction term. Hence, we have

$$\begin{aligned} \langle f|S_t^{(2)}|i\rangle &= \langle \mathbf{q}, \mathbf{q}' | \left[\frac{-g^2}{(3!)^2} \int d^4x \int d^4y (3!)^2 \overline{\varphi(x)} \varphi(y) \varphi^-(x) \varphi^-(y) \varphi^+(x) \varphi^+(y) \right] | \mathbf{p}, \mathbf{p}' \rangle. \end{aligned}$$

Now, as before we act on the initial state with the annihilators and the final state with the creation operators, giving

$$\varphi^+(x)\varphi^+(y)|\mathbf{p}, \mathbf{p}'\rangle = e^{-ip\cdot x}e^{-ip'\cdot y}|0\rangle, \quad (10.2.27)$$

$$\langle \mathbf{q}, \mathbf{q}' | \varphi^-(x)\varphi^-(y) = e^{iq\cdot x}e^{iq'\cdot y}\langle 0|. \quad (10.2.28)$$

Note that there is no factor of 2 this time, since the particle with momentum p must be destroyed at x in this interaction, and hence we cannot act on this particle with $\varphi^+(y)$, doing so corresponds instead to a u -channel process.

Putting this together with our results we have

$$\langle f | S_t^{(2)} | i \rangle = -g^2 \int d^4x \int d^4y e^{iq\cdot x} e^{iq'\cdot y} \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(y-x)}}{k^2 - m^2 + i\epsilon} e^{-ip\cdot x} e^{-ip'\cdot y}.$$

Collecting the exponents into x and y exponentials we have

$$\langle f | S_t^{(2)} | i \rangle = -g^2 \int d^4x \int d^4y \int \frac{d^4k}{(2\pi)^4} e^{i(q+k-p)\cdot x} e^{i(q'-k-p')\cdot y} \frac{1}{k^2 - m^2 + i\epsilon}.$$

Performing the x and y integrals gives Dirac deltas:

$$\langle f | S_t^{(2)} | i \rangle = -g^2 \int \frac{d^4k}{(2\pi)^4} (2\pi)^4 \delta^4(q+k-p) (2\pi)^4 \delta^4(q'-k-p') \frac{1}{k^2 - m^2 + i\epsilon}.$$

Performing the k integral with the first Dirac delta fixes $k = p - q$, and so we get

$$\langle f | S_t^{(2)} | i \rangle = (2\pi)^4 \delta^4(q + q' - p - p') \frac{-g^2}{(p - q)^2 - m^2}. \quad (10.2.29)$$

As before we have the Dirac delta enforcing energy conservation, and then the amplitude is

$$\mathcal{M}_t = \frac{-ig^2}{(p - q)^2 - m^2}. \quad (10.2.30)$$

Similarly we can compute the u -channel process, and we find

$$\langle f | S_u^{(2)} | i \rangle = (2\pi)^4 \delta^4(q + q' - p - p') \frac{-ig^2}{(p - q')^2 - m^2}. \quad (10.2.31)$$

That is, the amplitudes for the u -channel process is

$$\mathcal{M}_u = \frac{-ig^2}{(p - q')^2 - m^2}. \quad (10.2.32)$$

We now introduce the **Mandelstam invariants**, s , t , and u , defined as

$$s = (p + p')^2 = (q + q')^2, \quad (10.2.33)$$

$$t = (p - q)^2 = (p' - q')^2, \quad (10.2.34)$$

$$u = (p - q')^2 = (p' - q)^2. \quad (10.2.35)$$

These are Lorentz invariants characterising each of the three channels.

Consider the sum of these three invariants:

$$s + t + u = (p + p')^2 + (p - q)^2 + (p - q')^2 \quad (10.2.36)$$

$$\begin{aligned} &= p^2 + p'^2 + 2p \cdot p' + p^2 + q^2 - 2p \cdot q + p^2 + q'^2 - 2p \cdot q' \\ &= 6m^2 + 2(p \cdot p' - p \cdot q - p \cdot q') \end{aligned} \quad (10.2.37)$$

having used the on-shell condition, $p^2 = p'^2 = q^2 = q'^2 = m^2$. Now using conservation of momentum we have $q' = p + p' - q$, and so

$$s + t + u = 6m^2 + 2(p \cdot p' - p \cdot q - p \cdot [p + p' - q]) \quad (10.2.38)$$

$$= 6m^2 + 2(p \cdot p' - p \cdot q - p \cdot p - p \cdot p' + p \cdot q) \quad (10.2.39)$$

$$= 6m^2 + 2m^2 \quad (10.2.40)$$

$$= 4m^2. \quad (10.2.41)$$

This shows that, in a theory where all particles have the same mass, the three Mandelstam invariants are not independent.

We are interested in the total matrix element at second order which, ignoring loops, is

$$\langle f | S^{(2)} | i \rangle = \langle \mathbf{q}, \mathbf{q}' | (S_s^{(2)} + S_t^{(2)} + S_u^{(2)}) | \mathbf{p}, \mathbf{p}' \rangle. \quad (10.2.42)$$

Combing our results we have

$$\langle f | S^{(2)} | i \rangle = (2\pi)^4 \delta^4(q + q' - p - p') (-ig^2) \left[\frac{1}{s - m^2} + \frac{1}{t - m^2} + \frac{1}{u - m^2} \right] \quad (10.2.43)$$

and the amplitude is

$$\mathcal{M}^{(2)} = \mathcal{M}_s + \mathcal{M}_t + \mathcal{M}_u = -ig^2 \left[\frac{1}{s - m^2} + \frac{1}{t - m^2} + \frac{1}{u - m^2} \right]. \quad (10.2.44)$$

Notice that the amplitude is invariant under swapping p and p' and/or swapping q and q' , since this simply swaps t and u if we make one swap, or does nothing if we make two. This implies that our particles are bosons.

Now consider scattering in the centre of mass frame. Here the incoming particles enter with equal and opposite momenta, and therefore also with equal energy since they have the same mass and are on shell. So, $p^\mu = (E, \mathbf{p})$ and $p'^\mu = (E, -\mathbf{p})$. The energy must split equally between the outgoing particles, so they also have energy E , and the total momentum after the interaction must vanish so they have equal and opposite momenta, $q^\mu = (E, \mathbf{q})$ and $q'^\mu = (E, -\mathbf{q})$.

Calculating s we have $s = (p + p')^2 = 4E^2 \geq 4m^2$, where the inequality comes from $E^2 = m^2 + \mathbf{p}^2 \geq m^2$. This is what justified us dropping the $i\epsilon$ earlier, since if $s \geq 4m^2$ then $s - m^2$ will never be zero.

Since all particles have the same mass and energy and are on-shell we must have that $q^2 = p^2$, hence

$$t = -(\mathbf{p} - \mathbf{q})^2 = -2|\mathbf{p}|^2(1 - \cos \vartheta) \leq 0 \quad (10.2.45)$$

where $\cos \vartheta = \mathbf{p} \cdot \mathbf{q} / (|\mathbf{p}||\mathbf{q}|)$. Similarly,

$$u = -(\mathbf{p} + \mathbf{q})^2 = -2\mathbf{p}^2(1 + \cos \vartheta) \leq 0. \quad (10.2.46)$$

This means that $t - m^2$ and $u - m^2$ will always be negative, and in particular are never zero allowing us to drop the $i\epsilon$ term. This also implies that virtual particles are *always* off-shell, as otherwise we would not be able to conserve momentum in 2-1 and 1-2 vertices.

10.3 Feynman Rules

The calculations above all follow roughly the same steps, this suggests that we can formulate a way to go from the diagrams to the final result without having to do these steps every time. This is what Feynman rules do.

For any ϕ^3 process we have

$$\langle f|S|i\rangle = \delta_{fi} + (2\pi)^4 \delta^4\left(\sum p_f - \sum p_i\right) \sum_{n=2}^{\infty} \mathcal{M}^{(n)}. \quad (10.3.1)$$

The Kronecker delta accounts for the case where nothing happens, the Dirac delta enforces energy conservation, where the sums are over all final and then all initial states. The interesting bit is the amplitude, $\mathcal{M}^{(n)}$. Wick's theorem tells us that the total amplitude, \mathcal{M} , has contributions from all topologically distinct diagrams¹, and $\mathcal{M}^{(n)}$ has contributions from all topologically distinct n vertex diagrams. Further, we have to consider all diagrams with equal weight due to Wick's theorem, but the factor of g^n means that we can drop higher order terms, since g is taken to be small.

¹that is, all diagrams which are not related by rearranging where nodes and edges are

From what we've seen so far we can guess most of the Feynman rules, and the others aren't too bad to demonstrate. The **Feynman rules** for ϕ^3 theory are

- Every vertex contributes a factor of $-ig$ to the amplitude.
- Every internal line contributes a factor of $i\tilde{\Delta}_F(k) = i/(k^2 - m^2 + i\varepsilon)$ to the amplitude.
- Every external line contributes a factor of 1 to the amplitude.
- Impose conservation of four-momentum at every vertex.
- Every momentum, k , which is not fixed by conservation of four-momentum gives a factor of

$$\int \frac{d^4k}{(2\pi)^4}. \quad (10.3.2)$$

- A symmetry factor (trivial for simple diagrams).

Note that $\tilde{\Delta}_F$ is the Fourier transform of Δ_F , which is to say it is the integrand in the contour representation.

Eleven

Getting Measurable Results

Experiments can't measure the S matrix directly. Instead we can measure transition rates between states, including decay rates and cross sections. In this chapter we'll discuss how to compute these from the S matrix for comparison with experimental results.

Our starting point will be

$$S_{fi} = \delta_{fi} + (2\pi)^4 \delta^4 \left(\sum p_f - \sum p_i \right) \mathcal{M}. \quad (11.0.1)$$

The δ_{fi} term corresponds to the case where nothing happens. The Dirac delta, where the sums are over all final and initial particles, enforces conservation of momentum, and \mathcal{M} is the Feynman amplitude, which is what we compute in QFT.

11.1 Transition Rate

Consider the probability, P_{fi} , of transitioning from the initial state $|i\rangle$ to the final state $|f\rangle$. This is simply the square of the S matrix, which is given by

$$P_{fi} = |S_{fi}|^2 = (2\pi)^4 \delta^4(0) (2\pi)^4 \sum_f \delta^4 \left(\sum p_f - \sum p_i \right) |\mathcal{M}|^2. \quad (11.1.1)$$

assuming $i \neq f$, that is that something happens. We have used here $\delta(x)^2 = \delta(0)\delta(x)$ as distributions, which follows since

$$\int dx f(x) \delta(x)^2 = \delta(0) f(0), \quad (11.1.2)$$

having used the sifting property of one of the Dirac deltas. The first sum over f , outside of the Dirac delta, corresponds to a sum, which will turn out to really be an integral, over allowed values of the final momentum.

There's a problem here, in that $\delta^4(0)$ is infinite. Consider

$$(2\pi)^4 \delta^4(p) = \int d^4x e^{-ip \cdot x}. \quad (11.1.3)$$

This suggests that

$$(2\pi)^4 \delta^4(0) = \int d^4x = VT \quad (11.1.4)$$

where V is the infinite volume of space and T is the infinite length of time. So, VT is the volume of spacetime. To get around the problem we treat the entire system

as if it is in a finite box in spacetime, and then proceed to compute quantities per unit spacetime volume.

We therefore look to compute the transition rate per unit volume,

$$w_{fi} = \frac{P_{fi}}{VT} = \sum_f (2\pi)^4 \delta(\sum p_f - \sum p_i) |\mathcal{M}|^2. \quad (11.1.5)$$

We can proceed using the completeness relation,

$$1 = \int d\mathbf{p} |\mathbf{p}\rangle \langle \mathbf{p}|, \quad (11.1.6)$$

which holds if

$$\langle \mathbf{p} | \mathbf{p}' \rangle = \delta(\mathbf{p} - \mathbf{p}'). \quad (11.1.7)$$

Writing out the measure in full,

$$1 = \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} |\mathbf{p}\rangle \langle \mathbf{p}|. \quad (11.1.8)$$

The incident flux is

$$\langle \mathbf{p} | \mathbf{p} \rangle = 2\omega(\mathbf{p}) (2\pi)^3 \delta^3(\mathbf{0}). \quad (11.1.9)$$

Identifying $(2\pi)^3 \delta^3(\mathbf{0}) = V$ we have

$$\frac{\langle \mathbf{p} | \mathbf{p}' \rangle}{V} = 2\omega(\mathbf{p}). \quad (11.1.10)$$

This implies we have $2\omega(\mathbf{p})$ states per unit volume.

Now consider the decay of a single particle, p , into n particles, p_i . The decay rate in the frame of the particle, that is the transition rate per decaying particle, is then

$$d\Gamma = (2\pi)^4 \delta(p - \sum p_f) \frac{1}{2\omega(\mathbf{p})} \prod_f \frac{d^3 \mathbf{p}_f}{(2\pi)^3 2E_f} |\mathcal{M}|^2, \quad (11.1.11)$$

where we write $E_f = \omega(\mathbf{p}_f)$. We can package the measure up into a single thing called the n particle **phase space measure**:

$$(dPS)_n := (2\pi)^4 \delta^4(p - \sum p_f) \prod_f \frac{d^3 p_f}{(2\pi)^3 2E_f}, \quad (11.1.12)$$

so

$$d\Gamma = \frac{1}{2E(\mathbf{p})} |\mathcal{M}|^2 (dPS)_n, \quad (11.1.13)$$

where $E(\mathbf{p}) = \omega(\mathbf{p})$.

11.1.1 One to Two Decay

Consider the decay $p \rightarrow p_1 + p_2$. When doing an experiment we have detectors covering some solid angle, and so we are interested in not just the decay rate, but the decay rate into some particular solid angle, which is

$$\frac{d\Gamma}{d\Omega}. \quad (11.1.14)$$

We have two final particles, so we need to compute $(dPS)_2$, which is given by

$$(dPS)_2 = (2\pi)^4 \delta^4(p - p_1 - p_2) \frac{d^3 \mathbf{p}_1}{(2\pi)^3 2\omega(\mathbf{p}_1)} \frac{d^3 \mathbf{p}_2}{(2\pi)^3 2\omega(\mathbf{p}_2)}. \quad (11.1.15)$$

We can perform the p_2 integral using one of the Dirac deltas:

$$\int d^3 \mathbf{p}_2 \delta^4(p_1 + p_2 - p) = \delta(E_1 + E_2 - E), \quad (11.1.16)$$

where we now fix $\mathbf{p}_2 = \mathbf{p} - \mathbf{p}_1$ and $E_2 = |\mathbf{p}_2|^2 + m_2^2$. We can then express the momentum space volume element for p_1 in terms of polar variables:

$$d^3 p_1 = |\mathbf{p}_1|^2 dp_1 d\Omega. \quad (11.1.17)$$

Now use

$$E_1^2 = |\mathbf{p}_1|^2 + m_1^2 \implies E_1 dE_1 = p_1 dp_1, \quad (11.1.18)$$

since $dm_1 = 0$. Then, the partially integrated phase space is

$$\int (dPS)_2 = \frac{1}{4\pi} d\Omega \frac{p_1 dE_1}{E} \delta(E_1 + E_2 - E). \quad (11.1.19)$$

We can then write the phase space measure as

$$\int (dPS)_2 = \frac{1}{(4\pi)^2} \frac{|\mathbf{p}|}{m} d\Omega. \quad (11.1.20)$$

Thus in the centre of mass frame the decay rate into some solid angle is

$$\left(\frac{d\Gamma}{d\Omega} \right)_{\text{CoM}} = \frac{1}{32\pi^2} \frac{|\mathbf{p}|}{m^2} |\mathcal{M}|^2. \quad (11.1.21)$$

11.2 Cross Sections

When we consider collisions there is almost no point in considering collisions with three or more particles, since these are so unlikely to occur. So, consider a scattering of two particles, p_1 and p_2 . We can work in the lab frame, where, say, the second particle is stationary, so

$$p_1 = (\omega(\mathbf{p}_1), \mathbf{p}_1), \quad \text{and} \quad p_2 = (m_2, \mathbf{0}). \quad (11.2.1)$$

We can also write $E_1 = \omega(\mathbf{p}_1)$.

The incident flux, that is the number of particles crossing unit area per unit time, is given by the volume of particles which would cross through said area in

unit time times the density of particles. The volume is simply $|\mathbf{v}|$, where \mathbf{v} is the velocity of the particle, since in one unit time the particles will have moved a distance $|\mathbf{v}|$ forward. The density of particles is $2E_1$, so the incident flux is $2E_1|\mathbf{v}| = 2|\mathbf{p}_1|$, where we've used $\mathbf{p} = \gamma m \mathbf{v}$ and $E = \gamma m$, so $|\mathbf{v}| = |\mathbf{p}|/E$ as expected.

The number of scattering centres in the target per unit volume is the density of particle states, which is $2E_2 = 2m_2$.

The differential cross section is defined as the transition rate per scattering centre per unit incident flux. The idea being that the number of scattering centres and the incident flux are both dependent on the experiment, and the differential cross section is independent of experimental details. The differential cross section is then

$$d\sigma = (2\pi)^4 \delta^4 \left(\sum p_i - \sum p_f \right) \frac{1}{4|\mathbf{p}_1|m_2} \left(\prod_f \frac{d^3 \mathbf{p}_f}{(2\pi)^3 2E_f} \right) |\mathcal{M}|^2. \quad (11.2.2)$$

Now we can use $(p_1 \cdot p_2)^2 = (E_1 m_2)^2 = (\mathbf{p}_1 + m_1)^2 m_2^2$ which gives $|\mathbf{p}_1|^2 m_2^2 = (p_1 \cdot p_2)^2 - m_1^2 m_2^2$, which is Lorentz invariant. This means that we can write the result in a frame independent way as

$$d\sigma = \frac{|\mathcal{M}|^2}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} (dPS)_n. \quad (11.2.3)$$

11.2.1 Two to Two Scattering

Consider the case where we have two incoming and two outgoing particles, $p_1 + p_2 \rightarrow p_3 + p_4$. In the centre of mass frame we have

$$\int (dPS)_2 = \frac{1}{(4\pi)^2} \frac{p'}{W} \quad (11.2.4)$$

where $p' = |\mathbf{p}_3| = |\mathbf{p}_4|$, where equality between the magnitude of the momenta of the particles follows by conservation of momentum in the centre of mass frame, meaning $\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{0}$, so $\mathbf{p}_3 + \mathbf{p}_4 = \mathbf{0}$ also. As well, $W = E_1 + E_2 = E_3 + E_4$, where the second equality is just conservation of energy.

In the centre of mass frame we have

$$p_1 = (E_1, \mathbf{p}), \quad \text{and} \quad p_2 = (E_2, -\mathbf{p}). \quad (11.2.5)$$

Define $[=|\mathbf{p}|]$, and so

$$\mathbf{p}_1 \cdot \mathbf{p}_2 = E_1 E_2 + \mathbf{p}^2 \implies (p_1 \cdot p_2)^2 - m_1^2 m_2^2 = p^2 W^2, \quad (11.2.6)$$

where we've used the on-shell requirement to show this.

Hence, the cross section for scattering per unit solid angle is

$$\left(\frac{d\sigma}{d\Omega} \right)_{\text{CoM}} = \frac{1}{64\pi^2} \frac{1}{W^2} \left(\frac{p'}{p} \right)^2 |\mathcal{M}|^2. \quad (11.2.7)$$

Note that if all particles have the same mass then $p' = p$ and so often we see this equation without the p'/p factor. This is also a valid approximation in the high energy case where the masses provide a negligible contribution to the energy-momentum relation.

11.2.1.1 In φ^3 Theory

In φ^3 theory the two to two scattering gives us the amplitude

$$\mathcal{M} = (-ig^2) \left[\frac{1}{s-m^2} + \frac{1}{t-m^2} + \frac{1}{u-m^2} \right]. \quad (11.2.8)$$

Here $s = W^2$. Hence

$$\left(\frac{d\sigma}{d\Omega} \right)_{\text{CoM}} = \frac{1}{64\pi^2} \frac{g^4}{s} \left| \frac{1}{s-m^2} + \frac{1}{t-m^2} + \frac{1}{u-m^2} \right|^2. \quad (11.2.9)$$

Now consider high energy fixed angle scattering. Then $|\mathbf{p}|^2 \gg m^2$, so $|\mathbf{p}|^2 \approx E^2$. We also have $s = 4E^2 \approx 4\mathbf{p}^2$, $t \approx -2E^2(1 - \cos \vartheta)$ and $u \approx -2E^2(1 + \cos \vartheta)$. This means that the ratios s/u and s/t are fixed since ϑ is fixed and the factors of E^2 cancel. The angular distribution can be written as

$$d\Omega = 2\pi d\cos \vartheta, \quad (11.2.10)$$

and so

$$\left(\frac{d\sigma}{d(\cos \vartheta)} \right)_{\text{CoM}} = \frac{1}{32\pi} \frac{g^4}{s^3} \left(1 + \frac{s}{t} + \frac{s}{u} \right)^2, \quad (11.2.11)$$

where we get an extra factor of 2π from the $\int_0^{2\pi} d\varphi$ integral and the term in brackets follows from neglecting the masses and some trig identities.

Note that dimensionally this all works out since the action, S , has dimensions of $[S] = [\text{length}]^4 [g\varphi^3]$, which follows from the action being the spacetime integral of $-g\varphi^3/4!$, and so since action is dimensionless (when $\hbar = 1$) and length and energy have inverse units of each other (when $c = 1$ also) we have $[g\varphi^3] = [\text{energy}]^4$. We know that $[(\partial\varphi)^2] = [\text{energy}]^4$, and so we have $[\varphi] = [\text{energy}]$, since $[\partial] = [\text{length}]^{-1} = [\text{energy}]$. Hence, we must have that $[g] = [\text{energy}]$ and so this entire quantity has units of $[g^4]/[s^3] = [\text{energy}]^4/([\text{momentum}]^2)^3 = [\text{momentum}]^{-2}$, since $[\text{energy}] = [\text{momentum}]$ (when $c = 1$). This tells us that the cross section drops off inversely proportionally to the momentum squared, so at high momentums interactions are much less likely, which is what we would expect, high momentum particles move past each other very quickly without much of a chance for interactions.

Part III

Complex Scalar Fields

Twelve

Charge

Classical fields, such as the electromagnetic field, are always real since they are observables. Wave functions on the other hand are always complex, since their time evolution, $|\psi, t\rangle = e^{iEt}|\psi\rangle$, demands it. In quantum field theory fields can be either real or complex, since they are not observable. However, the action, Lagrangian, and Hamiltonian must be real, this is needed to ensure that the S matrix is unitary, which is required for conservation of probability. Also, the Hamiltonian corresponds to the energy of the system which is observable, and hence real.

12.1 Multiple Real Fields

Consider the common definition of complex numbers which starts something like “a complex number is an ordered pair of real numbers with operations of ...”. Much like this definition a complex field can be thought of as a pair of real fields. So, we need to be able to deal with multiple fields. Fortunately our work so far generalises very well to multiple real fields, we just have to sum over fields in certain expressions.

Suppose we have N real scalar fields, φ_r , with $r = 1, \dots, N$. For simplicity, and because it's all we need for the complex case, we assume that all N fields correspond to particles of mass m . The Lagrangian is then

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi_r(x))(\partial^\mu \varphi_r(x)) - m^2 \varphi_r(x)^2, \quad (12.1.1)$$

where we have an implied sum over r . We can then define the conjugate fields, and we find

$$\pi_r(x) := \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_r(x)} = \dot{\varphi}_r(x). \quad (12.1.2)$$

Any two distinct fields commute, and for two of the same field we have the usual commutation relations. That is, we have the equal time commutation relation

$$[\varphi_r(t, \mathbf{x}), \pi_s(t, \mathbf{x}')] = i\delta_{rs}\delta(\mathbf{x} - \mathbf{x}'), \quad (12.1.3)$$

and all other commutators vanish.

The mode expansion is the same as for a single field, but we have independent creation and annihilation operators for each field:

$$\varphi_r(x) = \int \frac{d^3[p]}{(2\pi)^3} \frac{1}{2\pi(\mathbf{p})} [a_r(\mathbf{p})e^{-ip \cdot x} + a_r^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (12.1.4)$$

These have the covariant commutation relations

$$[a_r(\mathbf{p}), a_s^\dagger(\mathbf{p}')] = \delta_{rs} 2\omega(\mathbf{p})(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}'), \quad (12.1.5)$$

and all other commutators vanish.

From here we proceed as before, we can define the energy-momentum tensor,

$$T^{\mu\nu}(x) := :(\partial^\mu \varphi_r(x))(\partial^\nu \varphi_r(x)) - \eta^{\mu\nu} \mathcal{L}(x):, \quad (12.1.6)$$

with an implied sum over r . We then get the conserved momentum

$$P^\nu = \int d^3x T^{0\nu}(x), \quad (12.1.7)$$

and the conserved current

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_r)} \delta \varphi_r, \quad (12.1.8)$$

with an implied sum over r .

The interpretation of $a_r^\dagger(\mathbf{p})$ and $a_r(\mathbf{p})$ as creation and annihilation operators is the same, each creating/annihilating particles of type r . This follows since we have

$$[H, a_r^\dagger(\mathbf{p})] = E a_r^\dagger(\mathbf{p}), \quad \text{and} \quad [H, a_r(\mathbf{p})] = -E a_r(\mathbf{p}). \quad (12.1.9)$$

So, $a_r^\dagger(\mathbf{p})$ creates a particle of type r and momentum \mathbf{p} , which we might write as

$$a_r^\dagger(\mathbf{p})|0\rangle = |\mathbf{p}, r\rangle. \quad (12.1.10)$$

The only time when we *don't* sum over the fields is when defining the number operator, which we want to count particles of type r , not the total number of particles, so we define

$$N_r(\mathbf{p}) := a_r^\dagger(\mathbf{p})a_r(\mathbf{p}) \quad (\text{no sum on } r). \quad (12.1.11)$$

12.2 Complex Scalar Field

Now, consider a complex scalar field, φ . We can write this as a sum of two real fields, φ_1 and φ_2 , one multiplied by i :

$$\varphi(x) = \frac{1}{\sqrt{2}}[\varphi_1(x) + i\varphi_2(x)]. \quad (12.2.1)$$

The factor of $1/\sqrt{2}$ here is so that the free Lagrangian comes out as the sum of two free Lagrangians for the real fields φ_1 and φ_2 .

Motivated by the need for the Lagrangian to be real we take the usual real scalar field Lagrangian and replace one of every pair of φ s with φ^\dagger . We then treat φ and φ^\dagger as two independent fields, since from these we have

$$\varphi_1 = \frac{1}{\sqrt{2}}(\varphi + \varphi^\dagger), \quad \text{and} \quad i\varphi_2 = \frac{1}{\sqrt{2}}(\varphi - \varphi^\dagger), \quad (12.2.2)$$

and we assume φ_1 and φ_2 are independent. Doing so we see that if we also drop the factor of 2 then we get

$$\mathcal{L} = (\partial_\mu \varphi^\dagger)(\partial^\mu \varphi) - m^2 \varphi^\dagger \varphi \quad (12.2.3)$$

$$= \frac{1}{2}(\partial_\mu \varphi_i)(\partial^\mu \varphi_i) - \frac{1}{2}m^2 \varphi_i^2, \quad (12.2.4)$$

with an implied sum over i . Note that we assume both fields have particles of the same mass, since otherwise we cannot guarantee that the Lagrangian will be real.

We can define the conjugate field

$$\pi := \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \dot{\varphi}^\dagger, \quad \text{and} \quad \pi^\dagger = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^\dagger} = \dot{\varphi}. \quad (12.2.5)$$

We then have the equal time commutation relations

$$[\varphi(t, \mathbf{x}), \pi(t, \mathbf{x}')] = [\varphi^\dagger(t, \mathbf{x}), \pi^\dagger(t, \mathbf{x}')] = i\delta^3(\mathbf{x} - \mathbf{x}'), \quad (12.2.6)$$

and all other commutators vanish.

We can define the Hamiltonian as before:

$$H = \int d^3\mathbf{x} (\pi \dot{\varphi} + \pi^\dagger \dot{\varphi}^\dagger - \mathcal{L}) \quad (12.2.7)$$

$$= \int d^3\mathbf{x} (\pi^\dagger \pi + (\nabla \varphi^\dagger) \cdot (\nabla \varphi) + m^2 \varphi^\dagger \varphi). \quad (12.2.8)$$

This second form makes it clear that this is real.

Computing the Heisenberg equations of motion gives

$$(\partial^2 + m^2)\varphi = 0, \quad \text{and} \quad (\partial^2 + m^2)\varphi^\dagger = 0. \quad (12.2.9)$$

12.3 Mode Expansion

The mode expansion is similar to the real case, except, in the real case we needed a and a^\dagger to give a real field. We don't need a real field here, but we do require a second term since this then causes things to cancel when the canonical commutation relations are evaluated outside of the light cone, which enforces causality. As a result we still include two terms in the mode expansion, but we no longer require one to be the Hermitian conjugate of the other, so we introduce some other operator, b^\dagger , to replace a^\dagger , with the conjugate for consistency with the real case:

$$\varphi(x) = \int \tilde{d}p [a(\mathbf{p})e^{-ip \cdot x} + b^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (12.3.1)$$

We then have

$$\varphi^\dagger(x) = \int \tilde{d}p [b(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (12.3.2)$$

We then have the equal time commutation relations

$$[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = [b(\mathbf{p}), b^\dagger(\mathbf{p}')] = \delta(\mathbf{p} - \mathbf{p}'), \quad (12.3.3)$$

and the two independent mode operators commute:

$$[a(\mathbf{p}), b(\mathbf{p})] = [a^\dagger(\mathbf{p}), b^\dagger(\mathbf{p})] = 0, \quad (12.3.4)$$

as we as all other commutators vanishing.

We can then show that

$$[H, a^\dagger(\mathbf{p})] = Ea^\dagger \quad [H, a(\mathbf{p})] = -Ea(\mathbf{p}), \quad (12.3.5)$$

$$[H, b^\dagger(\mathbf{p})] = Eb^\dagger \quad [H, b(\mathbf{p})] = -Eb(\mathbf{p}). \quad (12.3.6)$$

Recall that a real scalar field can be interpreted as an infinite set of harmonic oscillators, one for each momentum, \mathbf{p} . We can then interpret the complex scalar field as two independent sets of harmonic oscillators, with two for each momentum, \mathbf{p} . This leads to the interpretation of a^\dagger and b^\dagger as creating independent particles. These particles are bosons, both with respect to exchanging particles made by the same mode operator, so, for example, both created by a^\dagger , and with respect to exchanging particles made by different mode operators, so for example, exchanging a particle created by a^\dagger with a particle created by b^\dagger .

The energy-momentum operator is

$$T^{\mu\nu} = :(\partial^\mu \varphi^\dagger)(\partial^\nu \varphi) + (\partial^\nu \varphi^\dagger)(\partial^\mu \varphi) - \eta^{\mu\nu} \mathcal{L} :. \quad (12.3.7)$$

The conserved momentum is

$$P^\nu = \int d^3\mathbf{x} T^{0\nu} = \int \mathbf{p} p^\nu (a^\dagger(\mathbf{p})a(\mathbf{p}) + b^\dagger(\mathbf{p})b(\mathbf{p})) = \int \mathbf{p} p^\nu (N_a(\mathbf{p}) + N_b(\mathbf{p})).$$

This means it is the momentum of the combined two sets of particles which is conserved, not the momentum of each set individually.

12.4 Charge Conservation

There must be something that sets apart particles created by a^\dagger and particles created by b^\dagger . In this section we work out what this is. Consider the Lagrangian again,

$$\mathcal{L} = (\partial_\mu \varphi^\dagger)(\partial^\mu \varphi) - m^2 \varphi^\dagger \varphi. \quad (12.4.1)$$

Notice that this is invariant under the transformation

$$\varphi \mapsto e^{i\alpha} \varphi \quad (12.4.2)$$

for some constant α . Notice that α cannot depend on x , since then the derivatives would cause problems. Under this transformation we have $\varphi^\dagger \mapsto e^{-i\alpha} \varphi^\dagger$ as well. This is a U(1) symmetry¹. This is an internal symmetry, acting on the fields, rather than on spacetime.

A symmetry implies a conserved quantity, so lets find out what it is². Assume $\alpha \ll 1$, so we have $\exp[i\alpha] \approx 1 + i\alpha$, and hence

$$\varphi \mapsto e^{i\alpha} \varphi \approx (1 + i\alpha) \varphi = \varphi + i\alpha \varphi, \quad (12.4.3)$$

and hence

$$\delta \varphi = e^{i\alpha} \varphi - \varphi \approx i\alpha \varphi. \quad (12.4.4)$$

¹see the *Symmetries of Quantum Mechanics* or *Symmetries of Particles and Fields* course for details

²If the title of this section hasn't already given it away!

Similarly,

$$\delta\varphi^\dagger \approx -i\alpha\varphi. \quad (12.4.5)$$

We then have

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi = i\alpha(\partial^\mu\varphi^\dagger)\varphi, \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^\dagger)}\delta\varphi^\dagger = -i\alpha\varphi^\dagger(\partial^\mu\varphi). \quad (12.4.6)$$

³actually, normal ordering isn't strictly necessary here, but it is in other similar cases so we'll do so here for consistency

The conserved current is then, normal ordering to avoid later issues³,

$$j^\mu = : \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^\dagger)}\delta\varphi^\dagger : \quad (12.4.7)$$

$$= : i\alpha(\partial^\mu\varphi^\dagger)\varphi - i\alpha\varphi^\dagger(\partial^\mu\varphi) :. \quad (12.4.8)$$

We can rescale this by a factor of $-\alpha$ to get:

$$j^\mu = i : \varphi^\dagger(\partial^\mu\varphi) - (\partial^\mu\varphi^\dagger)\varphi : \quad (12.4.9)$$

Notice that we can easily show this is conserved, that is that $\partial_\mu j^\mu = 0$:

$$\partial_\mu j^\mu = i : [(\partial_\mu\varphi^\dagger)(\partial^\mu\varphi) + \varphi^\dagger(\partial^2\varphi) - (\partial^2\varphi^\dagger)\varphi - (\partial^\mu\varphi^\dagger)(\partial_\mu\varphi)] : \quad (12.4.10)$$

$$= i : [\varphi^\dagger(\partial^2\varphi) - (\partial^2\varphi^\dagger)\varphi] :. \quad (12.4.11)$$

The equations of motion for the fields give $\partial^2\varphi = -m^2\varphi$ and $\partial^2\varphi^\dagger = -m^2\varphi^\dagger$, so

$$\partial_\mu j^\mu = i : [\varphi^\dagger(-m^2\varphi) - (-m^2\varphi^\dagger)\varphi] : = 0, \quad (12.4.12)$$

where we've used the normal ordering to allow us to swap the order of φ and φ^\dagger .

Along with this conserved current we have a conserved charge:

$$Q = \int d^3\mathbf{x} j^0 = i \int d^3\mathbf{x} : (\varphi^\dagger\dot{\varphi} - \dot{\varphi}^\dagger\varphi) :. \quad (12.4.13)$$

We can write φ and φ^\dagger in terms of mode operators, with the time derivatives bringing down a factor of $\pm i\omega(\mathbf{p})$, giving

$$Q = i \int d^3\mathbf{x} \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3p'}{(2\pi)^3} \left\{ [a^\dagger(\mathbf{p})e^{ip\cdot x} + b e^{-ip\cdot x}] i\omega(\mathbf{p}') [-a(\mathbf{p}')e^{-ip'\cdot x} + b^\dagger(\mathbf{p}')e^{ip'\cdot x}] : \right. \\ \left. - i\omega(\mathbf{p}) : [a^\dagger(\mathbf{p})e^{ip\cdot x} - b(\mathbf{p})e^{-ip\cdot x}] [a(\mathbf{p}')e^{-ip'\cdot x} + b^\dagger(\mathbf{p}')e^{ip'\cdot x}] : \right\}. \quad (12.4.14)$$

After expanding this, integrating over x to get Dirac deltas, and then using these to perform the p' integral we're left with

$$Q = \int \frac{d^3p}{(2\pi)^3} : (a^\dagger(\mathbf{p})a(\mathbf{p}) - b(\mathbf{p})b^\dagger(\mathbf{p})) : \quad (12.4.15)$$

$$= \int \frac{d^3p}{(2\pi)^3} (a^\dagger(\mathbf{p})a(\mathbf{p}) - b^\dagger(\mathbf{p})b(\mathbf{p})). \quad (12.4.16)$$

From this we can then show that the following commutation relations hold:

$$[Q, a^\dagger(\mathbf{p})] = a^\dagger \quad [Q, a(\mathbf{p})] = -a(\mathbf{p}), \quad (12.4.17)$$

$$[Q, b^\dagger(\mathbf{p})] = -b^\dagger \quad [Q, b(\mathbf{p})] = b(\mathbf{p}). \quad (12.4.18)$$

So finally we have a difference between a and b . This leads to the interpretation that

- $a^\dagger(\mathbf{p})$ creates a particle with momentum \mathbf{p} and “charge” +1,
- $b^\dagger(\mathbf{p})$ creates a particle with momentum \mathbf{p} and “charge” −1,
- $a(\mathbf{p})$ annihilates a particle with momentum \mathbf{p} and “charge” +1,
- $b(\mathbf{p})$ annihilates a particle with momentum \mathbf{p} and “charge” −1.

Here we write “charge”, since this may not be electrical charge, but it can be.

So, a complex scalar field gives a theory with a pair of charged particles. Conservation of Q implies that particles must always be created and destroyed in pairs. This all follows simply from recognising that the Lagrangian must be real and so must depend only on products like $\varphi^\dagger\varphi$.

There is another interpretation of this, which is that a^\dagger creates particles, and b^\dagger creates antiparticles (or vice versa). An antiparticle being a particle with the same mass but opposite charge. Then

- $a^\dagger(\mathbf{p})$ creates a particle with momentum \mathbf{p} and “charge” +1,
- $b^\dagger(\mathbf{p})$ creates an antiparticle with momentum \mathbf{p} and “charge” −1,
- $a(\mathbf{p})$ annihilates a particle with momentum \mathbf{p} and “charge” +1,
- $b(\mathbf{p})$ annihilates an antiparticle with momentum \mathbf{p} and “charge” −1.

Notice that destroying a particle of charge +1 is the same as creating a particle of charge −1 and vice versa.

For a real field we have $a = b$ and we can interpret this as the statement that the particles in this theory are their own antiparticles.

12.5 Feynman Rules

It's useful for you to get used to confusing notation in this course, because it will get a whole lot worse.

Richard Ball

Consider the mode expansions of the fields:

$$\varphi = \int \mathrm{d}p [a(\mathbf{p})e^{-ip \cdot x} + b^\dagger(\mathbf{p})e^{ip \cdot x}], \quad (12.5.1)$$

$$\varphi^\dagger = \int \mathrm{d}p [b(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (12.5.2)$$

Using these and the commutation relations for the mode operators we can derive the covariant commutator

$$[\varphi(x), \varphi^\dagger(x')] = i\Delta(x - x') \quad (12.5.3)$$

where Δ is the same function as for real scalar fields. The other commutators, $[\varphi(x), \varphi(x')]$ and $[\varphi^\dagger(x), \varphi^\dagger(x')]$, vanish.

A result of this vanishing of commutators between φ s or between φ^\dagger s is that contractions between two φ s or two φ^\dagger s also vanish, and so the only nonvanishing contractions are

$$\overline{\varphi(x)\varphi^\dagger(x')} = \langle 0 | T[\varphi(x)\varphi^\dagger(x')] | 0 \rangle = i\Delta_F(x - x'), \quad (12.5.4)$$

where Δ_F is the same as for a real scalar field.

As before our state space is given by direct products of single particle states, which are now labelled with the charge as well as the momentum of the particle:

$$|\mathbf{p}, +\rangle = a^\dagger(\mathbf{p})|0\rangle, \quad \text{and} \quad |\mathbf{p}, -\rangle = b^\dagger(\mathbf{p})|0\rangle. \quad (12.5.5)$$

⁴We lower the + and - labels to make room for \dagger . Writing $\varphi = \varphi_+ + \varphi_-$ as before⁴ we have

$$\varphi_+(x)|\mathbf{p}, +\rangle = e^{-ip \cdot x}|0\rangle \quad \varphi_+^\dagger(x)|\mathbf{p}, -\rangle = e^{-ip \cdot x}|0\rangle, \quad (12.5.6)$$

$$\langle \mathbf{p}, + | \varphi_-^\dagger(x) = e^{ip \cdot x} \langle 0 | \quad \langle \mathbf{p}, - | \varphi_-(x) = e^{ip \cdot x} \langle 0 |. \quad (12.5.7)$$

So, we conclude that

- φ_+ destroys an incoming particle of charge +1,
- φ_+^\dagger destroys an incoming particle of charge -1,
- φ_- creates an outgoing particle of charge -1,
- φ_-^\dagger creates an outgoing particle of charge +1.

From this we can compute all the Feynman rules for any given interaction.

Thirteen

Interactions and Symmetries

13.1 Types of Interaction

We choose to look for interactions which conserve charge. While we can easily construct theories where this isn't the case charge conservation is a common phenomena and we wish our theories to exhibit it. This means we can no longer have a φ^3 interaction, since this would not conserve charge. Even if we consider a $\varphi^3 + (\varphi^\dagger)^3$ interaction, so the Lagrangian is real, we still don't get charge conservation, since the first term transforms as $\varphi^3 \mapsto e^{3i\alpha}\varphi^3$, and the second as $(\varphi^\dagger)^3 \mapsto e^{-3i\alpha}(\varphi^\dagger)^3$. The only way to have a theory conserve charge is if every term in the Lagrangian has as many factors of φ as it does φ^\dagger . There are a few ways to do this.

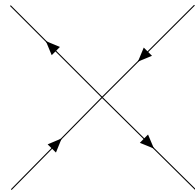
13.1.1 Four Point Interactions

Instead of the three point interaction, φ^3 , we can consider a four point interaction, $(\varphi^\dagger\varphi)^2$, where interactions involve four particles at once, this would lead to the interaction Lagrangian

$$\mathcal{L}_1^{(4)} = -\frac{\lambda}{2}(\varphi^\dagger\varphi)^2. \quad (13.1.1)$$

To keep track of charge in our Feynman diagrams it is conventional to put an arrow on our Feynman diagrams pointing in the direction of charge flow. Note that charge flowing backwards in time is the same as negative charge flowing forward in time, that is antiparticles correspond to arrows pointing left and particles to arrows pointing right.

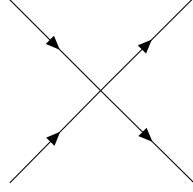
There are three vertices in this theory which are all slightly different, but have the same shape. The first is



(13.1.2)

This corresponds to a particle and antiparticle coming in and a particle and antiparticle going out. The net charge in and out is then 0. So charge is conserved in this vertex.

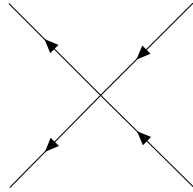
The next vertex is



(13.1.3)

This corresponds to two particles coming in and two particles going out. The net charge in and out is then 2, so charge is conserved in this vertex.

The final vertex is



(13.1.4)

This corresponds to two antiparticles coming in and two antiparticles going out. The net charge in and out is then -2 , so charge is conserved in this vertex.

As with the real case the constant factor in this interaction has been chosen such that every vertex gives a factor of $-i\lambda$.

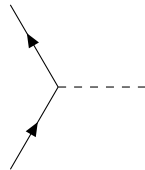
13.1.2 Three Point Interactions

If we introduce a new, neutral, real scalar field, Φ , then we can have a three point interaction of the form

$$\mathcal{L}_I^{(3)} = -y\Phi\varphi^\dagger\varphi. \quad (13.1.5)$$

This is known as a **Yukawa interaction**.

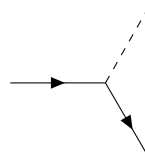
As with φ^3 theory we get three particles in each interaction. There are a variety of vertices we can create in this case. Denoting the neutral particle with a dashed line one possibility is



(13.1.6)

This corresponds to a particle and antiparticle annihilating into a neutral particle.

Another vertex is



(13.1.7)

This corresponds to a particle emitting a neutral particle. Similarly the vertex



corresponds to an antiparticle emitting a neutral particle. Notice that in all of these diagrams the total charge is conserved at each vertex.

As before the constant factors are chosen such that each vertex contributes a factor of $-iy$ to the amplitude.

A full Lagrangian in a theory with this interaction may take the form

$$\mathcal{L} = (\partial_\mu \varphi^\dagger)(\partial_\mu \varphi) - m^2 \varphi^\dagger \varphi + \frac{1}{2}(\partial_\mu \Phi)(\partial^\mu \Phi) - \frac{1}{2}M^2 \Phi^2 - y\Phi \varphi^\dagger \varphi, \quad (13.1.9)$$

corresponding to a neutral particle, call it s_0 , of mass M and charged (anti)particles, call them s_\pm , of mass m .

In many ways this theory is not that different from the real φ^3 theory. For example, we still have 2-2 scattering, such as $s_+s_- \rightarrow s_+s_-$, $s_+s_+ \rightarrow s_+s_+$, and $s_-s_- \rightarrow s_-s_-$. The exchange particle in these cases is a neutral scalar, s_0 . Note that in the $s_+s_- \rightarrow s_+s_-$ case, since the final two particles are distinct, there is no u -channel diagram. Similarly for $s_\pm s_\pm \rightarrow s_\pm s_\pm$ there is no s -channel diagram, since the intermediate particle would have to carry a charge of ± 2 .

There is one sort of interaction which we don't have in φ^3 theory, and that is decays. If $M > 2m$ then it is possible to have the interaction $s_0 \rightarrow s_+s_-$. This would correspond to an $n = 1$ contribution to the Dyson series, and these terms vanished for φ^3 theory. This would look like



The amplitude for this is $\mathcal{M} = -iy$. We can work out the decay rate starting with

$$\left(\frac{d\Gamma}{d\Omega} \right)_{\text{CoM}} = \frac{1}{32\pi^2} \frac{p}{M^2} |\mathcal{M}|^2, \quad (13.1.11)$$

where p is the magnitude of the three-momentum of the two charged particles in the centre of mass frame. The amplitude squared will give $|\mathcal{M}|^2 = |-iy|^2 = y^2$. We then have

$$\left(\frac{d\Gamma}{d\Omega} \right)_{\text{CoM}} = \frac{y^2}{64\pi^2} \frac{\sqrt{M^2 - 4m^2}}{M^2}. \quad (13.1.12)$$

So we see that we must have $M > 2m$ for this to be nonzero and real, assuming $M, m \geq 0$.

We just said “assuming $M, m \geq 0$ ”, but what if they aren't? In fact, what if they aren't even real? We can analytically extend this decay rate as a function of the

complex variable M . Recall that $\sqrt{f(z)}$ has a branch cut from $f(z) = 0$ to infinity. So the decay rate has a branch cut from $M = 2m$ to infinity. This is part of a general rule, particles correspond to poles, such as the pole at $m = \sqrt{s}$ in $1/(s - m^2)$, and decays correspond to branch cuts, such as the one starting at $M = 2m$ here.

13.2 Symmetries

In this section we'll study the effect of parity, time reversal, and charge conjugation on scalar fields. We'll start from their action on spacetime, and then derive their action on the fields and mode operators. These three symmetries are all \mathbb{Z}_2 symmetries, meaning that, apart from doing nothing, we have a single operation and applying this operation twice is the same as doing nothing. The technical word for this is that these symmetries are **idempotent**.

13.2.1 Parity

Parity acts on spacetime by sending \mathbf{x} to $-\mathbf{x}$. That is

$$\mathcal{P} : x^\mu = (t, \mathbf{x}) \mapsto \bar{x}^\mu = (t, -\mathbf{x}). \quad (13.2.1)$$

Since the direction of space is also changed the momentum is also changed by sending \mathbf{p} to $-\mathbf{p}$:

$$\mathcal{P} : p^\mu = (E, \mathbf{p}) \mapsto \bar{p}^\mu = (E, -\mathbf{p}). \quad (13.2.2)$$

We will use the notation \bar{x} and \bar{p} with this meaning throughout this section, including when discussing time reversal and charge conjugation.

The definition of parity above is classical. In quantum field theory we have operators and states. Parity is promoted to be a unitary operator, \mathcal{P} , on the states, that is $\mathcal{P}^\dagger \mathcal{P} = 1$. Starting with the simplest case, the vacuum state should be invariant under parity, since there's nothing there to change:

$$\mathcal{P}|0\rangle = |0\rangle. \quad (13.2.3)$$

The next simplest state is a single particle state, $|\mathbf{p}\rangle$. We expect that the momentum of this particle should be reversed under parity:

$$\mathcal{P}|\mathbf{p}\rangle = |-\mathbf{p}\rangle. \quad (13.2.4)$$

We can write this momentum state as $|\mathbf{p}\rangle = a^\dagger(\mathbf{p})|0\rangle$, and so we can identify the action of the parity operator on the state with an equivalent action on $a^\dagger(\mathbf{p})$. Since $a^\dagger(\mathbf{p})$ is an operator it transforms with two copies of \mathcal{P} as

$$\mathcal{P}a^\dagger(\mathbf{p})\mathcal{P}^\dagger = a^\dagger(-\mathbf{p}), \quad (13.2.5)$$

so

$$(\mathcal{P}a^\dagger(\mathbf{p})\mathcal{P}^\dagger)|0\rangle = a^\dagger(-\mathbf{p})|0\rangle = |-\mathbf{p}\rangle. \quad (13.2.6)$$

Taking the adjoint of this we have

$$[\mathcal{P}a^\dagger(\mathbf{p})\mathcal{P}^\dagger]^\dagger = (\mathcal{P}^\dagger)^\dagger a(\mathbf{p})\mathcal{P}^\dagger = \mathcal{P}a(\mathbf{p})\mathcal{P}^\dagger, \quad (13.2.7)$$

and so

$$\mathcal{P}a(\mathbf{p})\mathcal{P}^\dagger = [a^\dagger(-\mathbf{p})]^\dagger = a(-\mathbf{p}). \quad (13.2.8)$$

To derive the effect on the field, which is an operator and so transforms as $\varphi \mapsto \mathcal{P}\varphi\mathcal{P}^\dagger$, we can expand φ in the mode operators:

$$\mathcal{P}\varphi(x)\mathcal{P}^\dagger = \int \mathfrak{d}p [\mathcal{P}a(\mathbf{p})\mathcal{P}^\dagger e^{-ip \cdot x} + \mathcal{P}a^\dagger(\mathbf{p})\mathcal{P}^\dagger e^{ip \cdot x}] \quad (13.2.9)$$

$$= - \int \mathfrak{d}p [a(-\mathbf{p})e^{-ip \cdot x} + a^\dagger(-\mathbf{p})e^{ip \cdot x}]. \quad (13.2.10)$$

Note that the parity operator doesn't act on the phase factors, even though they depend on position and momentum, they're just scalars and so aren't affected by operators on the Hilbert space. Now take $\mathbf{p} \rightarrow -\mathbf{p}$ in the integral, this doesn't change the measure or region of integration and takes $p \rightarrow \bar{p}$, giving

$$\mathcal{P}\varphi(x)\mathcal{P}^\dagger = \int \mathfrak{d}p [a(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (13.2.11)$$

Now notice that $\bar{p} \cdot x = Et + \mathbf{p} \cdot \mathbf{x} = p \cdot \bar{x}$, and so

$$\mathcal{P}\varphi(x)\mathcal{P}^\dagger = \int \mathfrak{d}p [a(\mathbf{p})e^{-ip \cdot \bar{x}} + a^\dagger(\mathbf{p})e^{ip \cdot \bar{x}}] \quad (13.2.12)$$

$$= \varphi(\bar{x}). \quad (13.2.13)$$

So under parity the field transforms as

$$\mathcal{P} : \varphi(x) \mapsto \mathcal{P}\varphi(x)\mathcal{P}^\dagger = \varphi(\bar{x}). \quad (13.2.14)$$

This is what we would expect. Note that exactly the same argument can be applied to the conjugate field, and to complex scalar fields, they all transform the same.

One thing that we didn't consider is that quantum states are only defined up to phase. This means that its possible to have an equivalent transformation which introduces a phase, so we could have that parity acts on the single particle state as

$$\mathcal{P}|\mathbf{p}\rangle = \eta_p|-\mathbf{p}\rangle. \quad (13.2.15)$$

Since \mathcal{P} is idempotent we have $\mathcal{P}\mathcal{P}|\mathbf{p}\rangle = \mathcal{P}\eta_p|-\mathbf{p}\rangle = \eta_p^2|\mathbf{p}\rangle$, and we must have $\mathcal{P}\mathcal{P}|\mathbf{p}\rangle = |\mathbf{p}\rangle$, so $\eta_p^2 = 1$, meaning $\eta_p = \pm 1$. Above we assumed $\eta_p = 1$, suppose instead that $\eta_p = -1$, then the mode operators must transform as

$$\mathcal{P} : a(\mathbf{p}) \mapsto \mathcal{P}a(\mathbf{p})\mathcal{P}^\dagger = -a(-\mathbf{p}), \quad (13.2.16)$$

and hence the field transforms as

$$\mathcal{P} : \varphi(x) \mapsto \mathcal{P}\varphi(x)\mathcal{P}^\dagger = -\varphi(\bar{x}). \quad (13.2.17)$$

In this case we say that φ is a **pseudoscalar** field.

13.2.2 Time Reversal

Time reversal acts on spacetime by sending t to $-t$. That is

$$\mathcal{T}: x^\mu = (t, \mathbf{x}) \mapsto -\bar{x}^\mu = (-t, \mathbf{x}). \quad (13.2.18)$$

Since the direction of time reverses the direction of momentum, proportional to the velocity, \dot{x} , and hence first order in time derivatives, changes direction. The energy is unaffected. The simplest explanation for why is that the kinetic energy is proportional to \dot{x}^2 and so picks up two negatives upon time reversal, cancelling out for no overall change. So, momentum transforms under time reversal as

$$\mathcal{T}: p^\mu = (E, \mathbf{p}) \mapsto \bar{p}^\mu = (E, -\mathbf{p}). \quad (13.2.19)$$

Again when moving to quantum field theory time reversal is promoted to an operator, \mathcal{T} on the states. One effect of time reversal is that the initial and final state are swapped. This means that amplitudes change:

$$\mathcal{T}: \langle f|i \rangle \mapsto \langle i|f \rangle = \langle f|i \rangle^*. \quad (13.2.20)$$

So, time reversal means we have to take the complex conjugate of everything. Another way to view this is to consider the time evolution operator, $\exp[-iEt]$. Under time reversal this changes to $\exp[iEt] = \exp[-iEt]^*$. We say that \mathcal{T} is an antiunitary operator.

Definition 13.2.21 — Antiunitary Operator Let \mathbb{H} be a complex Hilbert space with inner product $\langle -, - \rangle$, and let $U: \mathbb{H} \rightarrow \mathbb{H}$ be an operator. Then U is an **antiunitary operator** if

$$\langle Ux, Uy \rangle = \langle x, y \rangle^* \quad (13.2.22)$$

for all $x, y \in \mathbb{H}$.

This means that under time reversal the exponential phase factor transforms as

$$\mathcal{T}: e^{ip \cdot x} \mapsto e^{(-i)\bar{p} \cdot (-\bar{x})} = e^{i\bar{p} \cdot \bar{x}} = e^{ip \cdot x}. \quad (13.2.23)$$

Consider the transformation of a single particle state:

$$\mathcal{T}|\mathbf{p}\rangle = |-\mathbf{p}\rangle. \quad (13.2.24)$$

We must therefore have

$$\mathcal{T}a^\dagger(\mathbf{p})\mathcal{T}^\dagger = a^\dagger(-\mathbf{p}), \quad \text{and} \quad \mathcal{T}a(\mathbf{p})\mathcal{T}^\dagger = a(-\mathbf{p}). \quad (13.2.25)$$

The field then transforms as

$$\mathcal{T}\varphi(x)\mathcal{T}^\dagger = \int \bar{d}p [a(-\mathbf{p})e^{ip \cdot x} + a^\dagger(-\mathbf{p})e^{-ip \cdot x}] \quad (13.2.26)$$

$$= \int \bar{d}p [a(\mathbf{p})e^{i\bar{p} \cdot \bar{x}} + a(\mathbf{p})e^{-i\bar{p} \cdot \bar{x}}] \quad (13.2.27)$$

$$= \int \bar{d}p [a(\mathbf{p})e^{-ip \cdot (-\bar{x})} + a(\mathbf{p})e^{ip \cdot (-\bar{x})}] \quad (13.2.28)$$

$$= \varphi(-\bar{x}). \quad (13.2.29)$$

Here we changed \mathbf{p} to $-\mathbf{p}$ in the second line. This means the field transforms as

$$\mathcal{T}: \varphi(x) \mapsto \mathcal{T}\varphi(x)\mathcal{T}^\dagger = \varphi(-\bar{x}). \quad (13.2.30)$$

Again, we can include a phase factor, η_t , which is restricted to be ± 1 . Also all of this logic applies to both the conjugate field and to complex scalar fields.

13.2.3 Charge Conjugation

Classically charge conjugation changes the sign of electric charge. So

$$\mathcal{C}: Q \mapsto -Q. \quad (13.2.31)$$

In quantum field theory this corresponds to exchanging particles and antiparticles. Hence, charge conjugation is only interesting for complex scalar fields. Charge conjugation acts on the single particle states according to

$$\mathcal{C}|\mathbf{p}, +\rangle = |\mathbf{p}, -\rangle, \quad \text{and} \quad \mathcal{C}|\mathbf{p}, -\rangle = |\mathbf{p}, +\rangle. \quad (13.2.32)$$

Exchanging particles and antiparticles is the same as exchanging mode operators $a \leftrightarrow b$:

$$\mathcal{C}: a(\mathbf{p}) \mapsto \mathcal{C}a(\mathbf{p})\mathcal{C}^\dagger = b(\mathbf{p}), \quad (13.2.33)$$

$$\mathcal{C}: a^\dagger(\mathbf{p}) \mapsto \mathcal{C}a^\dagger(\mathbf{p})\mathcal{C}^\dagger = b^\dagger(\mathbf{p}), \quad (13.2.34)$$

$$\mathcal{C}: b(\mathbf{p}) \mapsto \mathcal{C}b(\mathbf{p})\mathcal{C}^\dagger = a(\mathbf{p}), \quad (13.2.35)$$

$$\mathcal{C}: b^\dagger(\mathbf{p}) \mapsto \mathcal{C}b^\dagger(\mathbf{p})\mathcal{C}^\dagger = a^\dagger(\mathbf{p}). \quad (13.2.36)$$

The action on the field is then

$$\mathcal{C}\varphi(x)\mathcal{C}^\dagger = \int d\mathbf{p} [b(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}] = \varphi^\dagger(x), \quad (13.2.37)$$

so

$$\mathcal{C}: \varphi(x) \mapsto \mathcal{C}\varphi(x)\mathcal{C}^\dagger = \varphi^\dagger(x). \quad (13.2.38)$$

Similarly,

$$\mathcal{C}: \varphi^\dagger(x) \mapsto \mathcal{C}\varphi^\dagger(x)\mathcal{C}^\dagger = \varphi(x). \quad (13.2.39)$$

For a complex scalar field we also have the current, j^μ . Acting on this with charge conjugation exchanges φ^\dagger and φ , giving

$$\mathcal{C}j^\mu\mathcal{C}^\dagger = i\mathcal{C}:\varphi^\dagger\partial^\mu\varphi - \varphi\partial^\mu\varphi^\dagger:\mathcal{C}^\dagger \quad (13.2.40)$$

$$= i:\varphi\partial^\mu\varphi^\dagger - \varphi^\dagger\partial^\mu\varphi: \quad (13.2.41)$$

$$= -j^\mu, \quad (13.2.42)$$

which is what we would expect. Since the total charge is defined as the spatial integral of j^0 we also have

$$\mathcal{C}Q\mathcal{C}^\dagger = -Q, \quad (13.2.43)$$

which is the same as the classical action of charge conjugation.

13.2.4 Acting on the Lagrangian

The free Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2 \varphi^2 \quad (13.2.44)$$

for a real scalar field, or

$$\mathcal{L} = (\partial_\mu \varphi^\dagger)(\partial^\mu \varphi) - m^2 \varphi^\dagger \varphi \quad (13.2.45)$$

for a complex scalar field. In either case the result is invariant under \mathcal{P} , \mathcal{T} , and \mathcal{C} .

Now suppose that the interaction Lagrangian is also invariant under \mathcal{P} , \mathcal{T} and \mathcal{C} . Then, under \mathcal{P} or \mathcal{C} the amplitude transforms as

$$\mathcal{P}: \mathcal{M} \mapsto \mathcal{M}, \quad (13.2.46)$$

$$\mathcal{C}: \mathcal{M} \mapsto \mathcal{M}. \quad (13.2.47)$$

That is, the amplitude is unchanged. Under \mathcal{T} the amplitude transforms as

$$\mathcal{T}: \mathcal{M} \mapsto \mathcal{M}^*. \quad (13.2.48)$$

However, in all three cases $|\mathcal{M}|^2$ is unchanged and so the physics doesn't change.

Invariance of the interaction term under these symmetries is not a given. For example, a φ^3 interaction with a pseudoscalar field gains a minus sign under parity. However, for the purposes of this course we will only consider theories invariant under \mathcal{P} , \mathcal{T} and \mathcal{C} separately. In *Gauge Theories in Particle Physics* we will consider theories which aren't invariant under these symmetries individually, but instead are invariant under the combined $\mathcal{CP}\mathcal{T}$ symmetry.

Part IV

Dirac Equation

Fourteen

Deriving the Dirac Equation

In this chapter we'll derive the Dirac equation in the way that Dirac did historically. He was motivated by looking for an equation without the negative energies of the Klein–Gordon equation. He failed to find such an equation, but it turns out that the negative energy solutions are actually important, and not just a problem to be eliminated.

14.1 The Dirac Equation

To avoid having to take square roots when computing the energy, which is the source of the negative energy terms in the Klein–Gordon equation, Dirac looked for a wave equation linear in ∂_t . For this to be relativistic Dirac knew that the equation would also have to be linear in ∂_x . Dirac also imposed one further constraint. The equation he was deriving should imply the Klein–Gordon equation since this is how we get wave equations with the relativistic energy momentum relation.

Dirac started with the ansatz

$$\frac{\partial \psi}{\partial t} + \boldsymbol{\alpha} \cdot \nabla \psi + im\beta \psi = 0, \quad (14.1.1)$$

which can be written in a Schrödinger-like form as

$$i \frac{\partial \psi}{\partial t} = H\psi, \quad \text{where} \quad H = -i\boldsymbol{\alpha} \cdot \nabla + \beta m = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m \quad (14.1.2)$$

is the **Dirac Hamiltonian**. It should be noted that this is *not* actually the Hamiltonian for a system satisfying the Dirac equation.

Looking for real energies Dirac posited that H should be Hermitian, and hence $\boldsymbol{\alpha}$ and β should be real. Next Dirac looked for other conditions on $\boldsymbol{\alpha}$ and β which would enforce this equation implying the Klein–Gordon equation. Noticing that the Klein–Gordon equation has second derivatives we need to find a way to extract second derivatives from the Dirac equation. To do this we square the operator, but in such a way that the cross terms cancel, so we compute the product of the operator and its adjoint:

$$\left(\frac{\partial}{\partial t} - \boldsymbol{\alpha} \cdot \nabla - im\beta \right) \left(\frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \nabla + im\beta \right) \psi = 0. \quad (14.1.3)$$

Expanding this we have

$$\left(\frac{\partial^2}{\partial t^2} - \alpha_i \alpha_j \frac{\partial^2}{\partial x_i \partial x_j} + m^2 \beta^2 - im(\alpha_i \beta + \beta \alpha_i) \frac{\partial}{\partial x_i} \right) \psi = 0. \quad (14.1.4)$$

Looking at this we see that if $\alpha_i \alpha_j = 1$ the cross term vanishes then we get the Klein–Gordon equation. The symmetry of $\partial^2/\partial x_i \partial x_j$ implies that we should look for $\alpha_i \alpha_j$ to be symmetric, so in terms of setting $\alpha_i \alpha_j$ to 1 we impose

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij} \quad (14.1.5)$$

where $\{A, B\} := AB + BA$ is the **anticommutator**. Similarly to get the Klein–Gordon equation we need $\beta^2 = 1$. The cross term will vanish if

$$\{\alpha_i, \beta\} = 0 \quad (14.1.6)$$

Clearly these requirements cannot be satisfied if α_i and β are numbers. We can have relations like these if α_i and β are *matrices* though. Looking for H to be Hermitian means that α_i and β must be *Hermitian* matrices. Suppose α_i and β are $N \times N$ matrices. Then ψ must be an object with N components.

We can go on to derive further properties of α_i and β which will help us to narrow down what matrices they are. First, $\alpha_i^2 = \beta^2 = 1$ the eigenvalues must square to 1, so the eigenvalues of α_i and β are ± 1 .

Next consider the trace of α_i . Using the anticommutation relations for α_i and β we have

$$\text{tr } \alpha_i = \text{tr}(\beta^2 \alpha_i) \quad \beta^2 = 1 \quad (14.1.7)$$

$$= \text{tr}(\beta \alpha_i \beta) \quad \text{tr is cyclic} \quad (14.1.8)$$

$$= \text{tr}(-\alpha_i \beta^2) \quad \alpha_i \text{ and } \beta \text{ anticommute} \quad (14.1.9)$$

$$= -\text{tr}(\alpha_i \beta^2) \quad \text{tr is linear} \quad (14.1.10)$$

$$= -\text{tr } \alpha_i \quad \beta^2 = 1 \quad (14.1.11)$$

So we can conclude that $\text{tr } \alpha_i = -\text{tr } \alpha_i$ and so $\text{tr } \alpha_i = 0$. We can do exactly the same swapping β and α_i and so $\text{tr } \beta = 0$. This means that α_i and β are *traceless* Hermitian matrices.

In the basis where α_i is diagonal the diagonal is just the eigenvalues, and so the trace is the sum of the eigenvalues. The trace is basis independent so the trace is *always* the sum of eigenvalues. The fact that α_i and β have eigenvalues ± 1 and the sum of the eigenvalues, that is the trace, must vanish means that α_i and β must have an even number of eigenvalues, and hence α_i and β are traceless Hermitian matrices of *even dimension*.

At this point we can just start trying to find a solution for α_i and β . We start with the simplest case, $N = 2$. We know that the Pauli spin matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (14.1.12)$$

form a basis for the 2×2 traceless Hermitian matrices, and so we posit that $\alpha_i = \sigma_i$, or more accurately σ_i is a *representation*¹ of α_i . The problem comes when we look for a 2×2 representation of β . There is no other 2×2 Hermitian traceless matrix, and so we cannot find a value for β in the 2×2 matrices. This may not be a problem. If we're looking at a massless particle then $m = 0$ gets rid of the β term, so we don't need a value for β and we can use 2×2 matrices and 2 component spinors to describe such particles.

¹see *Symmetries of Quantum Mechanics* or *Symmetries of Particles and Fields* for details on what a representation is.

Move on to the next simplest case, $N = 4$, here there are solutions. One very common representation is the **Dirac representation**, which is chosen such that β is diagonal:

$$\beta = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix}, \quad \text{and} \quad \alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}. \quad (14.1.13)$$

We call these the **Dirac matrices** in this representation.

These are complex, meaning that ψ is an object with four complex components. We call ψ a **spinor**, for reasons we'll see later. As a four component thing it may be more accurate to write ψ with an index, say as ψ_a , but spinor indices are traditionally suppressed unless there is ambiguity. When we do write spinor indices we'll choose values from the start of the alphabet, a, b, c, \dots , as opposed to the Greek letters, μ, ν, ρ, \dots we use for Lorentz indices or i, j, k, \dots for spatial indices.

14.2 Covariant Form

The Dirac equation as we've written it so far is not *manifestly* covariant. We can fix this by changing up our notation slightly. Start by multiplying through by $i\beta$ to get

$$\left(i\beta \frac{\partial}{\partial t} + i\beta \alpha_j \frac{\partial}{\partial x_j} - m \right) \psi = 0, \quad (14.2.1)$$

where we've used $\beta^2 = 1$, although really we now know that "1" here is an identity matrix, which we leave implicit. Now define $\gamma^0 = \beta$ and $\gamma^i = \beta \alpha^i$. These form a four-vector, $\gamma^\mu = (\gamma^0, \gamma^i) = (\beta, \beta \alpha)$. We can then write

$$\left(i\gamma^0 \frac{\partial}{\partial x_0} + i\gamma^j \frac{\partial}{\partial x_j} - m \right) \psi = 0, \quad (14.2.2)$$

or more compactly,

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \quad (14.2.3)$$

or even more compactly,

$$(i\not{\partial} - m)\psi = 0 \quad (14.2.4)$$

where $\not{\partial} := \gamma^\mu \partial_\mu$ is a shorthand called **slash notation**.

Like α_i and β these γ^μ are really $N \times N$ matrices, which we call the **gamma matrices** for obvious reasons. If we reinstate the spinor indices then we might write γ_{ab}^μ , where μ is an index telling us if we are considering $\gamma^0, \gamma^1, \gamma^2$, or γ^3 and the indices a and b tell us which component of the matrix we are considering.

We can determine the anticommutation relations of γ^μ using those of α_i and β , first, we have $\{\gamma^0, \gamma^0\} = \{\beta, \beta\} = 2\beta^2 = 2$, next $\{\gamma^0, \gamma^i\} = \{\beta, \beta \alpha^i\} = \beta^2 \alpha^i + \beta \alpha^i \beta = \beta^2 \alpha^i - \alpha^i \beta = 0$. Finally, $\{\gamma^i, \gamma^j\} = \{\beta \alpha^i, \beta \alpha^j\} = \beta \alpha^i \beta \alpha^j + \beta \alpha^j \beta \alpha^i = -\beta \alpha^i \alpha^j \beta - \beta \alpha^j \alpha^i \beta = -\beta \{\alpha^i, \alpha^j\} \beta = -\beta 2\delta^{ij} \beta = 2\delta^{ij}$. Hence,

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}. \quad (14.2.5)$$

This is called the **Clifford algebra** of the gamma matrices. We take this as the defining property of the gamma matrices, and it is this algebra which is being represented by a particular choice of matrices.

It is useful to define a **fifth gamma matrix**, $\gamma^5 = \gamma_5$. This is *not* part of the Clifford algebra. The 5 is just a label, it's not an index and no index will ever take the value of 5 (while we remain in 4 dimensions, even in more dimensions there are only ever four gamma matrices), we use the label 5 because some sources label the gamma matrices 1 through 4 instead of 0 through 3. This extra matrix is defined as

$$\gamma^5 = \gamma_5 := i\gamma^0\gamma^1\gamma^2\gamma^3. \quad (14.2.6)$$



The factor of i here is sometimes left out, in which case γ^5 will be anti-Hermitian.

Using the anticommutator relations we have $\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu$ whenever $\mu \neq \nu$, as well as $(\gamma^0)^2 = 1$ and $(\gamma^i)^2 = -1$, so

$$(\gamma^5)^2 = -\gamma^0\gamma^1\gamma^2\gamma^3\gamma^0\gamma^1\gamma^2\gamma^3 \quad (14.2.7)$$

$$= \gamma^1\gamma^2\gamma^3(\gamma^0)^2\gamma^1\gamma^2\gamma^3 \quad \gamma^0 \text{ past 3 } \gamma^\mu \quad (14.2.8)$$

$$= \gamma^1\gamma^2\gamma^3\gamma^1\gamma^2\gamma^3 \quad (\gamma^0)^2 = 1 \quad (14.2.9)$$

$$= \gamma^2\gamma^3(\gamma^1)^2\gamma^2\gamma^3 \quad \gamma^1 \text{ past 2 } \gamma^\mu \quad (14.2.10)$$

$$= -\gamma^2\gamma^3\gamma^2\gamma^3 \quad (\gamma^1)^2 = -1 \quad (14.2.11)$$

$$= \gamma^3(\gamma^2)^2\gamma^3 \quad \gamma^2 \text{ past 1 } \gamma^\mu \quad (14.2.12)$$

$$= -(\gamma^3)^2 \quad (\gamma^2)^2 = -1 \quad (14.2.13)$$

$$= 1 \quad (\gamma^3)^2 = -1. \quad (14.2.14)$$

We can also derive anticommutator relations for γ^5 with the other gamma matrices:

$$\{\gamma^5, \gamma^0\} = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^0 + i(\gamma^0)^2\gamma^1\gamma^2\gamma^3 \quad (14.2.15)$$

$$= -i(\gamma^0)^2\gamma^1\gamma^2\gamma^3 + i(\gamma^0)^2\gamma^1\gamma^2\gamma^3 \quad (14.2.16)$$

$$= 0, \quad (14.2.17)$$

$$\{\gamma^5, \gamma^1\} = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^1 + i\gamma^1\gamma^0\gamma^1\gamma^2\gamma^3 \quad (14.2.18)$$

$$= i\gamma^0(\gamma^1)^2\gamma^2\gamma^3 - i\gamma^0(\gamma^1)^2\gamma^2\gamma^3 \quad (14.2.19)$$

$$= 0, \quad (14.2.20)$$

$$\{\gamma^5, \gamma^2\} = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^2 + i\gamma^2\gamma^0\gamma^1\gamma^2\gamma^3 \quad (14.2.21)$$

$$= -i\gamma^0\gamma^1(\gamma^2)^2\gamma^3 + i\gamma^0\gamma^1(\gamma^2)^2\gamma^3 \quad (14.2.22)$$

$$= 0, \quad (14.2.23)$$

$$\{\gamma^5, \gamma^3\} = i\gamma^0\gamma^1\gamma^2(\gamma^3)^2 + i\gamma^3\gamma^0\gamma^1\gamma^2\gamma^3 \quad (14.2.24)$$

$$= i\gamma^0\gamma^1\gamma^2(\gamma^3)^2 - i\gamma^0\gamma^1\gamma^2(\gamma^3)^2 \quad (14.2.25)$$

$$= 0, \quad (14.2.26)$$

so in general

$$\{\gamma^5, \gamma^\mu\} = 0. \quad (14.2.27)$$

The Dirac representation for α_i and β also gives a representation of the gamma matrices, namely

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \text{and} \quad \gamma^5 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix}. \quad (14.2.28)$$

As with α_i and β this is just one representation of the gamma matrices, the real definition is just anything satisfying the Clifford algebra. In this representation γ^0 is Hermitian, $\gamma^{0\dagger} = \gamma^0$, γ^i are anti-Hermitian, $\gamma^{i\dagger} = -\gamma^i$, and γ^5 is hermitian, $\gamma^{5\dagger} = \gamma^5$.

We can now show that the Dirac equation implies the Klein–Gordon equation. Start with the Dirac equation,

$$(i\partial - m)\psi = 0. \quad (14.2.29)$$

Multiply by $i\partial + m$ to get

$$(i\partial + m)(i\partial - m)\psi = 0. \quad (14.2.30)$$

Expanding this the cross terms cancel and we get

$$(-\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + m^2)\psi = 0. \quad (14.2.31)$$

Since $\partial_\mu \partial_\nu$ is symmetric in μ and ν we can replace $\gamma^\mu \gamma^\nu$ with the symmetrised product of γ^μ and γ^ν , namely $\{\gamma^\mu, \gamma^\nu\}/2 = \eta^{\mu\nu}$, since any antisymmetric part of $\gamma^\mu \gamma^\nu$ vanishes in the product with something totally symmetric. Hence we can replace $\gamma^\mu \gamma^\nu$ with $\eta^{\mu\nu}$:

$$(\eta^{\mu\nu} \partial_\mu \partial_\nu + m^2)\psi = 0 \implies (\partial^2 + m^2)\psi = 0, \quad (14.2.32)$$

which is exactly the Klein–Gordon equation, so the Dirac equation does imply the Klein–Gordon equation as required.

14.3 Spinor Transformations

We’ve said that spinors are *not* four-vectors, despite having four components in the Dirac representation. The difference is in their transformation rules. Recall that a four-vector, x , is defined by transforming as

$$x \mapsto x' = \Lambda x \iff x^\mu = x'^\mu = \Lambda^\mu_\nu x^\nu \quad (14.3.1)$$

for some Lorentz transformation Λ . We now ask how a general spinor transforms under a Lorentz transformation. The Lorentz transformation acts on spacetime, not on the spinor, so we’re looking for some spinor, $\psi(x)$, to transform into another spinor, $\psi'(x')$, when x transforms to x' .

Our starting point is the Dirac equation. Before transformation we have

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0, \quad (14.3.2)$$

and after transformation we expect the form of the Dirac equation to remain unchanged, that is

$$(i\gamma'^\mu \partial'_\mu - m)\psi'(x') = 0. \quad (14.3.3)$$

Since m is a scalar it doesn't change under Lorentz transformations. The derivative is just a four-vector, so transforms as $\partial_\mu \mapsto \partial'_\mu = \Lambda^\nu{}_\mu \partial_\nu$. We assume that γ'^μ are still gamma matrices, that is they satisfy the Clifford algebra $\{\gamma'^\mu, \gamma'^\nu\} = 2\eta^{\mu\nu}$. We can then, without loss of generality, take $\gamma'^\mu = \gamma^\mu$, since we know that γ'^μ must be a representation of this Clifford algebra and the physics shouldn't depend on our choice of representation, so we may as well choose the same representation before and after the transformation.

We now assume that ψ transforms according to

$$\psi'(x') = S(\Lambda)\psi(x) \quad (14.3.4)$$

where $S(\Lambda)$ is some 4×4 matrix depending on the Lorentz transformation Λ , but $S(\Lambda)$ is not necessarily a Lorentz transformation. Since the two frames are equivalent descriptions of the same physics it must be possible to go between them in both directions, so $S(\Lambda)$ must be nonsingular. We can then write

$$\psi(x) = S^{-1}(\Lambda)\psi'(x'). \quad (14.3.5)$$

However, note that we can also go from the second frame to the first by the inverse Lorentz transformation, Λ^{-1} , so we must have that

$$S^{-1}(\Lambda) = S(\Lambda^{-1}). \quad (14.3.6)$$

We can now write out the original Dirac equation in terms of the transformed quantities by reversing our transformation, that is $\partial_\mu = \Lambda^\nu{}_\mu \partial'_\nu$ and $\psi(x) = S^{-1}(\Lambda)\psi'(x')$. Then the Dirac equation in the first frame can be written as

$$(i\gamma^\mu \Lambda^\nu{}_\mu \partial'_\nu - m)S^{-1}(\Lambda)\psi'(x') = 0. \quad (14.3.7)$$

Now multiply this on the left by $S(\Lambda)$ and, noting that everything apart from γ^μ commutes with $S(\Lambda)$, we get

$$(iS(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu \partial'_\nu - m)\psi'(x'). \quad (14.3.8)$$

We see that this takes the form of the Dirac equation if

$$S(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu = \gamma^\nu. \quad (14.3.9)$$

Multiplying on the left by $S^{-1}(\Lambda)$ and the right by $S(\Lambda)$ we get

$$\gamma^\mu \Lambda^\nu{}_\mu = S^{-1}(\Lambda)\gamma^\nu S(\Lambda). \quad (14.3.10)$$

We can think of this as a sort of consistency equation between Lorentz transformations of four-vectors on the left and Lorentz transformations of spinors on the right. Note that γ^μ is a four-vector as far as Lorentz indices are concerned, and so transforms under a single copy of Λ . On the other hand γ^μ_{ab} is a rank 2 tensor for the purposes of spinor indices, and so transforms under two copies of $S(\Lambda)$ and $S^{-1}(\Lambda)$. We can write this out in terms of spinor components as

$$\gamma^\mu_{ab} \Lambda^\nu{}_\mu = S^{-1}(\Lambda)^{ac} \gamma^\nu_{cd} S(\Lambda)^{db}. \quad (14.3.11)$$

The question now is, what is $S(\Lambda)$? To answer this we consider the two basic types of Lorentz transformations.

14.3.1 Boosts

Consider a boost

$$\Lambda = \begin{pmatrix} \cosh \omega & \sinh \omega & 0 & 0 \\ \sinh \omega & \cosh \omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (14.3.12)$$

Then we claim that the corresponding spinor transformation matrix, S_B , is given by

$$S_B = \exp \left[\alpha^1 \frac{\omega}{2} \right] = \exp \left[\gamma^0 \gamma^1 \frac{\omega}{2} \right]. \quad (14.3.13)$$

Since $(\alpha^1)^2 = 1$ we can expand the exponential into odd and even parts, and since everything is real those parts will be hyperbolic, so

$$\exp \left[\alpha^1 \frac{\omega}{2} \right] = \cosh \frac{\omega}{2} + \alpha^1 \sinh \frac{\omega}{2} = \begin{pmatrix} \mathbb{1}_2 \cosh \frac{\omega}{2} & \sigma^1 \sinh \frac{\omega}{2} \\ \sigma^1 \cosh \frac{\omega}{2} & \mathbb{1}_2 \sinh \frac{\omega}{2} \end{pmatrix}. \quad (14.3.14)$$

In the last step we choose to work in the Dirac representation. Now using $S(\Lambda^{-1}) = S^{-1}(\Lambda)$ and considering the $\mu = 0$ case we can compute the product $S_B^{-1} \gamma^0 S_B$, using

$$S_B^{-1} = \begin{pmatrix} \mathbb{1}_2 \cosh \frac{\omega}{2} & -\sigma^1 \sinh \frac{\omega}{2} \\ -\sigma^1 \sinh \frac{\omega}{2} & \mathbb{1}_2 \cosh \frac{\omega}{2} \end{pmatrix} \quad (14.3.15)$$

and we find

$$S_B^{-1} \gamma^0 S_B = \begin{pmatrix} \mathbb{1}_2 \cosh \omega & \sigma^1 \sinh \omega \\ -\sigma^1 \sinh \omega & -\mathbb{1}_2 \cosh \omega \end{pmatrix} = \gamma^0 \cosh \omega + \gamma^1 \sinh \omega. \quad (14.3.16)$$

This shows that

$$S_B^{-1} \gamma^0 S_B = \Lambda^0_{\nu} \gamma^{\nu}. \quad (14.3.17)$$

We can make similar arguments for the $\mu = 1, 2, 3$ components by computing $S_B^{-1} \gamma^{\mu} S_B$. We find that

$$\Lambda^1_{\nu} \gamma^{\nu} = \gamma^0 \sinh \omega + \gamma^1 \cosh \omega, \quad \Lambda^2_{\nu} \gamma^{\nu} = \gamma^2, \quad \text{and} \quad \Lambda^3_{\nu} \gamma^{\nu} = \gamma^3. \quad (14.3.18)$$

This is what we would expect so S_B is the correct transformation matrix.

This generalises to a boost along some arbitrary direction given by the unit vector $\hat{\mathbf{n}}$, by

$$S_B = \exp \left[\boldsymbol{\alpha} \cdot \hat{\mathbf{n}} \frac{\omega}{2} \right]. \quad (14.3.19)$$

14.3.2 Rotations

Now consider a rotation,

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \vartheta & \sin \vartheta & 0 \\ 0 & -\sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (14.3.20)$$

We claim that the corresponding spinor transformation matrix, S_R , is given by

$$S_R = \exp \left[i \Sigma^3 \frac{\vartheta}{2} \right] \quad (14.3.21)$$

where

$$\Sigma^k = -\frac{i}{4} \varepsilon^{kij} [\gamma^i, \gamma^j] \implies \Sigma^3 = -\frac{1}{2} \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix}. \quad (14.3.22)$$

Choosing to work in the Dirac representation in the last step. More generally in the Dirac representation

$$\Sigma^k = \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \quad (14.3.23)$$

We can check that $(\Sigma^3)^2 = 1$, and so we can again expand the exponential in even and odd parts. This time the factor of i means we get normal trig:

$$S_R = \cos \frac{\vartheta}{2} + i \Sigma^3 \sin \frac{\vartheta}{2}. \quad (14.3.24)$$

We can then proceed similarly to the boosts case computing $S_R^{-1} \gamma^\mu S_R$ for each fixed value of μ . We find that

$$\Lambda^0_{\nu} \gamma^\nu = \gamma^0, \quad \Lambda^1_{\nu} \gamma^\nu = \gamma^1 \cos \vartheta + \gamma^2 \sin \vartheta, \quad (14.3.25)$$

$$\Lambda^2_{\nu} \gamma^\nu = -\gamma^1 \sin \vartheta + \gamma^2 \cos \vartheta, \quad \Lambda^3_{\nu} \gamma^\nu = \gamma^3. \quad (14.3.26)$$

This generalises to a rotation about an axis along \hat{n} as

$$S_R = \exp \left[i \Sigma \cdot \hat{n} \frac{\vartheta}{2} \right]. \quad (14.3.27)$$

14.3.3 General Lorentz Transformation

First note that S_R is unitary, but S_B is not. This corresponds to the fact that the rotations, $SO(3)$, form a compact subgroup of the Lorentz group, whereas the boosts form a noncompact subgroup². However, for both rotations and boosts we have the relation

²See *Symmetries of Particles and Fields* for details.

$$S^{-1} = \gamma^0 S^\dagger \gamma^0, \quad (14.3.28)$$

and since any Lorentz transformation can be formed from rotations and boosts this holds in general.

We can combine the two results into one for a general transformation:

$$S(\Lambda) = \exp \left[-\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right] \quad (14.3.29)$$

where

$$\sigma^{\mu\nu} := \frac{i}{2} [\gamma^\mu, \gamma^\nu] \quad (14.3.30)$$

and $\omega^{ij} = \varepsilon^{ijk} \vartheta^k$ where $\vartheta = \vartheta \hat{n}$ parametrises a rotation by angle ϑ about the unit vector \hat{n} .

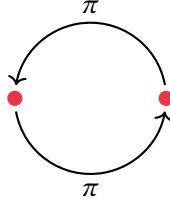


Figure 14.1: Rotating two particles into each other.

Considering just the spatial indices we have

$$-\frac{i}{4}\omega_{ij}\sigma^{ij} = -\frac{i}{4}\varepsilon^{ijk}\mathfrak{g}^k \frac{i}{2}[\gamma^i, \gamma^j] = \frac{i}{2}\boldsymbol{\vartheta} \cdot \boldsymbol{\Sigma}, \quad (14.3.31)$$

so we recover the rotation result. Similarly, considering $\mu = 0$ and $\nu = i$ we recover the boost result.

Notice that for a rotation by 2π we have

$$S_R(2\pi) = -S_R(0) = -\mathbb{1}, \quad (14.3.32)$$

so under a 2π rotation a spinor transforms as $\psi(x) \mapsto -\psi(x)$. This means that it takes a rotation of 4π to get back to the start, that is $S_R(4\pi) = \mathbb{1}$. We can interpret this as the Dirac equation describing fermions. The factors of $1/2$ in the exponentials defining S_R and S_B tell us that in particular these fermions are spin $1/2$. One way to convince us that these are fermions is to consider two identical particles described by the Dirac equation. If we rotate about the midpoint between them by π then the particles swap places, as shown in [Figure 14.1](#). However, two rotations by π are equivalent to one rotation by 2π , and so we get a minus sign. This is exactly what we would expect for fermions following Fermi–Dirac statistics, swapping identical fermions gives a negative sign in the wave function.

Fifteen

Dirac Lagrangian

Our goal in this section will be to derive a Lagrangian which produces the Dirac equation as its equation of motion. To do so we first need to look at the sort of quantities that can appear in the Lagrangian. Presumably it includes a spinor, and since the Lagrangian must be real, and spinors are complex, we need to look for a way to take products with spinors in a way that results in a real value, sort of like how we consider products of φ^\dagger and φ in a complex scalar Lagrangian. However, since spinors transform in a somewhat more complicated manner than scalars we need to carefully consider which combinations of spinors give us Lorentz scalars.

15.1 Dirac Adjoint

We start by defining the **Dirac adjoint** of a spinor, ψ , to be

$$\bar{\psi} := \psi^\dagger \gamma^0. \quad (15.1.1)$$

Recall that under a Lorentz transformation, Λ , a spinor, ψ , transforms as

$$\psi \mapsto \psi' = S(\Lambda)\psi. \quad (15.1.2)$$

The Hermitian conjugate of ψ then transforms as

$$\psi^\dagger \mapsto \psi'^\dagger = \psi^\dagger S(\Lambda)^\dagger. \quad (15.1.3)$$

We want the Lagrangian to be real. For a complex scalar field this was achieved by building it from products of φ^\dagger and φ . Consider then $(\gamma^\mu)^\dagger$. Every representation of the gamma matrices is equivalent to a unitary representation. This follows from Maschke's theorem¹ since the gamma matrices generate a finite group, the elements of which are the identity, products of gamma matrices, and their negatives, that this group is finite follows since any product of gamma matrices can be reduced to a product containing each gamma matrix at most once using the anti-commutation relations. Thus we have $(\gamma^\mu)^\dagger = (\gamma^\mu)^{-1}$. The commutation relations tell us that $(\gamma^0)^2 = 1$, so $(\gamma^0)^{-1} = \gamma^0$, and $(\gamma^i)^2 = -1$, so $(\gamma^i)^{-1} = -\gamma^i$. Now using $-\gamma^i = -\gamma^i(\gamma^0)^2 = \gamma^0\gamma^i\gamma^0$, as well as $(\gamma^0)^3 = \gamma^0$, we have the identity

$$(\gamma^\mu)^\dagger = \gamma^0 \gamma^\mu \gamma^0. \quad (15.1.4)$$

Now consider $S(\Lambda)^\dagger$. If we write $S(\Lambda)$ as

$$S(\Lambda) = \exp \left[-\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right] \quad (15.1.5)$$

¹see *Symmetries of Quantum Mechanics* or *Symmetries of Particles and Fields*

then, using the fact that the parameters, $\omega_{\mu\nu}$, are real we have

$$S(\Lambda)^\dagger = \exp \left[\frac{i}{4} \omega_{\mu\nu} (\sigma^{\mu\nu})^\dagger \right]. \quad (15.1.6)$$

Now, we need to compute $(\sigma^{\mu\nu})^\dagger$. Start from the definition, $\sigma^{\mu\nu} = i[\gamma^\mu, \gamma^\nu]/2$. Taking the Hermitian conjugate turns i into $-i$, and reverses the arguments of the commutator and sends them to their Hermitian conjugates, since taking the transpose reverses the order of a product:

$$(\sigma^{\mu\nu})^\dagger = -\frac{i}{2}[(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger] = \frac{i}{2}[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger] = \frac{i}{2}[\gamma^0 \gamma^\mu \gamma^0, \gamma^0 \gamma^\nu \gamma^0] \quad (15.1.7)$$

where we've used the antisymmetry of the commutator to exchange the order of the arguments. Expanding the commutator whenever we have two neighbouring factors of γ^0 they square to the identity and so we're left only with the outer factors of γ^0 , giving

$$(\sigma^{\mu\nu})^\dagger = \frac{i}{2} \gamma^0 [\gamma^\mu, \gamma^\nu] \gamma^0 = \gamma^0 \sigma^{\mu\nu} \gamma^0. \quad (15.1.8)$$

Hence, we have

$$S(\Lambda)^\dagger = \exp \left[\frac{i}{4} \omega_{\mu\nu} (\sigma^{\mu\nu})^\dagger \right] = \exp \left[\frac{i}{4} \omega_{\mu\nu} \gamma^0 \sigma^{\mu\nu} \gamma^0 \right]. \quad (15.1.9)$$

Now considering $\exp[\gamma^0 X \gamma^0]$ we have

$$\begin{aligned} \exp[\gamma^0 X \gamma^0] &= \sum_{n=0}^{\infty} \frac{1}{n!} (\gamma^0 X \gamma^0)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{\gamma^0 X \gamma^0 \gamma^0 X \gamma^0 \dots \gamma^0 X \gamma^0}_{n \text{ factors of } X} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \gamma^0 X^n \gamma^0 = \gamma^0 \exp[X] \gamma^0, \end{aligned} \quad (15.1.10)$$

since the paired factors of γ^0 square to the identity. Hence,

$$S(\Lambda)^\dagger = \exp \left[\frac{i}{4} \omega_{\mu\nu} \gamma^0 \sigma^{\mu\nu} \gamma^0 \right] = \gamma^0 \exp \left[\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right] \gamma^0 = \gamma^0 S^{-1}(\Lambda) \gamma^0. \quad (15.1.11)$$

Hence the Hermitian conjugate transforms as

$$\psi^\dagger \mapsto \psi'^\dagger = \psi^\dagger S(\Lambda)^\dagger = \psi^\dagger \gamma^0 S^{-1}(\Lambda) \gamma^0 \quad (15.1.12)$$

and so the Dirac adjoint transforms as

$$\bar{\psi} = \psi^\dagger \gamma^0 \mapsto \bar{\psi}' = \psi'^\dagger \gamma^0 = \underbrace{\psi^\dagger \gamma^0}_{=\bar{\psi}} S^{-1}(\Lambda) \underbrace{\gamma^0 \gamma^0}_{=1} = \bar{\psi} S^{-1}(\Lambda). \quad (15.1.13)$$

15.2 Dirac Bilinears

In order to form the Lagrangian we first need to consider what sort of quantities can appear in it. Clearly, the Lagrangian must be a real Lorentz scalar, so we look for ways to form Lorentz scalars from **Dirac bilinears**, which are objects of the form $\bar{\psi} \Gamma \psi$ where ψ is some spinor satisfying the Dirac equation and Γ is a product

of gamma matrices. It turns out that there are only a finite number of linearly independent possibilities for such objects.

We start with the simplest case, where Γ is a product of zero gamma matrices, $\bar{\psi}\psi$. We can easily see that this is a Lorentz scalar, since it transforms as

$$\bar{\psi}\psi \mapsto \bar{\psi}S^{-1}(\Lambda)S(\Lambda)\psi = \bar{\psi}\psi. \quad (15.2.1)$$

The next simplest case is when Γ is a single gamma matrix, so let $j^\mu = \bar{\psi}\gamma^\mu\psi$. Under a Lorentz transformation we have

$$j^\mu \mapsto j'^\mu = \bar{\psi}'\gamma^\mu\psi' = \bar{\psi}S^{-1}(\Lambda)\gamma^\mu S(\Lambda)\psi. \quad (15.2.2)$$

Now recall that, by definition

$$S^{-1}(\Lambda)\gamma^\nu S(\Lambda) = \gamma^\nu \Lambda^\mu{}_\nu, \quad (15.2.3)$$

so

$$j^\mu \mapsto \bar{\psi}\Lambda^\mu{}_\nu\gamma^\nu\psi = \Lambda^\mu{}_\nu j^\nu. \quad (15.2.4)$$

Hence j^μ is a four-vector.

We can form a scalar from this four vector by contracting with another four-vector. An important case is contraction with the derivative, we have a scalar

$$\bar{\psi}\gamma^\mu\partial_\mu\psi = \bar{\psi}\not{\partial}\psi. \quad (15.2.5)$$

In order to form a Lagrangian it's not enough for the quantities to be scalars, they also need to be real. Starting with $\bar{\psi}\psi$ we have

$$(\bar{\psi}\psi)^\dagger = (\psi^\dagger\gamma^0\psi)^\dagger = \psi^\dagger\gamma^0(\psi^\dagger)^\dagger = \psi^\dagger\gamma^0\psi = \bar{\psi}\psi. \quad (15.2.6)$$

So $\bar{\psi}\psi$ is real. On the other hand we have

$$(\bar{\psi}\not{\partial}\psi)^\dagger = (\psi^\dagger\gamma^0\gamma^\mu\partial_\mu\psi)^\dagger = (\partial_\mu\psi^\dagger)(\gamma^\mu)^\dagger(\gamma^0)^\dagger\psi = (\partial_\mu\bar{\psi})\gamma^\mu\psi. \quad (15.2.7)$$

Here we've used $(\gamma^\mu)^\dagger(\gamma^0)^\dagger = \gamma^0\gamma^\mu\gamma^0\gamma^0 = \gamma^0\gamma^\mu$. It isn't immediately clear what this means for the realness of $\bar{\psi}\not{\partial}\psi$. When forming the action we integrate over the Lagrangian, so we can consider this quantity under an integral:

$$\int d^4x (\bar{\psi}\not{\partial}\psi)^\dagger = \int d^4x (\partial_\mu\bar{\psi})\gamma^\mu\psi = \int d^4x [\partial_\mu(\bar{\psi}\gamma^\mu\psi) - \bar{\psi}\gamma^\mu\partial_\mu\psi]. \quad (15.2.8)$$

Now, the first term is a total derivative, and we are free to add a total derivative to the Lagrangian without changing anything, so we can drop this term without changing the physics. The second term is just $-\bar{\psi}\not{\partial}\psi$. So, for the purpose of forming our Lagrangian $(\bar{\psi}\not{\partial}\psi)$ is equivalent to $-\bar{\psi}\not{\partial}\psi$. Hence, we can think of $\bar{\psi}\not{\partial}\psi$ as a purely imaginary quantity, and so $i\bar{\psi}\not{\partial}\psi$ is a real quantity.

15.3 Dirac Lagrangian

We now have enough to form the Dirac Lagrangian. We know that we can use $\bar{\psi}\psi$ and $i\bar{\psi}\not{\partial}\psi$ in our Lagrangian and after some thought we find that

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi. \quad (15.3.1)$$

The action is then

$$S = \int d^4x \bar{\psi}(i\partial - m)\psi. \quad (15.3.2)$$

We can readily check that this is the correct Lagrangian by finding the equations of motion. Since ψ is complex we vary ψ and $\bar{\psi}$ independently. First varying $\bar{\psi}$ we have

$$\frac{\partial \mathcal{L}}{\partial \bar{\psi}} = (i\partial - m)\psi, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi})} = 0 \implies (i\partial - m)\psi = 0, \quad (15.3.3)$$

which is exactly the Dirac equation. Now varying ψ we have

$$\frac{\partial \mathcal{L}}{\partial \psi} = -m\bar{\psi}, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = i\bar{\psi}\gamma^\mu \implies i\partial_\mu \bar{\psi}\gamma^\mu + m\bar{\psi} = 0. \quad (15.3.4)$$

This is simply the Hermitian conjugate of the Dirac equation:

$$0 = [i\gamma^\mu(\partial_\mu \psi) - m\psi]^\dagger \quad (15.3.5)$$

$$= -i(\partial_\mu \psi)^\dagger (\gamma^\mu)^\dagger - m\psi^\dagger \quad (15.3.6)$$

$$= -i\partial_\mu \psi^\dagger \gamma^0 \gamma^\mu \gamma^0 - m\psi^\dagger \quad (15.3.7)$$

$$= -i\partial_\mu \bar{\psi}\gamma^\mu \gamma^0 - m\psi^\dagger. \quad (15.3.8)$$

Now multiply by $-\gamma^0$ on the right and we have

$$0 = -[i\gamma^\mu(\partial_\mu \psi) - m\psi]^\dagger \gamma^0 \quad (15.3.9)$$

$$= i\partial_\mu \bar{\psi}\gamma^\mu (\gamma^0)^2 + m\psi^\dagger \gamma^0 \quad (15.3.10)$$

$$= i\partial_\mu \bar{\psi}\gamma^\mu + m\bar{\psi}, \quad (15.3.11)$$

which is the equation derived by varying ψ .

As a complex object the spinor has a $U(1)$ symmetry given by

$$\psi \mapsto e^{-i\alpha}\psi \implies \bar{\psi} \mapsto e^{i\alpha}\bar{\psi}, \quad (15.3.12)$$

and so $\bar{\psi}\psi$ and $\bar{\psi}\partial\psi$ are invariant under this symmetry. Hence there is a conserved charge. To find it we consider the case of infinitesimal α , where we have

$$\delta\psi = e^{-i\alpha}\psi - \psi \approx (1 - i\alpha)\psi - \psi = -i\alpha\psi. \quad (15.3.13)$$

Hence the conserved current is, scaling out the factor of α ,

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \delta\psi = \bar{\psi}\gamma^\mu \psi. \quad (15.3.14)$$

The associated conserved charge is

$$Q = \int d^3\mathbf{x} j^0 = \int d^3\mathbf{x} \bar{\psi}\gamma^0\psi = \int d^3\mathbf{x} \psi^\dagger \gamma^0 \gamma^0 \psi = \int d^3\mathbf{x} \psi^\dagger \psi. \quad (15.3.15)$$

We can also derive a Hamiltonian for the Dirac equation. First, we define the conjugate variables

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger, \quad \text{and} \quad \bar{\pi} = \frac{\partial \mathcal{L}}{\partial \dot{\bar{\psi}}} = 0. \quad (15.3.16)$$

Then we form the Hamiltonian density in the usual way:

$$\mathcal{H} = \pi\dot{\psi} + \bar{\pi}\dot{\bar{\psi}} - \mathcal{L} \quad (15.3.17)$$

$$= i\psi^\dagger\dot{\psi} - i\bar{\psi}\not{\partial}\psi + m\bar{\psi}\psi \quad (15.3.18)$$

$$= \bar{\psi}(-i\gamma^i\partial_i + m)\psi. \quad (15.3.19)$$



The Hamiltonian derived by integrating this over spacetime is *not* the “Dirac Hamiltonian”, $H = -i\boldsymbol{\alpha}\cdot\nabla + \beta m$. This second quantity was mistakenly identified as the Hamiltonian by Dirac when he first derived his equation and the name has stuck.

It is also possible to derive the energy–momentum tensor for the Dirac Lagrangian, and we find that it is

$$T^{\mu\nu} = \bar{\psi}\gamma^\mu\partial^\nu\psi. \quad (15.3.20)$$

Sixteen

Solutions to the Dirac Equation

16.1 Free Particle Solution

We look for a free particle solution to the Dirac equation,

$$(i\partial - m)\psi = 0. \quad (16.1.1)$$

We'll work in the Dirac representation in this chapter. We start with the ansatz of plane wave solutions, multiplied by some spinor so that ψ is a spinor. Assuming that the particle is on-shell, $p^2 = m^2$, we expect this spinor to depend on the three-momentum and mass of the particles, and be independent of position since we're looking for free solutions. So we look for solutions of the form

$$\psi(x) = u(\mathbf{p})e^{-ip \cdot x} \quad (16.1.2)$$

for some spinor u . Substituting this into the Dirac equation we have

$$(i\partial - m)u(\mathbf{p})e^{-ip \cdot x} = iu(\mathbf{p})\partial e^{-ip \cdot x} - mu(\mathbf{p})e^{-ip \cdot x} = 0. \quad (16.1.3)$$

Consider the derivative term:

$$\partial e^{-ip \cdot x} = \gamma^\mu \partial_\mu e^{-ip_\nu x^\nu} = -i\gamma^\mu p_\mu e^{-ip \cdot x} = -i\mathbf{p} e^{-ip \cdot x}. \quad (16.1.4)$$

Hence, substituting this back into the above equation and dividing through by the phase factor we have

$$(\mathbf{p} - m)u(\mathbf{p}) = 0. \quad (16.1.5)$$

In the Dirac representation u is a four-component spinor. We can further split it into two two-component spinors, φ and χ :

$$u = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}. \quad (16.1.6)$$

We then use the block form of the gamma matrices:

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix}, \quad \text{and} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad (16.1.7)$$

to write

$$\mathbf{p} = \gamma^\mu p_\mu = \gamma^0 p_0 - \gamma^i p_i = \begin{pmatrix} p^0 & 0 \\ 0 & -p^0 \end{pmatrix} - \begin{pmatrix} 0 & \sigma^i p_i \\ \sigma^i p_i & 0 \end{pmatrix} = \begin{pmatrix} p^0 & -\boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & -p^0 \end{pmatrix}. \quad (16.1.8)$$

Then we can write

$$(\mathbf{p} - m) = \begin{pmatrix} p^0 & -\boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & -p^0 \end{pmatrix} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = m \begin{pmatrix} \varphi \\ \chi \end{pmatrix}. \quad (16.1.9)$$

16.1.1 Positive Energy Solutions

Take $E = p^0$, and then we have

$$\begin{pmatrix} E & -\boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & -E \end{pmatrix} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = m \begin{pmatrix} \varphi \\ \chi \end{pmatrix}. \quad (16.1.10)$$

Doing the matrix multiplication on the left we get

$$\begin{pmatrix} E\varphi - \boldsymbol{\sigma} \cdot \mathbf{p}\chi \\ \boldsymbol{\sigma} \cdot \mathbf{p}\varphi - E\chi \end{pmatrix} = m \begin{pmatrix} \varphi \\ \chi \end{pmatrix}. \quad (16.1.11)$$

So we must have

$$E\varphi - \boldsymbol{\sigma} \cdot \mathbf{p}\chi = m\varphi, \quad (16.1.12)$$

$$\boldsymbol{\sigma} \cdot \mathbf{p}\varphi - E\chi = m\chi, \quad (16.1.13)$$

rearranging gives

$$(E - m)\varphi = \boldsymbol{\sigma} \cdot \mathbf{p}\chi, \quad (16.1.14)$$

$$(E + m)\chi = \boldsymbol{\sigma} \cdot \mathbf{p}\varphi. \quad (16.1.15)$$

Solving the second for χ we get

$$\chi = \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E + m} \varphi. \quad (16.1.16)$$

Substituting this into the first we get

$$(E - m)\varphi = \frac{(\boldsymbol{\sigma} \cdot \mathbf{p})^2}{E + m} \varphi. \quad (16.1.17)$$

We can calculate $(\boldsymbol{\sigma} \cdot \mathbf{p})^2$:

$$(\boldsymbol{\sigma} \cdot \mathbf{p})^2 = \sigma^i p^i \sigma^j p^j \quad (16.1.18)$$

$$= p^i p^j \sigma^i \sigma^j \quad (16.1.19)$$

$$= p^i p^j \frac{1}{2} \{\sigma^i, \sigma^j\} \quad (16.1.20)$$

$$= p^i p^j \delta^{ij} \quad (16.1.21)$$

$$= p^i p^j \quad (16.1.22)$$

$$= \mathbf{p}^2. \quad (16.1.23)$$

Here we've used the fact that $p^i p^j$ is symmetric in i and j to replace $\sigma^i \sigma^j$ with its symmetric part, $\{\sigma^i, \sigma^j\}/2$, since any antisymmetric part vanishes in the product with something symmetric.

Hence, we have

$$(E - m)\varphi = \frac{\mathbf{p}^2}{E + m} \varphi \implies \varphi = \frac{\mathbf{p}}{E^2 - m^2} \varphi = \varphi. \quad (16.1.24)$$

since $E^2 - m^2 = \mathbf{p}^2$ for an on-shell particle. This means that φ is not fixed by the Dirac equation, only its relation to χ is fixed. We can therefore choose φ , and as a two-component spinor we make the choice that

$$\varphi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \varphi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (16.1.25)$$

This is a basis for the space of two-component spinors allowing us to express arbitrary φ . We can then consider two positive energy solutions, one for each basis spinor:

$$u(\mathbf{p}, s) = N \begin{pmatrix} \varphi^s \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \varphi^s \end{pmatrix}. \quad (16.1.26)$$

Here N is some normalisation factor which we'll discuss later.

Note that in the rest frame of the particles we get

$$u(\mathbf{0}, s) = N \begin{pmatrix} \varphi^s \\ 0 \end{pmatrix} \quad (16.1.27)$$

which is an eigenvector of the spin operator

$$\frac{1}{2} \Sigma^3 = \frac{1}{2} \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} \quad (16.1.28)$$

with eigenvalues $1/2$ and $-1/2$ for $s = 1$ and $s = 2$ respectively. We therefore interpret φ^1 with spin up particles and φ^2 with spin down particles.

16.1.2 Negative Energy Solutions

If instead we take $p^\mu = (-E, -\mathbf{p})$ then we get slightly different solutions. First, we make the alternative ansatz

$$\psi(x) = v(\mathbf{p}) e^{+ip \cdot x}. \quad (16.1.29)$$

Substituting this into the Dirac equation and cancelling the phase factor we get

$$(\not{p} + m)v(\mathbf{p}) = 0. \quad (16.1.30)$$

Following the same steps of writing this as a block matrix and computing the two components and rearranging we get

$$\varphi = \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E + m} \chi, \quad (16.1.31)$$

so φ and χ have swapped compared to the positive energy case. Again χ is not fixed after substituting this into the equation from the other component, and we have the solution

$$v(\mathbf{p}, s) = N \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \chi^s \\ \chi^s \end{pmatrix} \quad (16.1.32)$$

where

$$\chi^1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad \chi^2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad (16.1.33)$$

is a basis for the space of two-component spinors. We choose this basis, rather than the more obvious choice of $\chi^1 = (1, 0)^\top$ and $\chi^2 = (0, 1)^\top$, since we want to interpret the negative energy particles as antiparticles, which have opposite spin to the corresponding particles, and then the negative sign comes from requiring charge conjugation symmetry.

The important thing is that, despite Dirac's best attempts, there are still negative energy solutions. It turns out that these are unavoidable, since they correspond to antiparticles.

16.2 Spin, Angular Momentum and Helicity

We've interpreted

$$S_3 = \frac{1}{2}\Sigma^3 = \frac{1}{2}\begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} \quad (16.2.1)$$

as measuring the spin along the z-axis. We can further define

$$\mathbf{S} := \frac{1}{2}\boldsymbol{\Sigma} = \frac{1}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \quad (16.2.2)$$

as the spin operator with eigenvalues $\pm 1/2$, and

$$\mathbf{S}^2 := \frac{3}{4}(\boldsymbol{\Sigma}^3)^2 = \frac{3}{4}\mathbb{1}_4 \quad (16.2.3)$$

is the spin operator with eigenvalues $3/4 = 1/2(1/2 + 1)$.

The operator $\boldsymbol{\Sigma}$ only commutes with the Dirac Hamiltonian, $\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m$, in the rest frame of the particle. The angular momentum operator, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, also only commutes with the Dirac Hamiltonian in the rest frame of the particle. However,

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (16.2.4)$$

commutes with the Dirac Hamiltonian in any frame, suggesting we interpret it as the total angular momentum operator.

Another choice of operator which commutes in any frame is the **helicity** operator

$$h(\mathbf{p}) = \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{|\mathbf{p}|}. \quad (16.2.5)$$

This measures the spin along the direction of propagation of the particle.

16.3 Normalisation

There are different conventions for how to normalise spinors. We'll make the choice to define

$$N = \sqrt{E + m} \quad (16.3.1)$$

so

$$u(\mathbf{p}, s) = \sqrt{E + m} \begin{pmatrix} \varphi^s \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E + m} \varphi^s \end{pmatrix}, \quad \text{and} \quad v(\mathbf{p}, s) = \sqrt{E + m} \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E + m} \chi^s \\ \chi^s \end{pmatrix}. \quad (16.3.2)$$

We then have

$$u^\dagger u = (E + m) \left[\varphi^\dagger \varphi + \varphi^\dagger \left(\frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E + m} \right)^2 \varphi \right] \quad (16.3.3)$$

and using $\varphi^\dagger \varphi = 1$ for either spin and $(\boldsymbol{\sigma} \cdot \mathbf{p})^2 = \mathbf{p}^2 = E^2 - m^2$ we have

$$u^\dagger u = (E + m) \left[1 + \frac{E^2 - m^2}{(E + m)^2} \right] \quad (16.3.4)$$

$$= (E + m) \left[1 + \frac{(E + m)(E - m)}{(E + m)^2} \right] \quad (16.3.5)$$

$$= E + m + E - m \quad (16.3.6)$$

$$= 2E. \quad (16.3.7)$$

Similarly $v^\dagger v = 2E$.

Defining the Dirac adjoints, $\bar{u} = u^\dagger \gamma^0$ and $\bar{v} = v^\dagger \gamma^0$ we can compute $\bar{u}u$ in the rest frame, since $\bar{u}u$ is a Lorentz invariant scalar, and we find

$$\bar{u}u = (E + m) \left[(\varphi^\dagger \quad \chi^\dagger) \gamma^0 \begin{pmatrix} \varphi \\ \chi \end{pmatrix} + (\varphi^\dagger \quad \chi^\dagger) \frac{(\boldsymbol{\sigma} \cdot \mathbf{p}) \gamma^0 (\boldsymbol{\sigma} \cdot \mathbf{p})}{(E + m)^2} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \right] \quad (16.3.8)$$

$$= E + m \quad (16.3.9)$$

$$= 2m \quad (16.3.10)$$

since the second term vanishes in the rest frame, with $\mathbf{p} = \mathbf{0}$, and $E = m$ in the rest frame. Similarly we can show that

$$\bar{v}v = -2m. \quad (16.3.11)$$

Other quantities, such as $\bar{u}v$, $\bar{v}u$, and $\bar{u}(\mathbf{p}, s)u(\mathbf{p}, s')$ with $s \neq s'$ vanish.

We've just computed the inner products of our spinor solutions, we can also compute outer products, $u\bar{u}$. These are most interesting if we sum over spins, where we get the results

$$\sum_s u(\mathbf{p}, s)\bar{u}(\mathbf{p}, s) = \not{p} + m, \quad (16.3.12)$$

$$\sum_s v(\mathbf{p}, s)\bar{v}(\mathbf{p}, s) = \not{p} - m. \quad (16.3.13)$$

$$(16.3.14)$$

Note that the sign is the opposite to the sign in $(\not{p} - m)u = (\not{p} + m)v = 0$.

Finally, we can define the projection operators

$$\Lambda_\pm := \frac{1}{2m}(\pm \not{p} + m). \quad (16.3.15)$$

These have the effect of projecting on to the positive and negative energy eigenspaces:

$$\Lambda_+ u(\mathbf{p}, s) = \frac{1}{2m}(\not{p} + m)u \quad (16.3.16)$$

$$= \frac{1}{2m} \sum_{s'} u(\mathbf{p}, s')\bar{u}(\mathbf{p}, s')u(\mathbf{p}, s) \quad (16.3.17)$$

$$= \frac{1}{2m} \delta_{ss'} u(\mathbf{p}, s') 2m \quad (16.3.18)$$

$$= u(\mathbf{p}, s), \quad (16.3.19)$$

and similarly

$$\Lambda_- v(\mathbf{p}, s) = v(\mathbf{p}, s). \quad (16.3.20)$$

Since they are projection operators they are idempotent, $\Lambda_\pm^2 = \Lambda_\pm$, and they are also orthogonal projectors, $\Lambda_+ \Lambda_- = \Lambda_- \Lambda_+ = 0$. In particular we can show that $\Lambda_+ u = u$, $\Lambda_+ v = 0$, $\Lambda_- u = 0$, and $\Lambda_- v = v$, so

$$\Lambda_+(\alpha u + \beta v) = \alpha u, \quad \text{and} \quad \Lambda_-(\alpha u + \beta v) = \beta v \quad (16.3.21)$$

for $\alpha, \beta \in \mathbb{C}$.

Part V

Gaussian Integrals

Seventeen

Gaussian Integrals

This part of the course is full of **Gaussian integrals**, so we'll spend some time to get comfortable with them, starting from the simplest one-dimensional case, and then adding in constants, before moving on to the arbitrary dimension case.

17.1 One Dimension

17.1.1 No Constant Coefficient

It is well known that

$$Z_1 := \int_{-\infty}^{\infty} dx \exp\left[-\frac{1}{2}x^2\right] = \sqrt{2\pi}. \quad (17.1.1)$$

This follows by considering

$$Z_1^2 = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp\left[-\frac{1}{2}(x^2 + y^2)\right] \quad (17.1.2)$$

$$= \int_0^{2\pi} d\vartheta \int_0^{\infty} dr r \exp\left[-\frac{1}{2}r^2\right] \quad (17.1.3)$$

$$= 2\pi \left[-\exp\left[-\frac{1}{2}r^2\right] \right]_0^{\infty} \quad (17.1.4)$$

$$= 2\pi, \quad (17.1.5)$$

and so $Z_1 = \sqrt{Z_1^2} = \sqrt{2\pi}$.

17.1.2 With Coefficient

If we include a coefficient, a , with $\text{Re}(a) > 0$ then we have

$$Z_a := \int_{-\infty}^{\infty} dx \exp\left[-\frac{1}{2}ax^2\right] = \sqrt{\frac{2\pi}{a}}, \quad (17.1.6)$$

For $a \in \mathbb{R}$ this follows by a change of variables, $x \mapsto u/\sqrt{a}$, so $dx = du/\sqrt{a}$, and since $a > 0$ and so $\sqrt{a} > 0$ the integration limits are unchanged so we have

$$Z_a = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} du \exp\left[-\frac{1}{2}u^2\right] = \frac{1}{\sqrt{a}} Z_1 = \sqrt{\frac{2\pi}{a}}. \quad (17.1.7)$$

The same thing holds for $a \in \mathbb{C}$ with $\text{Re}(a) > 0$, but we have to consider some contour integral and its not as straight forward to show.

17.1.3 With Linear Term

Now consider

$$Z_a(b) := \int_{-\infty}^{\infty} dx \exp\left[-\frac{1}{2}ax^2 + bx\right] = \sqrt{\frac{2\pi}{a}} \exp\left[-\frac{b^2}{2a}\right]. \quad (17.1.8)$$

This is shown by completing the square:

$$-\frac{1}{2}ax^2 + bx = -\frac{a}{2}\left(x + \frac{b}{a}\right)^2 + \frac{b^2}{2a}. \quad (17.1.9)$$

Hence,

$$Z_a(b) = \int_{-\infty}^{\infty} dx \exp\left[-\frac{1}{2}a\left(x + \frac{b}{a}\right)\right] \exp\left[-\frac{b^2}{2a}\right]. \quad (17.1.10)$$

Making a change of variables $x \mapsto u + b/a$, so $dx = du$ and the integration region is unchanged we get

$$Z_a(b) = \exp\left[-\frac{b^2}{2a}\right] \int_{-\infty}^{\infty} du \exp\left[-\frac{1}{2}au^2\right] = \exp\left[-\frac{b^2}{2a}\right] Z_a = \sqrt{\frac{2\pi}{a}} \exp\left[-\frac{b^2}{2a}\right].$$

17.2 Arbitrary Dimension

17.2.1 Diagonal Case

Let $x \in \mathbb{R}^n$, then

$$Z_I := \int d^n x \exp\left[-\frac{1}{2}x^\top x\right] = (2\pi)^{n/2} \quad (17.2.1)$$

where an integral without bounds is understood to be over all of \mathbb{R}^n . This follows since $x^\top x = x_i x_i$ and we then can factorise this into a product of n one-dimensional Gaussian integrals, each giving a factor of $\sqrt{2\pi}$:

$$Z_I = (Z_1)^n = (2\pi)^{n/2}. \quad (17.2.2)$$

Now let $A \in \mathcal{M}_n(\mathbb{C})$ be diagonal with $\text{Re}(A_{ij}) > 0$ for all $i, j = 1, \dots, n$. Then we have

$$Z_A := \int d^n x \exp\left[-\frac{1}{2}x^\top A x\right] = (2\pi)^{n/2} \left(\prod_{i=1}^n \frac{1}{\sqrt{A_{ii}}}\right). \quad (17.2.3)$$

This follows since $x^\top A x = x_i A_{ij} x_j$, and for A diagonal we have $x^\top A x = x_i^2 A_{ii}$, so Z_A factorises into a product of one-dimensional Gaussian integrals:

$$Z_A = \prod_{i=1}^n Z_{A_{ii}} = (2\pi)^{n/2} \frac{1}{\sqrt{A_{ii}}}. \quad (17.2.4)$$

Now notice that for a diagonal matrix

$$\det A = \prod_{i=1}^n A_{ii}, \quad (17.2.5)$$

so

$$Z_A = (2\pi)^{n/2} (\det A)^{-1/2}. \quad (17.2.6)$$

17.2.2 Diagonalisable Case

Suppose that instead of being diagonal A is symmetric, and hence diagonalisable. That is, there exists some $D \in \text{SO}(n)$ such that $A' = D^{-1}AD$ is diagonal. Consider the change of variables $x \mapsto Dx$, and $x^\top \mapsto x^\top D^\top = x^\top D^{-1}$. Since $\det D = 1$ we have $d^n x \mapsto \det(D) d^n x = d^n x$. Under this transformation we then have

$$Z_A = \int d^n x \exp \left[-\frac{1}{2} x^\top A x \right] \quad (17.2.7)$$

$$= \int d^n x \exp \left[-\frac{1}{2} x^\top D^{-1} A D x \right] \quad (17.2.8)$$

$$= (2\pi)^{n/2} (\det(D^{-1}AD))^{-1/2} \quad (17.2.9)$$

$$= (2\pi)^{n/2} (\det A)^{-1/2}, \quad (17.2.10)$$

having used

$$\det(D^{-1}AD) = \det(D^{-1}) \det(A) \det(D) = \det(D)^{-1} \det(A) \det(D) = \det A.$$

This shows that if $A \in \mathcal{M}_n(\mathbb{C})$ is a symmetric matrix with $\text{Re}(A_{ij}) \geq 0$ for all $i, j = 1, \dots, n$ and the eigenvalues of A , call them a_i , are nonzero, which guarantees that $\det A \neq 0$, then

$$Z_A := \int d^n x \exp \left[-\frac{1}{2} x^\top A x \right] = (2\pi)^{n/2} (\det A)^{-1/2}. \quad (17.2.11)$$

17.2.3 With Linear Term

Now let $b \in \mathbb{R}^n$ and consider

$$Z_A(b) := \int d^n x \exp \left[-\frac{1}{2} x^\top A x + b^\top x \right]. \quad (17.2.12)$$

Consider the change of variables

$$x = u + \Delta b \quad (17.2.13)$$

where $\Delta = A^{-1}$. We then have $d^n x = d^n u$, and there is no change to the region of integration. Note that since A is symmetric so is Δ . We then have

$$-\frac{1}{2} x^\top A x = -\frac{1}{2} (u^\top + b^\top \Delta^\top) A (u + \Delta b) \quad (17.2.14)$$

$$= -\frac{1}{2} (u^\top A u + u^\top A \Delta b + b^\top \Delta^\top A u + b^\top \Delta^\top A \Delta b). \quad (17.2.15)$$

Now we can use $\Delta^\top = \Delta$ and $u^\top b = u_i b_i = b^\top u$ to get

$$-\frac{1}{2} x^\top A x = -\frac{1}{2} (u^\top A u + u^\top b + b^\top u + b^\top \Delta b) \quad (17.2.16)$$

$$= -\frac{1}{2} u^\top A u - b^\top u - \frac{1}{2} b^\top \Delta b. \quad (17.2.17)$$

Similarly,

$$b^\top x = b^\top u + b^\top \Delta b, \quad (17.2.18)$$

and so

$$-\frac{1}{2}x^\top Ax + b^\top x = -\frac{1}{2}u^\top Au - b^\top u - \frac{1}{2}b^\top \Delta b + b^\top u + b^\top \Delta b \quad (17.2.19)$$

$$= -\frac{1}{2}u^\top Au + \frac{1}{2}b^\top \Delta b. \quad (17.2.20)$$

Hence, we have

$$\exp\left[-\frac{1}{2}x^\top Ax + b^\top x\right] = \exp\left[-\frac{1}{2}u^\top Au\right] \exp\left[\frac{1}{2}b^\top \Delta b\right]. \quad (17.2.21)$$

Note that these are just exponentials of numbers, so everything commutes and the usual exponential laws apply. Finally, we have

$$Z_A(b) = \int d^n x \exp\left[-\frac{1}{2}x^\top Ax + b^\top x\right] \quad (17.2.22)$$

$$= \exp\left[\frac{1}{2}b^\top \Delta b\right] \int d^n u \exp\left[-\frac{1}{2}u^\top Au\right] \quad (17.2.23)$$

$$= Z_A \exp\left[\frac{1}{2}b^\top \Delta b\right] \quad (17.2.24)$$

$$= (2\pi)^{n/2} (\det A)^{-1/2} \exp\left[-\frac{1}{2}b^\top \Delta b\right]. \quad (17.2.25)$$

17.3 Generating Function

Let μ be a measure on \mathbb{R}^n . For our purposes a measure, μ , is something such that $d\mu(x) = d^n x \Omega(x)$ where Ω is some integrable function. We define the **expectation value** of $F: \mathbb{R}^n \rightarrow \mathbb{R}$ with respect to μ to be

$$\langle F \rangle_\mu := \int d\mu(x) F(x) = \int d^n x \Omega(x) F(x). \quad (17.3.1)$$

We assume that the measure is normalised such that

$$\int d\mu(x) = 1. \quad (17.3.2)$$

This allows for a probability interpretation of our results.

We then define the **generating function**

$$Z_\mu(b) := \langle e^{b^\top x} \rangle_\mu = \int d\mu(x) \exp[b^\top x]. \quad (17.3.3)$$

We can expand the exponential in a series to get

$$\exp[b^\top x] = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (b^\top x)^\ell = \sum_{\ell=0}^{\infty} \sum_{i_1, \dots, i_\ell=1}^n b_{i_1} \cdots b_{i_\ell} x_{i_1} \cdots x_{i_\ell}. \quad (17.3.4)$$

Applying this to the generating function, and realising that we can use the linearity of the integral to pull out all the sums and factors of b_i , we have

$$Z_\mu(b) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \sum_{i_1, \dots, i_\ell=1}^n b_{i_1} \cdots b_{i_\ell} \langle x_{i_1} \cdots x_{i_\ell} \rangle_\mu \quad (17.3.5)$$

We call

$$\langle x_{i_1} \cdots x_{i_\ell} \rangle_\mu = \int d\mu(x) x_{i_1} \cdots x_{i_\ell} \quad (17.3.6)$$

the **correlation function** of x with respect to the measure μ . In particular it's called the ℓ -point correlation function, since it has ℓ points, x_{i_1} through x_{i_ℓ} .

From this work we can see that the generating function is simply a polynomial in b_i , after computing the integrals in the correlation functions they are just numbers. We can then take derivatives to get

$$\frac{\partial}{\partial b_k} Z_\mu(b) = \int d\mu(x) \frac{\partial}{\partial b_k} \exp[b_i x_i] = \int d\mu(x) x_k \exp[b_i x_i]. \quad (17.3.7)$$

This allows us to write the correlation function as

$$\langle x_{i_1} \cdots x_{i_\ell} \rangle_\mu = \frac{\partial}{\partial b_{i_1}} \cdots \frac{\partial}{\partial b_{i_\ell}} Z_\mu(b) \Big|_{b=0}, \quad (17.3.8)$$

since $\exp[b^\top x]|_{b=0} = 1$.

We can use this same method to find the expectation value of any function with a Taylor series. If

$$F(x) = \sum_{\ell=0}^{\infty} \sum_{i_1, \dots, i_\ell=1}^n F_{i_1, \dots, i_\ell} x_{i_1} \cdots x_{i_\ell} \quad (17.3.9)$$

for some constants F_{i_1, \dots, i_ℓ} , then

$$\langle F \rangle_\mu = \sum_{\ell=0}^{\infty} \sum_{i_1, \dots, i_\ell=1}^n F_{i_1, \dots, i_\ell} \langle x_{i_1} \cdots x_{i_\ell} \rangle_\mu. \quad (17.3.10)$$

17.4 Gaussian Generating Functions

Consider the **Gaussian measure**

$$d\mu_0(x) := d^n x \Omega_0(x) = d^n x \mathcal{N}_0 \exp\left[-\frac{1}{2} x^\top A x\right] \quad (17.4.1)$$

where $A \in \mathcal{M}_n(\mathbb{C})$ is a symmetric matrix with $\text{Re}(A_{ij}) \geq 0$ for all $i, j = 1, \dots, n$ and with nonzero eigenvalues, and \mathcal{N}_0 is the normalisation factor

$$\mathcal{N}_0 = (2\pi)^{-n/2} (\det A)^{1/2} = Z_A^{-1} \quad (17.4.2)$$

which ensures that

$$\int d\mu_0(x) = 1. \quad (17.4.3)$$

We can then define the **Gaussian generating function**,

$$Z_0(b) = \frac{Z_A(b)}{Z_A} = \exp\left[\frac{1}{2} b^\top \Delta b\right]. \quad (17.4.4)$$

Then the **Gaussian correlation function** is

$$\langle x_{i_1} \cdots x_{i_\ell} \rangle_0 := \langle x_{i_1} \cdots x_{i_\ell} \rangle_{\mu_0} = \frac{\partial}{\partial b_{i_1}} \cdots \frac{\partial}{\partial b_{i_\ell}} \exp \left[\frac{1}{2} b^\top \Delta b \right] \Big|_{b=0}. \quad (17.4.5)$$

If F is an arbitrary function with a Taylor series as in Equation (17.3.9) then

$$\langle F \rangle_0 = \sum_{\ell=0}^{\infty} \sum_{i_1, \dots, i_\ell=1}^n F_{i_1, \dots, i_\ell} \frac{\partial}{\partial b_{i_1}} \cdots \frac{\partial}{\partial b_{i_\ell}} \exp \left[\frac{1}{2} b^\top \Delta b \right] \Big|_{b=0} \quad (17.4.6)$$

$$= F(\partial/\partial b) \exp \left[\frac{1}{2} b^\top \Delta b \right] \Big|_{b=0} \quad (17.4.7)$$

where $F(\partial/\partial b)$ is a formal operator expression given by “evaluating” the Taylor series for F at the partial derivatives:

$$F(\partial/\partial b) = \sum_{\ell=0}^{\infty} \sum_{i_1, \dots, i_\ell=1}^n F_{i_1, \dots, i_\ell} \frac{\partial}{\partial b_{i_1}} \cdots \frac{\partial}{\partial b_{i_\ell}}. \quad (17.4.8)$$

17.4.1 One Point Gaussian Correlator

Consider the one point Gaussian correlation function,

$$\langle x_k \rangle_0 = \frac{\partial}{\partial b_k} \langle \exp[b^\top x] \rangle_0 \Big|_{b=0} = \frac{\partial}{\partial b_k} \exp \left[\frac{1}{2} b^\top \Delta b \right] \Big|_{b=0}. \quad (17.4.9)$$

We have

$$\frac{\partial}{\partial b_k} \exp \left[\frac{1}{2} b_i \Delta_{ij} b_j \right] = \frac{1}{2} [\Delta_{kj} b_j + b_i \Delta_{ik}] \exp[b^\top \Delta b] = \Delta_{kj} b_j \exp[b^\top \Delta b], \quad (17.4.10)$$

where we’ve used the symmetry of Δ and $\partial b_i / \partial b_k = \delta_{ik}$. We then have

$$\langle x_k \rangle_0 = (\Delta_{kj} b_j) \exp \left[\frac{1}{2} b^\top \Delta b \right] \Big|_{b=0} = 0. \quad (17.4.11)$$

We could have realised that this must vanish since our integrand is an odd function of x over an even range. We can interpret the one point correlation as the mean of a normally distributed value.

17.4.2 Two Point Gaussian Correlator

Consider the two point Gaussian correlation function,

$$\langle x_k x_l \rangle_0 = \frac{\partial}{\partial b_k} \frac{\partial}{\partial b_l} \langle \exp[b^\top x] \rangle_0 \Big|_{b=0}. \quad (17.4.12)$$

We’ve already done the first derivative, so we just have to compute

$$\langle x_k x_l \rangle_0 = \frac{\partial}{\partial b_l} \left((\Delta_{kj} b_j) \exp \left[\frac{1}{2} b^\top \Delta b \right] \right) \Big|_{b=0} \quad (17.4.13)$$

$$= [\delta_{kk} \Delta_{kj} + (\Delta_{kj} b_j) (\Delta_{li} b_i)] \exp \left[\frac{1}{2} b^\top \Delta b \right] \Big|_{b=0} \quad (17.4.14)$$

$$= \Delta_{kl}. \quad (17.4.15)$$

We can interpret the two point correlation as the covariance of two normally distributed values.

17.4.3 General Gaussian Correlator

Now consider the ℓ -point Gaussian correlation function

$$\langle x_{i_1} \cdots x_{i_\ell} \rangle_0 = \frac{\partial}{\partial b_{i_1}} \cdots \frac{\partial}{\partial b_{i_\ell}} \langle \exp[b^\top x] \rangle_0 \Big|_{b=0}. \quad (17.4.16)$$

In general the derivatives will either act on the exponential term, giving a factor of Δb , or they can act on an existing Δb term giving a Δ term. The only nonzero terms after setting $b = 0$ are those where each Δb has been acted on by a derivative to remove the b dependence. This requires an even number of derivatives to get a nonzero result. In the case when ℓ is even the process for determining the result is as follows:

- Write all pairings (i_p, i_q) , which can be formed from the indices i_1, \dots, i_ℓ . Treat (i_p, i_q) and (i_q, i_p) as the same and no index can be in a pair with itself.
- Build the set, P , of all sets of pairings, so the elements of P are sets of $\ell/2$ pairs such that every index appears in exactly one pair.
- The correlator is given by

$$\langle x_{i_1} \cdots x_{i_\ell} \rangle_0 = \sum_P \Delta_{i_p i_q} \cdots \Delta_{i_r i_s}. \quad (17.4.17)$$

This is **Wick's theorem**.

To make this clearer let's consider some examples. First, the two point correlator, $\langle x_{i_1} x_{i_2} \rangle_0$. We have two indices, i_1 and i_2 , so there is a single pairing, (i_1, i_2) , and we have

$$\langle x_{i_1} x_{i_2} \rangle_0 = \Delta_{i_1 i_2}. \quad (17.4.18)$$

We can also write this as

$$\langle x_{i_1} x_{i_2} \rangle_0 = \overline{x_{i_1} x_{i_2}} \quad (17.4.19)$$

where

$$\overline{x_{i_1} x_{i_2}} := \Delta_{i_1 i_2}. \quad (17.4.20)$$

Now consider the four point correlator, $\langle x_{i_1} x_{i_2} x_{i_3} x_{i_4} \rangle_0$. The set of possible pairings is

$$P = \{(i_1, i_2), (i_3, i_4)\}, \{(i_1, i_3), (i_2, i_4)\}, \{(i_1, i_4), (i_2, i_3)\}. \quad (17.4.21)$$

The correlator is then

$$\langle x_{i_1} x_{i_2} x_{i_3} x_{i_4} \rangle_0 = \Delta_{i_1 i_2} \Delta_{i_3 i_4} + \Delta_{i_1 i_3} \Delta_{i_2 i_4} + \Delta_{i_1 i_4} \Delta_{i_2 i_3} \quad (17.4.22)$$

$$= \overline{x_{i_1} x_{i_2} x_{i_3} x_{i_4}} + \overline{x_{i_1} x_{i_2} x_{i_3} x_{i_4}} + \overline{x_{i_1} x_{i_2} x_{i_3} x_{i_4}}. \quad (17.4.23)$$

Finally, let's do the six point correlation function. We'll jump straight to the contractions:

$$\langle x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6} \rangle = \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,2)(3,4)(5,6)} + \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,2)(3,5)(4,6)} \quad (17.4.24)$$

$$+ \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,2)(3,6)(4,5)} + \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,3)(2,4)(5,6)} \quad (17.4.25)$$

$$+ \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,3)(2,5)(4,6)} + \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,3)(2,6)(4,5)} \quad (17.4.26)$$

$$+ \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,4)(2,3)(5,6)} + \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,4)(2,5)(3,6)} \quad (17.4.27)$$

$$+ \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,4)(2,6)(3,5)} + \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,5)(2,3)(4,6)} \quad (17.4.28)$$

$$+ \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,5)(2,6)(3,4)} + \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,6)(2,3)(4,5)} \quad (17.4.29)$$

$$+ \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(1,6)(2,5)(3,4)} + \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(2,3)(4,5)(6,1)} \quad (17.4.30)$$

$$+ \overbrace{x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} x_{i_6}}^{(2,3)(4,6)(5,1)} \quad (17.4.31)$$

Code 17.4.32 Here's some Mathematica code to generate all pairs. **WARNING:** Horribly inefficient code! It generates all permutations of the indices, groups them into pairs, then eliminates repeats, so its $\mathcal{O}(\ell!)$ in memory, I can't even run it for $\ell = 12$.

```
1 Union[Sort /@
2   Union[Sort /@ Partition[#[[1]], 2] & /@
3     Table[
4       {#[[i]], #[[i + 1]]}, {i, 1, Length[#], 2}] &@
5     Permutations[Range[ℓ]]
6   (* replace ℓ with number of indices (even) *)
```

As you can see the number of indices grows quickly. Suppose we have $2m$ indices for some integer m . Then there are $2m - 1$ options for what to pair the first variable with, since it can't be paired with itself. Then take the next unpaired variable, there are $2m - 3$ variables to pair it with, since it can't be paired with itself or either of the already paired variables. The next unpaired variable has $2m - 5$ potential pairings, and so on, until we have four variables left, one of these has three options for pairing, and then we're left with two variables which we have to pair. That is, the number of pairings is given by

$$\text{number of pairings} = (2m - 1)(2m - 3) \cdots 3 \cdots 1 = (2m - 1)!! \quad (17.4.33)$$

where $!!$ is the double factorial, defined recursively as $n!! = n(n - 2)!!$ with $1!! = 0!! = 1$. Here are the number of pairings for the first few nonzero correlators:

2	4	6	8	10	12	14	16	18	20
1	3	15	105	945	10395	135135	2027025	34459425	654729075

So it's fair to say that evaluating general correlators with more than a couple of points quickly becomes intractable.

Eighteen

Perturbed Correlators

18.1 General Theory

As with the Gaussian case we can write the expected value of any analytic function, F , with respect to some measure, μ , as a sum over correlators with respect to μ :

$$\langle F \rangle_\mu = \sum_{\ell=0}^{\infty} \sum_{i_1, \dots, i_\ell=1}^n F_{i_1 \dots i_\ell} \langle x_{i_1} \cdots x_{i_\ell} \rangle_\mu \quad (18.1.1)$$

$$= \sum_{\ell=0}^{\infty} \sum_{i_1, \dots, i_\ell=1}^n F_{i_1 \dots i_\ell} \frac{\partial}{\partial b_{i_1}} \cdots \frac{\partial}{\partial b_{i_\ell}} Z_\mu(b) \Big|_{b=0} \quad (18.1.2)$$

$$= F(\partial/\partial b) Z_\mu(b) \Big|_{b=0}. \quad (18.1.3)$$

We now consider a slightly more general measure, $d\mu(x) = d^n x \Omega(x)$ where

$$\Omega(x) = \frac{1}{Z_\lambda} e^{-S_\lambda(x)} \quad (18.1.4)$$

where

$$S_\lambda(x) = S_0(x) + \lambda V(x) \quad (18.1.5)$$

with $S_0(x) = x^T A x / 2$, giving the usual Gaussian measure, and V being some “potential” acting as a perturbation. We will assume λ is small and consider expansions in λ . The factor Z_λ is given by

$$Z_\lambda = \int d^n x \exp[-S_\lambda(x)] \quad (18.1.6)$$

$$= \int d^n x \exp[-S_0(x)] \exp[-\lambda V(x)] \quad (18.1.7)$$

$$= Z_0 \langle \exp[-\lambda V(x)] \rangle_0. \quad (18.1.8)$$

We will be interested in computing the value of

$$\frac{Z_\lambda}{Z_0} = \langle \exp[-\lambda V(x)] \rangle_0 = \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \langle V(x)^k \rangle_0. \quad (18.1.9)$$

We can work out the expectation value once we know V by expanding V .

18.2 x^4 Potential

The first instant of a long series of asking “how do we count?”

Luigi Del Debbio

Consider the example

$$V(x) = \frac{1}{4!} \sum_{i=1}^n x_i^4. \quad (18.2.1)$$

We can quickly write down a series for Z_λ/Z_0 in terms of correlators:

$$\frac{Z_\lambda}{Z_0} = 1 - \lambda \frac{1}{4!} \sum_i \langle x_i^4 \rangle_0 + \frac{\lambda^2}{2} \frac{1}{(4!)^2} \sum_i \sum_j \langle x_i^4 x_j^4 \rangle_0 + \mathcal{O}(\lambda^3). \quad (18.2.2)$$

Now we just need to compute the correlators.

Starting with $\langle x_i^4 \rangle_0$. This is just a special case of the four point correlator in Equation (17.4.22) where we set all indices equal. This means that we have

$$\langle x_i x_i x_i x_i \rangle_0 = \Delta_{ii} \Delta_{ii} + \Delta_{ii} \Delta_{ii} + \Delta_{ii} \Delta_{ii} \quad (18.2.3)$$

$$= \overbrace{x_i x_i x_i x_i}^{\square} + \overbrace{x_i x_i x_i x_i}^{\square} + \overbrace{x_i x_i x_i x_i}^{\square} \quad (18.2.4)$$

$$= 3\Delta_{ii}^2. \quad (18.2.5)$$

Notice that we could have worked this out without needing to do the full four point correlator for arbitrary indices, we just need to find a way to count the number of pairings. In this case there are 3 choices for what to pair the first x_i with, then the two remaining x_i s must be paired, so there are 3 terms, and all three are the same.

This is good because as we saw for the six point correlator things quickly get out of hand, and the next term involves an 8 point correlator, which would be even worse at $(8-1)!! = 105$ terms. So, what sort of pairings are there going to be in this correlator? First there are going to be pairings akin to two copies of the four point correlator, with all x_i s paired with x_i s, and all x_j s paired with x_j s. There are nine ways to do this, getting a factor of three from pairing up x_i s and a factor of three from pairing up x_j s. This means we have the term $9\Delta_{ii}^2 \Delta_{jj}^2$.

The next logical term to consider is one which pairs one x_i with one x_j , however, if we do so we're left with an odd number of x_i s, so another of these must pair with an x_j . So, our next term pairs two x_i s with two x_j s, and then the remaining x_i s are paired with each other and likewise for the remaining x_j s. How many ways are there to do this? There are actually (at least) two ways to count this. First, we have $\binom{4}{2}$ ways to choose the two x_i to pair with x_j s. Then we have four choices of which x_j to pair with the first and 3 choices of which x_j to pair with the second, for a total of $\binom{4}{2} \cdot 4 \cdot 3 = 72$ terms. The second way to count starts with choosing two x_i s, again giving a factor of $\binom{4}{2}$, and then choosing two x_j s, for another factor of $\binom{4}{2}$. We can then pair up between these pairs in two ways, pairing the first of each pair, or the first of the x_i s and the second of the x_j s, for a factor of 2, giving $\binom{4}{2}^2 \cdot 2 = 72$. Hence, we have the term $72\Delta_{ij}^2 \Delta_{ii} \Delta_{jj}$.

There's no way to pair up exactly three x_i s with x_j s, so the only other term is pairing all four x_i s with an x_j . This can be done by holding the x_i in some fixed

order, then considering all permutations of the four x_j s, and pairing them up in order for each permutation. So, we have a factor of $4! = 24$. The final term is then $24\Delta_{ij}^4$.

At this point there are some checks we can do. First, we should, including repeated terms, have 8 indices, and 4 Δ s in each term. Second, adding up the combinatorial factors should give 105, the total number of ways to pair 8 things, and indeed this does work out. So, we conclude that

$$\langle x_i^4 x_j^4 \rangle_0 = 9\Delta_{ii}^2 \Delta_{jj}^2 + 72\Delta_{ij}^2 \Delta_{ii} \Delta_{jj} + 24\Delta_{ij}^4. \quad (18.2.6)$$

Hence, to second order, we have

$$\frac{Z_\lambda}{Z_0} = 1 - \frac{\lambda}{4!} \sum_i 3\Delta_{ii}^2 + \frac{\lambda^2}{2(4!)^2} \sum_{i,j} [9\Delta_{ii}^2 \Delta_{jj}^2 + 72\Delta_{ij}^2 \Delta_{ii} \Delta_{jj} + 24\Delta_{ij}^4] \quad (18.2.7)$$

$$= 1 - \frac{\lambda}{8} \sum_i \Delta_{ii}^2 + \lambda^2 \sum_{i,j} \left[\frac{1}{128} \Delta_{ii}^2 \Delta_{jj}^2 + \frac{1}{16} \Delta_{ij}^2 \Delta_{ii} \Delta_{jj} + \frac{1}{48} \Delta_{ij}^4 \right] \quad (18.2.8)$$

18.2.1 Diagrammatic Representation

There is a diagrammatic representation of terms like this. For each propagator, Δ_{ij} , we draw a line connecting the nodes i and j :

$$\Delta_{ij} = i \text{ --- } j \quad (18.2.9)$$

Note that the position of the nodes doesn't matter, just how they're connected. For terms like Δ_{ii} with a repeated index we have a sum over i when these appear. As a diagram a repeated index means a closed loop, and so any closed loop implies a sum:

$$\sum_i \Delta_{ii} = i \text{ --- } i \quad (18.2.10)$$

Another example might be the term $\sum_j \Delta_{ij} \Delta_{jj}$, which corresponds to the diagram

$$\sum_j \Delta_{ij} \Delta_{jj} = i \text{ --- } j \text{ --- } j \quad (18.2.11)$$

Fun Fact™: Diagrams like this one are called tadpole diagrams!

Using this we can write the first order (in λ) Δ_{ii}^2 term as

$$\sum_i \Delta_{ii}^2 = i \text{ --- } i \text{ --- } i \quad (18.2.12)$$

Similarly we can write the following

$$\langle x_i^4 x_j^4 \rangle_0 = 9\Delta_{ii}^2 \Delta_{jj}^2 + 72\Delta_{ij}^2 \Delta_{ii} \Delta_{jj} + 24\Delta_{ij}^4. \quad (18.2.13)$$

as

$$\sum_{i,j} \langle x_i^4 x_j^4 \rangle = 9 \left(i \text{ --- } i \text{ --- } i \text{ --- } i \right) + 72 \left(i \text{ --- } j \text{ --- } j \text{ --- } i \text{ --- } i \text{ --- } i \right) + 24 \left(i \text{ --- } j \text{ --- } j \text{ --- } i \text{ --- } j \text{ --- } j \right) \quad (18.2.14)$$

18.2.2 Computing $\log(Z_\lambda/Z_0)$

Suppose we want to compute $\log(Z_\lambda/Z_0)$. To do this we notice that $Z_\lambda/Z_0 = 1 + \mathcal{O}(\lambda)$, and so we can use the Taylor expansion

$$\log(1 + \varepsilon) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \varepsilon^n \approx \varepsilon - \frac{1}{2} \varepsilon^2 + \frac{1}{3} \varepsilon^3 - \dots \quad (18.2.15)$$

Since we already have ε here as a power series in λ we should work this out in powers of λ . The first order term has only a single contribution, from the first order term in λ . That is, to first order

$$\log\left(\frac{Z_\lambda}{Z_0}\right) = -\frac{\lambda}{8} \sum_i \Delta_{ii}^2 + \mathcal{O}(\lambda^2). \quad (18.2.16)$$

For the second order term there are two contributions, we have the first order term squared, which also picks up a factor of $-1/2$ from the expansion of the log:

$$-\frac{1}{2} \frac{\lambda^2}{8^2} \sum_i \Delta_{ii}^2 \sum_j \Delta_{jj}^2 = -\frac{1}{2} \frac{\lambda^2}{128} \sum_{i,j} \Delta_{ii}^2 \Delta_{jj}^2. \quad (18.2.17)$$

We also have the first order in ε term from the second order in λ term:

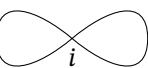


$$\lambda^2 \sum_{i,j} \left[\frac{1}{128} \Delta_{ii}^2 \Delta_{jj}^2 + \frac{1}{16} \Delta_{ij}^2 \Delta_{ii} \Delta_{jj} + \frac{1}{48} \Delta_{ij}^4 \right]. \quad (18.2.18)$$

Notice that the first term here cancels with the other λ^2 term, to give the second order result

$$\log\left(\frac{Z_\lambda}{Z_0}\right) = -\frac{\lambda}{8} \sum_i \Delta_{ii}^2 + \lambda^2 \left[\frac{1}{128} \Delta_{ii}^2 \Delta_{jj}^2 + \frac{1}{16} \Delta_{ij}^2 \Delta_{ii} \Delta_{jj} + \frac{1}{48} \Delta_{ij}^4 \right] + \mathcal{O}(\lambda^3). \quad (18.2.19)$$

In terms of diagrams, this corresponds to

$$\begin{aligned} \log\left(\frac{Z_\lambda}{Z_0}\right) = & -\frac{\lambda}{8} \text{diagram}_1 \\ & + \frac{\lambda^2}{16} \text{diagram}_2 + \frac{\lambda^2}{48} \text{diagram}_3 + \mathcal{O}(\lambda^3). \end{aligned} \quad (18.2.20)$$

Notice that there are no disconnected diagrams, they cancelled out. In fact disconnected diagrams will cancel out here to all orders, and this isn't just the case for this potential. Taking the log of the generating function yields the generating function of connected contributions. We'll prove this later in [Section 22.3](#).

18.3 Correlators

That's it. That's QFT done.

Luigi Del Debbio

Much of the rest of the course will be focused on computing various correlators with respect to some perturbed Gaussian, and of course linking these mathematical expressions to physics. We'll often be interested in calculating

$$\langle x_{i_1} \cdots x_{i_\ell} \rangle_\lambda = \frac{1}{Z_\lambda} \int d^n x \exp[-S_\lambda(x)] x_{i_1} \cdots x_{i_\ell} \quad (18.3.1)$$

$$= \frac{Z_0}{Z_\lambda} \frac{1}{Z_0} \int d^n x \exp[-S_\lambda(x)] x_{i_1} \cdots x_{i_\ell} \quad (18.3.2)$$

$$= \frac{Z_0}{Z_\lambda} \frac{1}{Z_0} \int d^n x \exp[-S_0(x)] \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} V(x)^k x_{i_1} \cdots x_{i_\ell} \quad (18.3.3)$$

$$= \frac{Z_0}{Z_\lambda} \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \langle V(x)^k x_{i_1} \cdots x_{i_\ell} \rangle_0. \quad (18.3.4)$$

The good news is that we've just seen how to compute Z_λ/Z_0 , so all we need to do is invert this result to get Z_0/Z_λ . This can be done in a similar way to the logarithm by again noticing that $Z_\lambda/Z_0 = 1 - \mathcal{O}(\lambda)$ and then using the result

$$\frac{1}{1 - \varepsilon} = \sum_{n=0}^{\infty} \varepsilon^n \approx 1 + \varepsilon + \varepsilon^2 + \cdots. \quad (18.3.5)$$

The results we get will be of the form

$$\langle x_{i_1} \cdots x_{i_\ell} \rangle_0 = \mathcal{O}(\lambda^0) + \mathcal{O}(\lambda) + \mathcal{O}(\lambda^2) + \cdots. \quad (18.3.6)$$

The order of λ^0 (i.e., constant) term will be the result we get with an unperturbed Gaussian. We then have first order corrections, and second order corrections, and so on.

Part VI

Path Integrals in Quantum Mechanics

Nineteen

Path Integrals in QM

19.1 Preliminaries

In this chapter we will develop the theory of path integrals in nonrelativistic quantum mechanics. This should then rapidly generalise to relativistic quantum field theory, but for now we stick to the nonrelativistic case to avoid unnecessary complications.

Working in quantum mechanics we have a position operator, \hat{Q} , with eigenstates $|q\rangle$, with eigenvalues q :

$$\hat{Q}|q\rangle = q|q\rangle. \quad (19.1.1)$$

We consider a system evolving under the Hamiltonian \hat{H} . We will be interested in computing the transition amplitude between two position eigenstates at two different times. That is, we want to compute

$$\langle q', t' | q, t \rangle \quad (19.1.2)$$

where $t' > t$. This amplitude can be computed in terms of the states $|q\rangle$ and $|q'\rangle$ and the time evolution operator:

$$\langle q', t' | q, t \rangle = \langle q' | e^{-i\hat{H}(t'-t)} | q \rangle. \quad (19.1.3)$$

This is where path integrals come in.

19.2 Constructing the Path Integral

We assume this limit exists, and I have no interest in proving it does.

Luigi Del Debbio

Let $T = t' - t$, so we're computing

$$\langle q', t' | q, t \rangle = \langle q' | e^{-i\hat{H}T} | q \rangle. \quad (19.2.1)$$

We can split up the time interval between t and t' into $n-1$ steps of length $\varepsilon = T/n$. Define $t_k = t + k\varepsilon$, so $t_0 = t$ and $t_n = t + n\varepsilon = t + T = t'$. We can then take our single factor of $\exp[-i\hat{H}T]$ and split it into n factors of $\exp[-i\hat{H}\varepsilon]$:

$$\langle q', t' | q, t \rangle = \langle q' | \underbrace{e^{-i\hat{H}\varepsilon} \dots e^{-i\hat{H}\varepsilon}}_{n \text{ factors}} | q \rangle. \quad (19.2.2)$$

Now take the completeness relation,

$$\int dq |q\rangle\langle q| = \mathbb{1}. \quad (19.2.3)$$

We can insert this $n - 1$ times between each of the factors of $\exp[-i\hat{H}t]$ giving

$$\int \left(\prod_{k=1}^{n-1} dq_k \right) \langle q | e^{-i\hat{H}\varepsilon} | q_{n-1} \rangle \langle q_{n-1} | e^{-i\hat{H}\varepsilon} | q_{n-2} \rangle \langle q_{n-2} | \cdots | q_1 \rangle \langle q_1 | e^{-i\hat{H}\varepsilon} | q \rangle. \quad (19.2.4)$$

Now suppose that ε is small, or equivalently n is large. We can expand the exponential:

$$e^{-i\hat{H}\varepsilon} = 1 - i\varepsilon\hat{H} + \mathcal{O}(\varepsilon^2). \quad (19.2.5)$$

Suppose we have the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2} + V(\hat{Q}) \quad (19.2.6)$$

where we choose units such that the mass is unity. First consider the terms with the potential, here we simply replace the argument of V with the appropriate position, using $V(\hat{Q})|q\rangle = \sum_{\ell=1}^{\infty} V^{(\ell)}(0)\hat{Q}^{\ell}|q\rangle = \sum_{\ell=1}^{\infty} V^{(\ell)}(0)q^{\ell} = V(q)|q\rangle$, and so

$$\langle q_k | V(\hat{Q}) | q_{k-1} \rangle = V(q_{k-1}) \langle q_k | q_{k-1} \rangle \quad (19.2.7)$$

$$= V(q_{k-1}) \delta(q_k - q_{k-1}) \quad (19.2.8)$$

$$= V\left(\frac{q_k + q_{k-1}}{2}\right) \delta(q_k - q_{k-1}) \quad (19.2.9)$$

where in the last step we note that for nonzero solutions we have $q_k = q_{k-1}$, and in this case $(q_k + q_{k-1})/2 = q_k$. This choice is required to make things converge later. We can then insert the integral representation of the Dirac delta,

$$\delta(q - q') = \int \frac{dp}{2\pi} e^{ip(q-q')}, \quad (19.2.10)$$

giving

$$\langle q_k | V(\hat{Q}) | q_{k-1} \rangle = \int \frac{dp_k}{2\pi} V\left(\frac{q_k + q_{k-1}}{2}\right) e^{ip_k(q_k - q_{k-1})}. \quad (19.2.11)$$

For the momentum term we can insert a complete set of states,

$$\int dp |p\rangle\langle p| = \mathbb{1}, \quad (19.2.12)$$

giving

$$\frac{1}{2} \langle q_k | \hat{p}^2 | q_{k-1} \rangle = \int dp_k \frac{1}{2} \langle q_k | \hat{p}^2 | p_k \rangle \langle p_k | q_{k-1} \rangle \quad (19.2.13)$$

$$= \int dp_k \frac{p_k^2}{2} \langle q_k | p_k \rangle \langle p_k | q_{k-1} \rangle \quad (19.2.14)$$

$$= \int \frac{dp_k}{2\pi} \frac{p_k^2}{2} e^{ip_k(q_k - q_{k-1})} \quad (19.2.15)$$

where in the last step we've recognised the amplitudes as the wave function of the state momentum eigenstate $|p_k\rangle$ in the position basis,

$$\langle q|p\rangle = \frac{1}{\sqrt{2\pi}} e^{ipq}. \quad (19.2.16)$$

We can take this and go back to writing $1 - i\varepsilon\hat{H}$ as $\exp[-i\hat{H}\varepsilon]$, and the results will agree to first order in ε , that is

$$\begin{aligned} \langle q_k|e^{-i\hat{H}\varepsilon}|q_{k-1}\rangle &= \\ &\int \frac{dp_k}{2\pi} \exp\left\{-i\varepsilon\left[\frac{p_k^2}{2} + V\left(\frac{q_k + q_{k-1}}{2}\right)\right]\right\} \exp[ip_k(q_k - q_{k-1})] + \mathcal{O}(\varepsilon^2). \end{aligned} \quad (19.2.17)$$

We can recognise the term in the first exponential as the Hamiltonian evaluated at $\tilde{q}_k = (q_k + q_{k-1})/2$ and p_k , allowing us to rewrite this as

$$\langle q_k|e^{-i\hat{H}\varepsilon}|q_{k-1}\rangle = \int \frac{dp_k}{2\pi} \exp[-i\varepsilon H(\tilde{q}_k, p_k)] \exp[ip_k(q_k - q_{k-1})] + \mathcal{O}(\varepsilon^2). \quad (19.2.18)$$

Now combine the two exponentials and factor out $i\varepsilon$:

$$\langle q_k|e^{-i\hat{H}\varepsilon}|q_{k-1}\rangle = \int \frac{dp_k}{2\pi} \exp\left\{i\varepsilon\left[p_k \frac{q_k - q_{k-1}}{\varepsilon} - H(\tilde{q}_k, p_k)\right]\right\} + \mathcal{O}(\varepsilon^2). \quad (19.2.19)$$

We can now put all of these factors together to get the desired result, although we have to take the limit of $n \rightarrow \infty$, so $\varepsilon \rightarrow 0$, in order to make it exact. We end up with a product of the above terms, with the product of exponentials becoming a sum in the exponential. We should take the limit in such a way that $T = \varepsilon n$ is constant. The result is

$$\begin{aligned} \langle q', t'|q, t\rangle &= \\ \lim_{n \rightarrow \infty} \int \left(\prod_{k=1}^{n-1} dq_k\right) \left(\prod_{j=1}^n \frac{dp_j}{2\pi}\right) \exp\left\{i\varepsilon \sum_{m=1}^n \left[p_m \frac{q_m - q_{m-1}}{\varepsilon} - iH(\tilde{q}_m, p_m)\right]\right\}. \end{aligned} \quad (19.2.20)$$

We assume that this limit exists. In the limit the sum becomes an integral and $(q_m - q_{m-1})/\varepsilon$ becomes the velocity, the distinction between q_m and \tilde{q}_m is not important in the limit. We introduce the following shorthand for this expression:

$$\langle q', t'|q, t\rangle = \int \mathcal{D}q \mathcal{D}p \exp\left\{i \int_t^{t'} d\tau [p(\tau)\dot{q}(\tau) - H(q(\tau), p(\tau))]\right\}. \quad (19.2.21)$$

This is the **path integral** form of the amplitude. Note that we absorb the $1/(2\pi)$ factor for each momentum integral into the measure.

19.2.1 Quadratic Momentum

For Hamiltonians, like the one we had above, which are quadratic in the momentum we can partially evaluate this integral. Consider the integral

$$\int \frac{dp_k}{2\pi} \exp\left[ip_k(q_k - q_{k-1}) - i\varepsilon \frac{p_k^2}{2}\right] = (2\pi i\varepsilon)^{-1/2} \exp\left[i\varepsilon \left(\frac{q_k - q_{k-1}}{\varepsilon}\right)^2\right],$$

where we've used Equation (17.1.8) with $a = i\varepsilon$ and $b = q_k - q_{k-1}$. Using this we can rewrite the amplitude as

$$\langle q', t' | q, t \rangle = \lim_{n \rightarrow \infty} \int \left(\prod_{k=1}^{n-1} dq_k \right) (2\pi i\varepsilon)^{-n/2} \exp \left\{ i\varepsilon \sum_{m=1}^n \left[\frac{1}{2} \left(\frac{q_m - q_{m-1}}{\varepsilon} \right)^2 - V(\tilde{q}_m) \right] \right\} \quad (19.2.22)$$

In the limit $(q_m - q_{m-1})/\varepsilon$ becomes \dot{q} , and the sum becomes an integral. We then have

$$\frac{1}{2} \dot{q}^2 - V(q) = L(q, \dot{q}), \quad (19.2.23)$$

so we can identify the exponent as the integral of the Lagrangian, which is just the action, $S[q]$. We can also absorb the $(2\pi i\varepsilon)^{-n/2}$ into the measure and, again assuming the limit exists, we get the shorthand

$$\langle q', t' | q, t \rangle = \int_{q, q'} \mathcal{D}q \exp \left[i \int_t^{t'} d\tau L(q, \dot{q}) \right] = \int_{q, q'} \mathcal{D}q e^{-iS[q]}. \quad (19.2.24)$$

Here we write q, q' as the limit of the integrals to remind us that q and q' are the fixed endpoints.

The interpretation of this integral is that we split time into very narrow slices. We then integrate over all positions at each time. An alternative way of viewing this, rather than considering one time slice at a time, is that we consider all possible paths between q and q' , integrating over them with the weighting $\exp\{iS[q]\}$. This gives the quantum amplitude.

Consider this equation in units where $\hbar \neq 1$. Then since the action has units of \hbar and the argument to the exponential must be dimensionless we must have that

$$\langle q' t' | q, t \rangle = \int_{q, q'} \mathcal{D}q e^{-iS[q]/\hbar}. \quad (19.2.25)$$

Now consider the classical limit, where \hbar is much smaller than any other quantity with the same units appearing in our expressions. We can treat this as taking the limit $\hbar \rightarrow 0$. In this limit $1/\hbar$ is very large, and so small changes in the rest of the exponent cause large oscillations in the value. Most of these oscillations cancel out, and only when $S[q]$ is stationary do we get a meaningful contribution. This is the root of the principal of least (or stationary) action. This can be made rigorous with the method of stationary phase, see the course *Methods of Mathematical Physics* for more details. Now set $\hbar = 1$ again.

19.3 Correlators

19.3.1 One Point Correlator

Suppose we wish to calculate the matrix element

$$\langle q', t' | \hat{Q}(\bar{t}) | q, t \rangle. \quad (19.3.1)$$

We assume that $t < \bar{t} = t + \bar{k}\varepsilon = t_{\bar{k}} < t'$, this is a reasonable assumption since we'll be taking time to be a continuous parameter in the limit. We can write the operator $\hat{Q}(\bar{t})$ in the Schrödinger picture as

$$\hat{Q}(t) = e^{i\hat{H}\bar{t}} \hat{Q} e^{-i\hat{H}\bar{t}}. \quad (19.3.2)$$

We can then write the correlator in terms of time independent states:

$$\langle q', t' | \hat{Q}(\bar{t}) | q, t \rangle = \langle q' | e^{-i\hat{H}(t'-\bar{t})} \hat{Q} e^{i\hat{H}(\bar{t}-t)} | q \rangle. \quad (19.3.3)$$

We follow the same procedure as before, splitting time into narrow slices. We then split the exponential into many identical factors of $\exp[-i\hat{H}\varepsilon]$:

$$\langle q', t' | \hat{Q}(\bar{t}) | q, t \rangle = \langle q' | \underbrace{e^{-i\hat{H}\varepsilon} \dots e^{-i\hat{H}\varepsilon}}_{\bar{k} \text{ factors}} \hat{Q} \underbrace{e^{-i\hat{H}\varepsilon} \dots e^{-i\hat{H}\varepsilon}}_{n-\bar{k} \text{ factors}} | q \rangle. \quad (19.3.4)$$

We can then proceed to insert the identity a bunch in terms of the completeness relation. The result will be mostly the same, but we'll have one factor different. Splitting into kinetic and potential terms again we'll have

$$\langle q_{\bar{k}} | \hat{Q}(\bar{t}) V(\hat{Q}) | q_{\bar{k}-1} \rangle = q_{\bar{k}-1} V(\tilde{q}_{\bar{k}}) \delta(q_{\bar{k}} - q_{\bar{k}-1}), \quad (19.3.5)$$

which differs from the other terms by a factor of $q_{\bar{k}}$. The result is that we get an extra factor of $q_{\bar{k}}$ in our integral:

$$\begin{aligned} \langle q', t' | \hat{Q}(\bar{t}) | q, t \rangle = \\ \lim_{n \rightarrow \infty} \left(\prod_{k=1}^{n-1} dq_k \right) \left(\prod_{j=1}^n \frac{dp_j}{2\pi} \right) q_{\bar{k}-1} \exp \left\{ i\varepsilon \sum_{m=1}^n \left[p_m \frac{q_m - q_{m-1}}{\varepsilon} - H(\tilde{q}_m, p_m) \right] \right\}. \end{aligned} \quad (19.3.6)$$

Which, assuming the limit exists, we write as

$$\langle q', t' | \hat{Q}(\bar{t}) | q, t \rangle = \int \mathcal{D}q \mathcal{D}p q(\bar{t}) \exp \left\{ i \int_t^{t'} d\tau [p(\tau) \dot{q}(\tau) - H(q, p)] \right\}. \quad (19.3.7)$$

If the Hamiltonian has quadratic dependence on momentum we can again partially complete the integral to get

$$\begin{aligned} \langle q', t' | \hat{Q}(\bar{t}) | q, t \rangle = \\ \lim_{n \rightarrow \infty} \int \left(\prod_{k=1}^{n-1} dq_k \right) (2\pi i\varepsilon)^{-n/2} q_{\bar{k}} \exp \left\{ i\delta \sum_{m=1}^n \left[\frac{1}{2} \left(\frac{q_m - q_{m-1}}{\varepsilon} \right)^2 - V(\tilde{q}_m) \right] \right\}. \end{aligned} \quad (19.3.8)$$

If this limit exists then we write this as

$$\langle q', t' | \hat{Q}(\bar{t}) | q, t \rangle = \int_{q, q'} \mathcal{D}q q(\bar{t}) \exp \left[i \int_t^{t'} d\tau L(q, \dot{q}) \right]. \quad (19.3.9)$$

19.3.2 Two Point Correlator

Suppose we wish to calculate

$$\langle q', t' | \hat{Q}(\bar{t}_1) \hat{Q}(\bar{t}_2) | q, t \rangle. \quad (19.3.10)$$

We assume that $t < \bar{t}_2 < \bar{t}_1 < t'$. Exactly the same logic as before applies, except we now have two of these altered factors, so get two factors of q in our final path integral. The result is

$$\langle q', t' | \hat{Q}(\bar{t}_1) \hat{Q}(\bar{t}_2) | q, t \rangle = \int_{q, q'} \mathcal{D}q q(\bar{t}_1) q(\bar{t}_2) \exp \left[i \int_t^{t'} d\tau L(q, \dot{q}) \right]. \quad (19.3.11)$$

Here we see something interesting. On the right hand side it doesn't matter if we write $q(\bar{t}_1)q(\bar{t}_2)$ or $q(\bar{t}_2)q(\bar{t}_1)$, since these are just numbers. However, the order of $\hat{Q}(\bar{t}_1)\hat{Q}(\bar{t}_2)$ is important. This ambiguity is due to the assumption that $\bar{t}_2 < \bar{t}_1$. If this isn't the case then the path integral here instead represents

$$\langle q', t' | \hat{Q}(\bar{t}_2)\hat{Q}(\bar{t}_1) | q, t \rangle. \quad (19.3.12)$$

We can compactly summarise this by writing

$$\langle q', t' | T[\hat{Q}(\bar{t}_1)\hat{Q}(\bar{t}_2)] | q, t \rangle = \int_{q, q'} \mathcal{D}q q(\bar{t}_1)q(\bar{t}_2) \exp \left[i \int_t^{t'} d\tau L(q, \dot{q}) \right] \quad (19.3.13)$$

where

$$T[A(t)B(t')] = \theta(t - t')A(t)B(t) + \theta(t' - t)B(t)A(t) \quad (19.3.14)$$

is the time ordered product of the operators $A(t)$ and $B(t')$.

Taking this process to its logical conclusion we have

$$\begin{aligned} \langle q', t' | T[\hat{Q}(\bar{t}_1) \cdots \hat{Q}(\bar{t}_\ell)] | q, t \rangle = \\ \int_{q, q'} \mathcal{D}q q(\bar{t}_1) \cdots q(\bar{t}_\ell) \exp \left[i \int_t^{t'} d\tau L(q, \dot{q}) \right]. \end{aligned} \quad (19.3.15)$$

Twenty

Generating Functions

20.1 Functional Derivatives

Consider a **functional**, F , which associates a number, $F[u]$, to a function, $u : \mathbb{R}^D \rightarrow \mathbb{R}$. The functional derivative, $\delta F / \delta u(x)$, is defined such that if we vary the function u ,

$$u(x) \mapsto u(x) + \delta u(x), \quad (20.1.1)$$

then the variation in F is

$$\delta F = F[u + \delta u] - F[u] = \int d^D x \frac{\delta F}{\delta u(x)} \delta u(x). \quad (20.1.2)$$

Compare this to the case of a function $f : \mathbb{R}^D \rightarrow \mathbb{R}$ when we vary each argument according to $x_k \mapsto x_k + \delta x_k$ and so f varies as

$$\delta f = f(x + \delta x) - f(x) = \sum_k \frac{\partial f}{\partial x_k} \delta x_k. \quad (20.1.3)$$

In this function case we have the identity

$$\frac{\partial}{\partial x_j} x_i = \delta_{ij}, \quad (20.1.4)$$

and in the functional case we have the analogous identity

$$\frac{\delta}{\delta f(x)} f(y) = \delta(x - y). \quad (20.1.5)$$

This, and the fact that the product rule and chain rule also apply to functional derivatives, is pretty much all we'll need.

20.2 Sources

Consider two time dependent functions, f and h . We can add terms proportional to these into our path integral. We call such terms source terms. In particular, we'll be interested in the path integral with sources given below:

$$\begin{aligned} & \langle q', t' | q, t \rangle_{f,h} \\ & := \int \mathcal{D}p \mathcal{D}q \exp \left\{ i \int_t^{t'} d\tau [p(\tau) \dot{q}(\tau) - H(q(\tau), p(\tau))] + f(\tau) q(\tau) + h(\tau) p(\tau) \right\}. \end{aligned} \quad (20.2.1)$$

Consider the following functional derivative:

$$A(\bar{\tau}) = \frac{\delta}{\delta f(\bar{\tau})} \exp \left[i \int_t^{t'} d\tau f(\tau) q(\tau) \right] \quad (20.2.2)$$

$$= \left(\frac{\delta}{\delta f(\bar{\tau})} i \int_t^{t'} d\tau f(\tau) q(\tau) \right) \exp \left[i \int_t^{t'} d\tau f(\tau) q(\tau) \right], \quad (20.2.3)$$

where the last step is an application of the chain rule. We now assume we can move the derivative through the integral to get

$$A(\bar{\tau}) = \left(i \int_t^{t'} d\tau \frac{\delta}{\delta f(\bar{\tau})} f(\tau) q(\tau) \right) \exp \left[i \int_t^{t'} d\tau f(\tau) q(\tau) \right] \quad (20.2.4)$$

$$= \left(i \int_t^{t'} d\tau \delta(\tau - \bar{\tau}) q(\tau) \right) \exp \left[i \int_t^{t'} d\tau f(\tau) q(\tau) \right] \quad (20.2.5)$$

$$= i q(\bar{\tau}) \exp \left[i \int_t^{t'} d\tau f(\tau) q(\tau) \right]. \quad (20.2.6)$$

Hence, we can generate factors of $q(\bar{\tau})$ through

$$q(\bar{\tau}) \exp \left[i \int_t^{t'} d\tau f(\tau) q(\tau) \right] = \frac{1}{i} \frac{\delta}{\delta f(\bar{\tau})} \exp \left[i \int_t^{t'} d\tau f(\tau) q(\tau) \right]. \quad (20.2.7)$$

We can also apply this to $\langle q', t' | q, t \rangle_{f,h}$ to get factors of $q(\bar{\tau})$, since all of the terms without $f(\tau)$ in them don't interact with the functional derivative. That is, we have

$$\begin{aligned} \frac{1}{i} \frac{\delta}{\delta f(\bar{\tau})} \langle q', t' | q, t \rangle_{f,h} &:= \int \mathcal{D}p \mathcal{D}q \, q(\bar{\tau}) \\ &\exp \left\{ i \int_t^{t'} d\tau [p(\tau) \dot{q}(\tau) - H(q(\tau), p(\tau))] + f(\tau) q(\tau) + h(\tau) p(\tau) \right\}. \end{aligned} \quad (20.2.8)$$

We can apply the functional derivative again to get another factor,

$$\begin{aligned} \left(\frac{1}{i} \frac{\delta}{\delta f(\bar{\tau}_1)} \right) \left(\frac{1}{i} \frac{\delta}{\delta f(\bar{\tau}_2)} \right) \langle q', t' | q, t \rangle_{f,h} &:= \int \mathcal{D}p \mathcal{D}q \, q(\bar{\tau}_1) q(\bar{\tau}_2) \\ &\exp \left\{ i \int_t^{t'} d\tau [p(\tau) \dot{q}(\tau) - H(q(\tau), p(\tau))] + f(\tau) q(\tau) + h(\tau) p(\tau) \right\}. \end{aligned} \quad (20.2.9)$$

Repeating this we have

$$\begin{aligned} \left(\frac{1}{i} \frac{\delta}{\delta f(\bar{\tau}_1)} \right) \cdots \left(\frac{1}{i} \frac{\delta}{\delta f(\bar{\tau}_n)} \right) \langle q', t' | q, t \rangle_{f,h} &:= \int \mathcal{D}p \mathcal{D}q \, q(\bar{\tau}_1) \cdots q(\bar{\tau}_n) \\ &\exp \left\{ i \int_t^{t'} d\tau [p(\tau) \dot{q}(\tau) - H(q(\tau), p(\tau))] + f(\tau) q(\tau) + h(\tau) p(\tau) \right\}. \end{aligned} \quad (20.2.10)$$

Comparing this to [Equation \(19.3.15\)](#) we see that

$$\langle q', t' | T[\hat{Q}(t_1) \cdots \hat{Q}(t_n)] | q, t \rangle = \left(\frac{1}{i} \frac{\delta}{\delta f(t_1)} \right) \cdots \left(\frac{1}{i} \frac{\delta}{\delta f(t_n)} \right) \langle q', t' | q, t \rangle \Big|_{f=h=0}. \quad (20.2.11)$$

20.3 Projection onto the Ground State

Often it is useful to compute the amplitude for a system to evolve from the vacuum state at time t to the vacuum state at time t' under some external sources, f and h . That is, we want to compute $\langle 0, t' | 0, t \rangle$, where $|0, t\rangle$ is the vacuum state at time t . Note that this is an eigenstate of the Hamiltonian, *not* the position. We can insert a complete set of states twice to get

$$\langle 0, t' | 0, t \rangle_{f,h} = \int dq dq' \langle 0, t' | q', t' \rangle_{f,h} \langle q', t' | q, t \rangle_{f,h} \langle q, t | 0, t \rangle_{f,h} \quad (20.3.1)$$

$$= \int dq dq' \varphi_0^*(q') \langle q', t' | q, t \rangle_{f,h} \varphi_0(q). \quad (20.3.2)$$

Here we use the ground state wave function in the position basis:

$$\varphi_0(q) = \langle q, t | 0, t \rangle, \implies \varphi_0^*(q) = \langle 0, t | q, t \rangle. \quad (20.3.3)$$

More generally if \hat{H} is the Hamiltonian for our system then we have the energy eigenstate $|n\rangle$, which is such that

$$\hat{H}|n\rangle = E_n, \quad (20.3.4)$$

and the energy eigenfunction

$$\varphi_n(q) = \langle q | n \rangle. \quad (20.3.5)$$

We assume that the energy levels are defined such that the ground state energy vanishes, that is $E_0 = 0$.

Now suppose that f and h have support in $[t, t']$. That is, outside the interval $[t, t']$ f and h vanish. Suppose also that we have times T and T' such that $T < t < t' < T'$. Then between T and t , and t' and T' , the sources vanish and we can easily compute $\langle q, t | Q, T \rangle$ by inserting a complete set of energy eigenstates:

$$\langle q, t | Q, T \rangle_{f,h} = \langle q | e^{-i\hat{H}(t-T)} | Q \rangle \quad (20.3.6)$$

$$= \sum_n \langle q | e^{-i\hat{H}(t-T)} | n \rangle \langle n | Q \rangle \quad (20.3.7)$$

$$= \sum_n e^{-iE_n(t-T)} \langle q | n \rangle \langle n | Q \rangle \quad (20.3.8)$$

$$= \sum_n e^{-iE_n(t-T)} \varphi_n(q) \varphi_n^*(Q). \quad (20.3.9)$$

We can analytically extend this result to time with a small imaginary part, $T_I = (1 - i\varepsilon)T$ where $\varepsilon > 0$. We then have

$$e^{-iE_n(t-T_I)} = e^{-iE_n(t-T+i\varepsilon T)} = e^{-iE_n(t-T)} e^{\varepsilon E_n T}. \quad (20.3.10)$$

If we consider the limit $T \rightarrow -\infty$ then all states apart from the ground state, where $E_0 = 0$, will be exponentially suppressed. We then have

$$\lim_{T \rightarrow -\infty} \langle q, t | Q, T_I \rangle_{f,h} = \langle q | 0 \rangle \langle 0 | Q \rangle = \varphi_0(q) \varphi_0^*(Q). \quad (20.3.11)$$

We can do the same with $\langle Q', T' | q', t' \rangle$, we get

$$\langle Q', T' | q', t' \rangle_{f,h} = \langle Q | e^{-i\hat{H}(T'-t')} | t \rangle \quad (20.3.12)$$

$$= \sum_n \langle Q' | e^{-i\hat{H}(T'-t')} | n \rangle \langle n | q' \rangle \quad (20.3.13)$$

$$= \sum_n e^{-iE_n(T'-t')} \langle Q' | n \rangle \langle n | q' \rangle \quad (20.3.14)$$

$$= \sum_n e^{-iE_n(T'-t')} \varphi_n(Q') \varphi_n^*(q'). \quad (20.3.15)$$

Analytically extending this result to $T'_I = (1 - i\varepsilon)T'$, and so

$$e^{-iE_n(T'_I - t')} = e^{-iE_n(T' - i\varepsilon T' - t')} e^{-iE_n(T' - t')} e^{-\varepsilon E_n T'}, \quad (20.3.16)$$

so this time all states apart from the ground state are exponentially suppressed in the limit $T' \rightarrow \infty$. We then have

$$\lim_{T' \rightarrow \infty} \langle Q', T'_I | q', t' \rangle_{f,h} = \langle Q' | 0 \rangle \langle 0 | q' \rangle = \varphi_0(Q') \varphi_0^*(q'). \quad (20.3.17)$$

Putting these two results together we find that

$$\lim_{\substack{T' \rightarrow -\infty \\ T' \rightarrow +\infty}} \frac{\langle Q', (1 - i\varepsilon)T' | Q, (1 - i\varepsilon)T' \rangle_{f,h}}{\varphi_0^*(Q) \varphi_0(Q')} = \langle 0, t' | 0, t \rangle_{f,h}. \quad (20.3.18)$$

Notice that we can get the same result by the transformation $\hat{H} \mapsto (1 - i\varepsilon)\hat{H}$. Note that the only dependence on the boundary is in the derivative on the left hand side, which is simply a normalisation factor which doesn't effect the values of derivatives with respect to the sources.

20.4 Weyl Ordering

For more complicated Hamiltonians we may have to consider terms with products of \hat{P} and \hat{Q} . We then need a prescription for ordering terms in the Hamiltonian so that the amplitude matches the path integral. We've used the mid-point prescription, evaluating the potential at $\tilde{q}_k = (q_k + q_{k-1})/2$, to the same effect in this simple case. In the more complex case we use the **Weyl ordering**. Here the order of a product of powers of some operators, \hat{A} and \hat{B} , is defined by considering the operator $(\alpha\hat{A} + \beta\hat{B})^n$ and expanding in powers of α and β :

$$(\alpha\hat{A} + \beta\hat{B})^n = \sum_k \frac{n!}{k!\ell!} \alpha^k \beta^\ell [\hat{A}^k \hat{B}^\ell]. \quad (20.4.1)$$

Here $[\hat{A}^k \hat{B}^\ell]$ is the Weyl ordered product. Notice that $\ell = n - k$ and

$$\frac{n!}{k!\ell!} = \binom{n}{k} = \binom{n}{\ell} \quad (20.4.2)$$

Example 20.4.3 Suppose we want to compute $[\hat{A}\hat{B}]$. This is quadratic in the operators so we consider $(\alpha\hat{A} + \beta\hat{B})^2$:

$$(\alpha\hat{A} + \beta\hat{B})^2 = \alpha^2\hat{A} + \alpha\beta(\hat{A}\hat{B} + \hat{B}\hat{A}) + \beta\hat{B}^2. \quad (20.4.4)$$

We then take the terms proportional to $\alpha\beta$ and, after dividing by $n!/(k!\ell!) = 2!/(1!1!) = 2$ we get the Weyl ordering,

$$[\hat{A}\hat{B}] = \frac{1}{2}(\hat{A}\hat{B} + \hat{B}\hat{A}). \quad (20.4.5)$$

Now suppose we want to compute all third order Weyl orderings. We consider

$$\begin{aligned} (\alpha\hat{A} + \beta\hat{B})^3 &= \alpha^3\hat{A}^3 + \alpha^2\beta(\hat{A}^2\hat{B} + \hat{A}\hat{B}\hat{A} + \hat{B}\hat{A}^2) \\ &\quad + \alpha\beta^2(\hat{A}\hat{B}^2 + \hat{B}\hat{A}\hat{B} + \hat{B}^2\hat{A}) + \beta^3\hat{B}^3. \end{aligned} \quad (20.4.6)$$

We then have

$$[\hat{A}^3] = \frac{3!1!}{3!}\hat{A}^3 = \hat{A}^3, \quad (20.4.7)$$

$$[\hat{A}^2\hat{B}] = \frac{2!1!}{3!}(\hat{A}^2\hat{B} + \hat{A}\hat{B}\hat{A} + \hat{B}\hat{A}^2) = \frac{1}{3}(\hat{A}^2\hat{B} + \hat{A}\hat{B}\hat{A} + \hat{B}\hat{A}^2), \quad (20.4.8)$$

$$[\hat{A}\hat{B}^2] = \frac{1!2!}{3!}(\hat{A}\hat{B}^2 + \hat{B}\hat{A}\hat{B} + \hat{B}^2\hat{A}) = \frac{1}{3}(\hat{A}\hat{B}^2 + \hat{B}\hat{A}\hat{B} + \hat{B}^2\hat{A}), \quad (20.4.9)$$

$$[\hat{B}^3] = \frac{1!3!}{3!}\hat{B}^3 = \hat{B}^3. \quad (20.4.10)$$

Finally, suppose we want to compute $[\hat{A}^2\hat{B}^2]$. Expanding $(\alpha\hat{A} + \beta\hat{B})^5$ we find the relevant term is

$$\begin{aligned} \alpha^3\beta^2 &(\hat{A}^3\hat{B}^2 + \hat{A}^2\hat{B}\hat{A}\hat{B} + \hat{A}^2\hat{B}^2\hat{A} + \hat{A}\hat{B}\hat{A}^2\hat{B} + \hat{A}\hat{B}\hat{A}\hat{B}\hat{A} \\ &\quad + \hat{A}\hat{B}^2\hat{A}^2 + \hat{B}\hat{A}^3\hat{B} + \hat{B}\hat{A}^2\hat{B}\hat{A} + \hat{B}\hat{A}\hat{B}\hat{A}^2 + \hat{B}^2\hat{A}^3) \end{aligned} \quad (20.4.11)$$

and so the Weyl ordered product is this result divided by $5!/(3!2!) = 10$

$$\begin{aligned} [\hat{A}^3\hat{B}^2] &= \frac{1}{10}(\hat{A}^3\hat{B}^2 + \hat{A}^2\hat{B}\hat{A}\hat{B} + \hat{A}^2\hat{B}^2\hat{A} + \hat{A}\hat{B}\hat{A}^2\hat{B} \\ &\quad + \hat{A}\hat{B}\hat{A}\hat{B}\hat{A} + \hat{A}\hat{B}^2\hat{A}^2 + \hat{B}\hat{A}^3\hat{B} + \hat{B}\hat{A}^2\hat{B}\hat{A} + \hat{B}\hat{A}\hat{B}\hat{A}^2 + \hat{B}^2\hat{A}^3) \end{aligned} \quad (20.4.12)$$

Part VII

Path Integrals in Scalar Field Theory

Twenty-One

Noninteracting Scalar Field Theory

21.1 Scalar Field Theory Recap

The Lagrangian for a free real scalar field, φ , is

$$\mathcal{L}_0(\varphi, \partial_\mu \varphi) = \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2 \varphi^2. \quad (21.1.1)$$

Applying the Euler–Lagrange equations,

$$\partial_\mu \left(\frac{\partial \mathcal{L}_0}{\partial(\partial_\mu \varphi)} \right) - \frac{\partial \mathcal{L}_0}{\partial \varphi} = 0 \quad (21.1.2)$$

gives the Klein–Gordon equation,

$$(\partial^2 + m^2)\varphi = 0. \quad (21.1.3)$$

The action associated with this Lagrangian is, in D dimensions,

$$S_0[\varphi] = \int d^D x \left[\frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2 \varphi^2 \right]. \quad (21.1.4)$$

There is already some physics going on here. The Lagrangian and action tell us what the degrees of freedom of the system are, in this case we just have one, the field, φ . We can also examine the symmetries of the action. Here we have one symmetry, Lorentz invariance. Recall that if x transforms as

$$x^\mu \mapsto x'^\mu = \Lambda^\mu_\nu x^\nu \iff x \mapsto x' = \Lambda x \quad (21.1.5)$$

for some Lorentz transformation Λ then the field transforms as

$$\varphi(x) \mapsto \varphi'(x') = \varphi(x), \quad (21.1.6)$$

since by definition the field is a scalar. Writing x' as $\Lambda x'$ in the middle of this we have $\varphi(x) \mapsto \varphi'(\Lambda x)$. Redefining our variable so that x is the position in the new frame we have $\varphi(x) \mapsto \varphi'(x) = \varphi(\Lambda^{-1}x)$. If this is confusing think of a one-dimensional function, $f: \mathbb{R} \rightarrow \mathbb{R}$. If we want to translate this function to the right, in the positive x direction, by, say a units, then we transform it as $f(x) \mapsto f'(x) = f(x - a)$, so we subtract. The logic is the same here. Our transformation moves the point of evaluation, so we have to move it back.

We define the conjugate momentum as

$$\Pi(x) = \frac{\partial \mathcal{L}_0}{\partial(\partial_0 \varphi)} = \partial_0 \varphi(x) = \dot{\varphi}(x). \quad (21.1.7)$$

The Hamiltonian is then

$$\mathcal{H}(\varphi, \Pi) = \Pi(x)\dot{\varphi}(x) - \mathcal{L}_0(\varphi, \partial_\mu \varphi) \quad (21.1.8)$$

$$= \frac{1}{2}\Pi(x)^2 + \frac{1}{2} \sum_{k=1}^3 (\partial_k \varphi(x))^2 + \frac{1}{2}m^2 \varphi(x)^2. \quad (21.1.9)$$

21.2 The Path Integral

We could derive the path integral in scalar field theory in a very similar way to how we derived it for quantum mechanics, but we won't. Instead we'll assume the form is the same after making the usual substitutions which take us from quantum mechanics to quantum field theory. These substitutions are as follows:

- $t \rightarrow x$, we change from a single coordinate to D coordinates;
- $q(t) \rightarrow \varphi(x)$, we change from position to the field, φ ;
- $\hat{Q}(t) \rightarrow \hat{\varphi}(x)$, we change from the position operator to the field operator.

Recall the quantum mechanical path integral:

$$\langle q, t | q', t' \rangle_f = \int \mathcal{D}q \exp\{i(S_0[q] + qf)\}. \quad (21.2.1)$$

Moving to scalar field theory we replace this with

$$Z_0[J] := \langle 0 | 0 \rangle_J = \int \mathcal{D}\varphi \exp\{i(S_0[\varphi] + J \cdot \varphi)\} \quad (21.2.2)$$

where

$$J \cdot \varphi := \int d^D x J(x) \varphi(x). \quad (21.2.3)$$

This path integral is just a Gaussian integral, the one complication being that it's an infinite dimensional Gaussian integral. We can make this more obvious with a bit of rewriting. First, consider the action:

$$S_0[\varphi] := \int d^D x \mathcal{L}_0 = \frac{1}{2} \int d^D x [(\partial_\mu \varphi)(\partial^\mu \varphi) - m^2 \varphi^2]. \quad (21.2.4)$$

Take the first term and integrate by parts to get

$$\int d^D x (\partial_\mu \varphi)(\partial^\mu \varphi) = \int d^D x [\partial_\mu (\varphi \partial^\mu \varphi) - \varphi \partial^2 \varphi] = - \int d^D x \varphi \partial^2 \varphi \quad (21.2.5)$$

where in the last step we assume that the field vanishes sufficiently quickly at infinity that we can ignore the boundary term. The action is then

$$S_0[\varphi] = \frac{1}{2} \int d^D x \varphi(x) [-\partial^2 - m^2] \varphi(x). \quad (21.2.6)$$

Now, we insert a factor of $-i\varepsilon\varphi^2$ for some $\varepsilon > 0$, which we will later take to zero. This ensures convergence. The modified action is then

$$S_0[\varphi] = \frac{1}{2} \int d^D x \varphi(x) [-\partial^2 - m^2 - i\varepsilon] \varphi(x). \quad (21.2.7)$$

Now define

$$K(x, x') := \delta(x - x') [-\partial^2 - m^2]. \quad (21.2.8)$$

Then we can write the action as

$$S_0[\varphi] = \frac{1}{2} \int d^D x d^D x' \varphi(x) K(x, x') \varphi(x'). \quad (21.2.9)$$

Hence,

$$Z_0[0] = \int \mathcal{D}\varphi \exp \left[i \frac{1}{2} \int d^D x d^D x' \varphi(x) K(x, x') \varphi(x') \right]. \quad (21.2.10)$$

$$Z_A = \int d^n x \exp \left[-\frac{1}{2} x^T A x \right] = \int d^n x \exp \left[-\sum_i \sum_j x_i A_{ij} x_j \right]. \quad (21.2.11)$$

We can see that the kernel, $K(x, x')$, plays the role of A and our sums become integrals, but other than that not much changes.

It's easier to compute the action if we move to momentum space, since in momentum space $\partial^2 \varphi$ simply becomes $-p^2 \tilde{\varphi}$. To do this we write φ as the inverse transform of $\tilde{\varphi}$:

$$\varphi(x) = \int \frac{d^D p}{(2\pi)^D} e^{-ip \cdot x} \tilde{\varphi}(p). \quad (21.2.12)$$

We then have

$$S_0[\varphi] = \frac{1}{2} \int_x \int_{p, p'} e^{-ip \cdot x} \tilde{\varphi}(p) (p'^2 - m^2 + i\varepsilon) e^{-ip' \cdot x} \tilde{\varphi}(p'). \quad (21.2.13)$$

Here we've introduced the shorthand notation

$$\int_x := \int d^D x, \quad \int_p := \int \frac{d^D p}{(2\pi)^D}, \quad \text{and} \quad \int_{p, p'} = \int \frac{d^D p}{(2\pi)^D} \int \frac{d^D p'}{(2\pi)^D} \quad (21.2.14)$$

and so on. We can combine the two exponentials into $\exp[-i(p + p') \cdot x]$, and then performing the x integral gives us a Dirac delta:

$$S_0[\varphi] = \frac{1}{2} \int_{p, p'} \tilde{\varphi}(p) (p'^2 - m^2 + i\varepsilon) \tilde{\varphi}(p') \delta(p + p'). \quad (21.2.15)$$

Using the Dirac delta we can perform the p' integral to get

$$S_0[\varphi] = \frac{1}{2} \int_p \tilde{\varphi}(p) (p^2 - m^2 + i\varepsilon) \tilde{\varphi}(-p). \quad (21.2.16)$$

We can simplify this slightly using a Fourier transform identity. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a real function whose Fourier transform exists. Then we have

$$\tilde{f}(p) = \int dx e^{ipx} f(x). \quad (21.2.17)$$

Hence, we have

$$\tilde{f}(-p) = \int dx e^{-ipx} f(x) = \left[\int dx e^{ipx} f(x) \right]^* = \tilde{f}(p)^*. \quad (21.2.18)$$

That is, for a real function, $f(-p) = f(p)^*$. This same result holds in arbitrary dimensions by an identical proof.

Using this we can write the action as

$$S_0[\varphi] = \frac{1}{2} \int_p \tilde{\varphi}(p)(p^2 - m^2 + i\varepsilon)\tilde{\varphi}(p)^*. \quad (21.2.19)$$

21.2.1 With a Source Term

Now we want to compute

$$Z_0[J] = \int \mathcal{D}\varphi \exp \left[i \frac{1}{2} \int_p \tilde{\varphi}(p)(p^2 - m^2 + i\varepsilon)\tilde{\varphi}(p)^* + \int d^D x J(x)\varphi(x) \right]. \quad (21.2.20)$$

This is the infinite dimensional analogue of

$$Z_A(b) = \int d^n x \exp \left[-\frac{1}{2} x^T A x + b^T x \right] \quad (21.2.21)$$

$$= \int d^n x \exp \left[-\frac{1}{2} \sum_i \sum_j x_i A_{ij} x_j + \sum_i b_i x_i \right]. \quad (21.2.22)$$

The source term simply has the sum turn into an integral. To compute this integral we completed the square by making a substitution, $x = y + \Delta b$, where $\Delta = A^{-1}$. We can do the same here, making the substitution in momentum space:

$$\tilde{\chi}(p) = \tilde{\varphi}(p) + \frac{J(p)}{p^2 - m^2 + i\varepsilon}. \quad (21.2.23)$$

Notice that it's much easier to do this in Fourier space since we can literally divide by $p^2 - m^2 + i\varepsilon$. To do this in real space we'd have to somehow invert the operator $-\partial^2 - m^2 + i\varepsilon$. The finite dimensional analogue of this is that going to Fourier space is a basis change into a basis in which A is diagonal, and so easily inverted. In a sense $\tilde{K}(p, p')$ is "diagonal" in momentum space, in particular we get independent modes, one for each momentum.

Before we do this we should Fourier transform the source term as well. We have

$$J(x) = \int \frac{d^D p}{(2\pi)^D} e^{-ip \cdot x} \tilde{J}(p). \quad (21.2.24)$$

Substituting this, and the Fourier transform of the field, into the source term we have

$$\int_x J(x)\varphi(x) = \int_x \int_{p,p'} \tilde{J}(p)\tilde{\varphi}(p')e^{i(p+p')\cdot x} \quad (21.2.25)$$

$$= \int_{p,p'} \tilde{J}(p)\tilde{\varphi}(p')(2\pi)^D \delta(p+p') \quad (21.2.26)$$

$$= \int_p \tilde{J}(p)\tilde{\varphi}(-p). \quad (21.2.27)$$

Hence,

$$S_0[\varphi] + J \cdot \varphi = \frac{1}{2} \int_p [\tilde{\varphi}(p)(p^2 - m^2 + i\varepsilon)\tilde{\varphi}(-p) + \tilde{J}(p)\tilde{\varphi}(-p) + \tilde{J}(-p)\tilde{\varphi}(p)]. \quad (21.2.28)$$

The seemingly extra term of $\tilde{J}(-p)\tilde{\varphi}(p)$ comes from factoring out the half, giving two $\tilde{J}(p)\tilde{\varphi}(-p)$ terms. We then recognise that since the integration range is symmetric all that matters is that the arguments of \tilde{J} and $\tilde{\varphi}$ in this term differ by a sign. So, we can make the substitution $p \rightarrow -p$ in one of the factors then giving the final term above.

Making this change of variables we can add the source term to the action:

$$S_0[\varphi] + J \cdot \varphi = \frac{1}{2} \int_p \left[\left(\tilde{\chi}(p) - \frac{\tilde{J}(p)}{p^2 - m^2 + i\varepsilon} \right) \right] \quad (21.2.29)$$

$$\times (p^2 - m^2 + i\varepsilon) \left(\tilde{\chi}(-p) - \frac{\tilde{J}(-p)}{p^2 - m^2 + i\varepsilon} \right) \quad (21.2.30)$$

$$+ \tilde{J}(p) \left(\tilde{\chi}(p) - \frac{\tilde{J}(p)}{p^2 - m^2 + i\varepsilon} \right) \quad (21.2.31)$$

$$+ \tilde{J}(-p) \left(\tilde{\chi}(-p) - \frac{\tilde{J}(-p)}{p^2 - m^2 + i\varepsilon} \right) \Big] \quad (21.2.32)$$

Expanding this most terms cancel and we're left with

$$S_0[\varphi] + J \cdot \varphi = \frac{1}{2} \int_p \left[\tilde{\chi}(p)(p^2 - m^2 + i\varepsilon)\tilde{\chi}(-p) - \tilde{J}(p)\frac{1}{p^2 - m^2 + i\varepsilon}\tilde{J}(-p) \right]. \quad (21.2.33)$$

Hence, we have

$$Z_0[J] = \int \mathcal{D}\chi \exp \left\{ i \frac{1}{2} \int_p \left[\tilde{\chi}(p)(p^2 - m^2 + i\varepsilon)\tilde{\chi}(-p) \right. \right. \quad (21.2.34)$$

$$\left. - \tilde{J}(p)\frac{1}{p^2 - m^2 + i\varepsilon}\tilde{J}(-p) \right] \Big\} \quad (21.2.35)$$

$$= \exp \left[-\frac{i}{2} \int_p \tilde{J}(p)\frac{1}{p^2 - m^2 + i\varepsilon}\tilde{J}(-p) \right] \quad (21.2.36)$$

$$\times \int \mathcal{D}\chi \exp \left[\frac{i}{2} \int_p \tilde{\chi}(p)(p^2 - m^2 + i\varepsilon) \tilde{\chi}(-p) \right] \quad (21.2.37)$$

$$= \mathcal{N} \exp \left[-\frac{i}{2} \int_p \tilde{J}(p) \frac{1}{p^2 - m^2 + i\varepsilon} \tilde{J}(-p) \right] \quad (21.2.38)$$

where

$$\mathcal{N} = \int \mathcal{D}\chi \exp \left[\frac{i}{2} \int_p \tilde{\chi}(p)(p^2 - m^2 + i\varepsilon) \tilde{\chi}(-p) \right] \quad (21.2.39)$$

is a normalisation factor. This normalisation factor is analogous to the $(\det A)^{-1/2}$ factor in the finite dimensional case. Exactly how we're taking the determinant of an infinite dimensional operator and then inverting this is not important, since we always consider ratios, $Z_0[J]/Z_0[0]$, where this normalisation vanishes.

We can write this result as

$$Z_0[J] = \mathcal{N} \exp \left[-\frac{i}{2} \int d^D x d^D x' J(x) \Delta(x, x') J(x') \right] \quad (21.2.40)$$

where

$$\Delta(x, x') = \Delta(x - x') = \int_p \frac{e^{-ip \cdot (x - x')}}{p^2 - m^2 + i\varepsilon} \quad (21.2.41)$$

is the **Feynman propagator**. We find this by doing the inverse transform

$$\tilde{J}(p) = \int d^D x e^{ip \cdot x} J(x), \quad \text{and} \quad \tilde{J}(-p) = \int d^D x e^{-ip \cdot x} J(x), \quad (21.2.42)$$

so

$$\begin{aligned} \int_p \tilde{J}(-p) \frac{1}{p^2 - m^2 + i\varepsilon} \tilde{J}(p) &= \int_p \int_{x, x'} e^{-ip \cdot x} J(x) \frac{1}{p^2 - m^2 + i\varepsilon} e^{ip \cdot x'} J(x') \\ &= \int_p \int_{x, x'} J(x) \frac{1}{p^2 - m^2 + i\varepsilon} J(x') e^{-ip \cdot (x - x')}. \end{aligned}$$

We can show that, in analogy to the finite dimensional integrals, Δ is, in a sense, the inverse of K , more specifically,

$$I = \int d^D z K(x, z) \Delta(z, x') = \delta(x - x'). \quad (21.2.43)$$

This follows by substituting in the definitions:

$$I = \int d^D z \int_p \delta(x - z) [-\partial^2 - m^2 - i\varepsilon] \frac{e^{-ip \cdot (z - x')}}{p^2 - m^2 + i\varepsilon} \quad (21.2.44)$$

$$= \int d^D z \delta(x - z) [-\partial^2 - m^2 - i\varepsilon] \frac{\delta(z - x')}{p^2 - m^2 + i\varepsilon} \quad (21.2.45)$$

The two terms $[-\partial^2 - m^2 - i\varepsilon]$ and $1/(p^2 - m^2 + i\varepsilon)$ cancel as they are inverses once we move the first to Fourier space, and we are left with the product $\delta(x - z)\delta(z - x')$, which reduces to $\delta(x - x')$ when we perform the z integral.

21.2.2 Using the Path Integral

We can use the path integral to compute correlators, much like in [Section 20.2](#). We normalise so that $\mathcal{N} = 1$. Consider, for example, the vacuum expectation of the time ordered product of two fields:

$$\begin{aligned}
 \langle 0 | T[\varphi(x_1)\varphi(x_2)] | 0 \rangle &= \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x_2)} \right) Z_0[J] \Big|_{J=0} \\
 &= \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x_2)} \right) \exp \left[-\frac{i}{2} \int_{x,x'} J(x) \Delta(x, x') J(x') \right] \\
 &= \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left[-\frac{1}{2} \int_{x'} \Delta(x_2, x') J(x') - \frac{1}{2} \int_x J(x) \Delta(x, x_2) \right] \\
 &\quad \times \exp \left[-\frac{i}{2} \int_{x,x'} J(x) \Delta(x, x') J(x') \right] \\
 &= -\frac{1}{i} \frac{1}{2} \{ \Delta(x_1, x_2) + \Delta(x_2, x_1) \} \\
 &= i \Delta(x_1, x_2)
 \end{aligned}$$

where we set $J = 0$ in the penultimate term and then used the fact that Δ is symmetric in its arguments. This must be the case since $T[\varphi(x_1)\varphi(x_2)]$ is symmetric in x_1 and x_2 , and we've used this in the past to *define* Δ (see [Equation \(9.4.11\)](#)). We can show its the case using the definition here also, since

$$\Delta(x) = \int \frac{d^D p}{(2\pi)^D} \frac{e^{-ip \cdot x}}{p^2 - m^2 + i\epsilon} \stackrel{p \rightarrow -p}{=} \int \frac{d^D p}{(2\pi)^D} \frac{e^{ip \cdot x}}{p^2 - m^2 + i\epsilon} = \Delta(-x). \quad (21.2.46)$$

Note that even if D is odd we still don't have to change the overall sign since the integration region, \mathbb{R}^D , is even so we're integrating over all possible value of p or $-p$ either way so the change in sign cancels with exchanging the limits.

In terms of physics this symmetry is due to invariance under translations, which means that the propagator can only depend on the *relative* positions of its arguments, not on the actual positions, meaning we can always write it as a function of the difference in positions, and then all that changes is the sign if we swap the two arguments and the argument above tells us the sign is not important.

Twenty-Two

Interacting Scalar Field Theory

22.1 General Theory

In this chapter we consider a scalar field theory with some interaction given by the potential, V , which is a function of the field, φ . This is described by the Lagrangian

$$\mathcal{L} = \mathcal{L}_0 + V(\varphi) \quad (22.1.1)$$

$$= \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2 \varphi^2 + V(\varphi). \quad (22.1.2)$$

The action for this Lagrangian is then

$$S[\varphi] = S_0[\varphi] + S_{\text{int}}[\varphi] \quad (22.1.3)$$

$$= \int d^D x \left[\frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2 \varphi^2 \right] + \int d^D x V(\varphi). \quad (22.1.4)$$

We will be interested in computing the path integral

$$Z[J] = \langle 0|0 \rangle_J = \int \mathcal{D}\varphi \exp \left\{ i \left[S_0[\varphi] + \int d^D x V(\varphi) + J \cdot \varphi \right] \right\}, \quad (22.1.5)$$

where

$$J \cdot \varphi = \int d^D x J(x) \varphi(x). \quad (22.1.6)$$

Now expand V term in a power series, first write

$$Z[J] = \int \mathcal{D}\varphi \exp \left\{ i \int d^D x V(\varphi) \right\} \exp \{ i(S_0[\varphi] + J \cdot \varphi) \}. \quad (22.1.7)$$

Then expand the first exponential:

$$Z[J] = \int \mathcal{D}\varphi \sum_{n=0}^{\infty} \frac{1}{n!} \left[i \int d^D x V(\varphi(x)) \right]^n \exp \{ i(S_0[\varphi] + J \cdot \varphi) \}. \quad (22.1.8)$$

Now expand the potential as well, giving

$$Z[J] = \int \mathcal{D}\varphi \sum_{n=0}^{\infty} \frac{1}{n!} \left[i \int d^D x \sum_{k=0}^{\infty} \frac{1}{k!} \frac{\partial V}{\partial \varphi} \varphi^k \right]^n \exp \{ i(S_0[\varphi] + J \cdot \varphi) \}. \quad (22.1.9)$$

We can obtain each factor of φ by acting on the exponential with $\delta/\delta J(x)$. This also brings down a factor of i , so we include a factor of $1/i$ for each derivative giving

$$Z[J] = \int \mathcal{D}\varphi \sum_{n=0}^{\infty} \frac{1}{n!} \left[i \int d^D x \sum_{k=0}^{\infty} \frac{1}{k!} \frac{\partial V}{\partial \varphi} \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right)^k \right]^n \exp\{i(S_0[\varphi] + J \cdot \varphi)\}.$$

We can package the derivatives up into $V(-i\delta/\delta J(x))$, which is just a formal term standing for exactly the expansion above, giving

$$Z[J] = \int \mathcal{D}\varphi \sum_{n=0}^{\infty} \frac{1}{n!} \left[i \int d^D x V\left(\frac{1}{i} \delta/\delta J(x)\right) \right]^n \exp\{i(S_0[\varphi] + J \cdot \varphi)\}. \quad (22.1.10)$$

Finally we can package the exponential back up giving

$$Z[J] = \int \mathcal{D}\varphi \exp\left\{ i \int d^D x V\left(\frac{1}{i} \frac{d}{dJ(x)}\right) \right\} \exp\{i(S_0[\varphi] + J \cdot \varphi)\}. \quad (22.1.11)$$

Now the first exponential is independent of φ , so we can pull it outside of the integral to get

$$Z[J] = \exp\left\{ i \int d^D x V\left(\frac{1}{i} \frac{\delta}{\delta J(x)}\right) \right\} \int \mathcal{D}\varphi \exp\{i(S_0[\varphi] + J \cdot \varphi)\} \quad (22.1.12)$$

$$= \exp\left\{ i \int d^D x V\left(\frac{1}{i} \frac{\delta}{\delta J(x)}\right) \right\} Z_0[J], \quad (22.1.13)$$

where we recognise

$$Z_0[J] = \int \mathcal{D}\varphi \exp\{i(S_0[\varphi] + J \cdot \varphi)\} \quad (22.1.14)$$

as the amplitude for the noninteracting system.

22.2 φ^3 Interactions

We will consider as an example the potential

$$V(\varphi) = \frac{g}{3!} \varphi^3. \quad (22.2.1)$$

Then we can use

$$V\left(\frac{1}{i} \frac{\delta}{\delta J(x)}\right) = \frac{g}{3!} \left(\frac{1}{i} \frac{\delta}{\delta J(x)}\right)^3. \quad (22.2.2)$$

Expand the exponential and $Z_0[J]$ in power series, using [Equation \(21.2.40\)](#) for $Z_0[J]$, this gives

$$Z[J] = \sum_{V=0}^{\infty} \frac{1}{V!} \left[i \frac{g}{3!} \int_x \left(\frac{1}{i} \frac{\delta}{\delta J(x)}\right)^3 \right]^V \sum_{P=0}^{\infty} \frac{1}{P!} \left[-\frac{1}{2} \int_{y,z} J(y) i\Delta(y-z) J(z) \right]^P. \quad (22.2.3)$$

This is a perturbative expansion for $Z[J]$, and is a function of J , all spacetime points, x , y , and z , are integrated over.

Consider what happens for some fixed values of V and P . We have $2P$ factors of J , and $3V$ derivatives with respect to J , so we end up with $E = 2P - 3V$ factors of J .

We will again introduce a diagrammatic notation to write the result of integrals like these. First, a propagator is written as a line:

$$i\Delta(y - z) = x \text{ ————— } y \quad (22.2.4)$$

We denote an integral of $J(x)$ as solid dot at the end of a line:

$$i \int d^D x J(x) = x \bullet \text{ ————— } y \quad (22.2.5)$$

Note that the line also gives a propagator, which we're not worried about here. We denote an integral and a factor of g with a vertex, here with three legs, each of which again represents a propagator which we aren't worried about here:

$$ig \int d^D x = \text{---} x \text{---} \quad (22.2.6)$$

We have to carefully count how many times each term contributes to the overall amplitude. We need to count the number of ways we can change the term without changing anything. We have the following symmetries to account for:

- Permutations of the derivatives with respect to $J(x)$, evaluated at the same x . That is, if we have the $V = 2$ term

$$\frac{1}{2!} \left(i \frac{g}{3!} \int_{x_1} \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right)^3 \right) \left(i \frac{g}{3!} \int_{x_2} \left(\frac{1}{i} \frac{\delta}{\delta J(x_2)} \right)^3 \right) \dots \quad (22.2.7)$$

then we can exchange two of the derivatives with respect to $J(x_1)$, but we can't exchange a derivative with respect to $J(x_1)$ and a derivative with respect to $J(x_2)$. In total this symmetry gives a factor of $(3!)^V$.

- Permutations of propagators and relabelling dummy integration variables. There are P propagators (which is why we called this index P in the first place) and so we have a factor of $P!$.
- Swapping the ends of propagators, since they're symmetric functions, this gives a factor of 2^P .
- Permuting integrals over x in the first term, which corresponds to permuting vertices. There are V vertices (which is why we called this index V in the first place) and so we have a factor of $V!$.

Notice that Equation (22.2.3) has factors of $1/(3!)^V$, $1/P!$, $1/2^P$, and $1/V!$, and so all of these factors cancel out, indeed this is why we choose to have the factor of $3!$ and $1/2$ in the initial definitions. However, this only works if all of the operations described above are distinct, and this isn't always the case. If two of these operations give the same term then we don't want to double count them.

For example, consider the $V = 1, P = 2$ term:

$$i \frac{g}{3!} \int d^D x \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \\ \times \frac{1}{2!} \left(-\frac{1}{2} \right)^2 \int_{y_1, z_1} J(y_1) i\Delta(y_1 - z_1) J(z_1) \int_{y_2, z_2} J(y_2) i\Delta(y_2 - z_2) J(z_2).$$

Now focus on the contribution from this term when the second derivative acts on the first factor of J and the third derivative acts on the second factor of J :

$$i \frac{g}{3!} \int d^D x \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \\ \times \frac{1}{2!} \left(-\frac{1}{2} \right)^2 \int_{y_1, z_1} J(y_1) i\Delta(y_1 - z_1) J(z_1) \int_{y_2, z_2} J(y_2) i\Delta(y_2 - z_2) J(z_2).$$

Now swap the second and third derivative, which changes nothing, but keep track of which J they act on

$$i \frac{g}{3!} \int d^D x \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \\ \times \frac{1}{2!} \left(-\frac{1}{2} \right)^2 \int_{y_1, z_1} J(y_1) i\Delta(y_1 - z_1) J(z_1) \int_{y_2, z_2} J(y_2) i\Delta(y_2 - z_2) J(z_2).$$

We can now swap the first two factors of J , and make a change of variables $y_1 \leftrightarrow z_1$, giving

$$i \frac{g}{3!} \int d^D x \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right) \\ \times \frac{1}{2!} \left(-\frac{1}{2} \right)^2 \int_{y_1, z_1} J(y_1) i\Delta(y_1 - z_1) J(z_1) \int_{y_2, z_2} J(y_2) i\Delta(y_2 - z_2) J(z_2).$$

This is just the term we started with but with the colours swapped. This means that these two swaps are inverses, and so swapping the derivatives and reverse swapping the J s (which is just swapping them in this case where there are only 2) are the same operation, and so if we include both the factor of $P! = 2$ and $(3!)^V = 6$ then we'll be overcounting the contributions from this term.

To see this we can perform three derivatives in this term. They will each act on one of four factors of J , specifically one of $J(y_1)$, $J(z_1)$, $J(y_2)$, and $J(z_2)$. The first derivative will have four options for which J to act on, the second derivative will have three options, and the third will have two, so we include a factor of $4 \cdot 3 \cdot 2$. It doesn't matter which J we don't act on, since we can always make a change of variables to choose the argument of J to be any of y_1, z_1, y_2 and z_2 . So without loss of generality we assume that nothing acts on $J(y_1)$. Then the three other J s are acted on, and so each gives a Dirac delta, $\delta(x - \xi)$ for each $J(\xi)$. Hence this term becomes

$$i \frac{g}{3!} \frac{1}{2!} \left(-\frac{1}{2} \right)^2 4 \cdot 3 \cdot 2 \int_{x, y_1, z_1, y_2, z_2} J(y_1) \delta(x - z_1) \delta(x - y_2) \delta(x - z_2) i\Delta(y_1 - z_1) i\Delta(y_2 - z_2). \quad (22.2.8)$$

Cancelling some factors and using the Dirac deltas to perform three of the integrals we get

$$g \frac{1}{2} \int_{x, y_1} J(y_1) \Delta(y_1 - x) \Delta(x - x). \quad (22.2.9)$$

Now we can draw the diagram representing this integral, we have a single source, so one solid dot, at y_1 this is then connected by a propagator to x , and then we have another propagator from x to x , which gives a loop:



$$(22.2.10)$$

Notice that this diagram has a single mirror symmetry along the propagator line from y_1 to x . This 2-fold symmetry exactly corresponds to the factor of $1/2$ that we are left with in the integral. In general the integral will have a factor of 1 over the order of the symmetry group of the diagram.

22.3 Disconnected Diagrams

A disconnected diagram is any diagram made from two or more pieces with no propagators connecting them. We can consider some arbitrary diagram, D , which may or may not be connected. We can always decompose D as a product of its connected components, which will just be D itself if it's connected. Let C_I be the I th connected diagram and suppose it appears n_I times in D . Then diagram C_I contributes a factor of $C_I^{n_I}$ to D , up to a symmetry factor. We include in C_I the symmetry factor for this diagram, and so the only symmetry not accounted for is permutations of repeated connected subdiagrams. In particular we get a factor of $n_I!$ for each connected subdiagram. Thus, we can express D as

$$D = \frac{1}{S_D} = \prod_I C_I^{n_I}, \quad \text{where} \quad S_D = \prod_I n_I!. \quad (22.3.1)$$

We can then work to all orders in perturbation theory and express $Z[J]$ as a sum over all diagrams, D , which we can regard as a function of $\{n_I\}$, the number of times the connected diagram C_I appears as a subdiagram of D :

$$Z[J] = \sum_{\{n_I\}} D(\{n_I\}) = \sum_{\{n_I\}} \prod_I \frac{1}{n_I!} C_I^{n_I}. \quad (22.3.2)$$

We can swap the order of the product and, then the sum is over fixed I and we take all possible diagrams so n_I runs from 0 to ∞ , giving

$$Z[J] = \prod_I \sum_{n_I=0}^{\infty} \frac{1}{n_I!} C_I^{n_I} = \prod_I \exp[C_I] = \exp \left[\sum_I C_I \right]. \quad (22.3.3)$$

When normalising $Z[J]$ we choose to have $Z[0] = 1$, which corresponds to replacing $Z[J]$ with $Z[J]/Z[0]$. Since $Z[0]$ consists of all diagrams with no sources we have

$$Z[0] = \exp \left[\sum_{I|E=0} C_I \right] \quad (22.3.4)$$

where the sum is taken over diagrams with $E = 0$. We then have

$$\frac{Z[J]}{Z[0]} = \exp \left[\sum_I C_I \right] \exp \left[- \sum_{I|E=0} C_I \right] = \exp \left[\sum_{I|E \neq 0} C_I \right] = \exp \left[\sum_I' C_I \right] \quad (22.3.5)$$

where the sum in the last two is taken over all connected diagrams with no sources.

Now consider the log of this value:

$$\log \left(\frac{Z[J]}{Z[0]} \right) = \sum_{I|E \neq 0} C_I =: iW[J]. \quad (22.3.6)$$

This is a sum over connected diagrams with no sources. The factor of i is chosen so that $W[J]$ is real for a real scalar field.

That disconnected diagrams don't contribute to this value proves a similar statement about Gaussian correlators made in [Section 18.2.2](#).

Part VIII

Scalar Field Correlators

Twenty-Three

Correlators

23.1 Correlators

The generating functional in φ^3 theory is

$$Z[J] = \sum_{V=0}^{\infty} \frac{1}{V!} \left[\frac{ig}{3!} \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right)^3 \right]^V \sum_{P=0}^{\infty} \frac{1}{P!} \left[-\frac{1}{2} \int d^D y \int d^D z J(y) i\Delta(y-z) J(z) \right]^P.$$

We want to calculate the n -point correlator

$$G^{(n)}(x_1, \dots, x_n) := \langle 0 | T\{\varphi(x_1) \cdots \varphi(x_n)\} | 0 \rangle \quad (23.1.1)$$

$$= \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \cdots \left(\frac{1}{i} \frac{\delta}{\delta J(x_n)} \right) Z[J] \Big|_{J=0}. \quad (23.1.2)$$

Since $Z[J]$ is just a power series in g we can also expand $G^{(n)}$ as a power series in g with coefficients $G^{(n,V)}$ defined by

$$G^{(n)}(x_1, \dots, x_n) = \sum_{V=0}^{\infty} G^{(n,V)}(x_1, \dots, x_n) \quad (23.1.3)$$

In order to have a nonzero contribution we need only consider terms with no sources left after performing the derivatives. That is, we want $n = E$, where $E = 2P - 3V$ is the number of sources.

23.2 Two Point Correlator

As an example we'll compute the two point correlator,

$$G^{(2)}(x_1, x_2) = \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x_2)} \right) Z[J] \Big|_{J=0}. \quad (23.2.1)$$

to order of g^2 . Since $n = 2$, and we require $E = n$ for a nonzero contribution, we need to choose P to satisfy $2 = 2P - 3V$.

23.2.1 $V = 0$

For $V = 0$ we must have $P = 1$ for a nonzero contribution. We can easily read off the $V = 0$ and $P = 1$ contribution to the generating functional, it is

$$-\frac{1}{2} \int_{y,z} J(y) i\Delta(y-z) J(z) = \bullet \text{---} \bullet \quad (23.2.2)$$

We could have gone the other way here. If we had first drawn the diagram then we can see from this that there are two sources, the dots, connected by a propagator

So we want to compute

$$G^{(2,0)}(x_1, x_2) = \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x_2)} \right) Z[J] \Big|_{J=0} \quad (23.2.3)$$

$$= -\frac{1}{2} \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left(\frac{1}{i} \frac{\delta}{\delta J(x_2)} \right) \int_{y,z} J(y) i\Delta(y-z) J(z) \quad (23.2.4)$$

$$= -\frac{1}{2} \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left[\int_{y,z} \delta(x_1 - y) \Delta(y-z) J(z) \right. \quad (23.2.5)$$

$$\left. + \int_{y,z} J(y) i\Delta(y-z) \delta(x_1 - z) \right] \quad (23.2.6)$$

$$= -\frac{1}{2} \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left[\int_z \Delta(x_1 - z) J(z) \right. \quad (23.2.7)$$

$$\left. + \int_y J(y) \Delta(y - x_1) \right] \quad (23.2.8)$$

$$= -\left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \int_y J(y) \Delta(y - x_1) \quad (23.2.9)$$

where in the last step we perform a change of integration variables $z \rightarrow y$ in the first integral and then use the symmetry of Δ , $\Delta(a-b) = \Delta(b-a)$, which is a result of translation invariance. Performing the last integral we get

$$G^{(2,0)}(x_1, x_2) = i \int_y \delta(x_2 - y) \Delta(y - x_1) \quad (23.2.10)$$

$$= i\Delta(x_2 - x_1). \quad (23.2.11)$$

We can represent this diagrammatically as

$$G^{(2,0)}(x_1, x_2) = i\Delta(x_1 - x_2) = x_1 \text{ --- } x_2 \quad (23.2.12)$$

where we now fix the two endpoints and so no longer have the symmetry factor of 2 from swapping the ends of the propagator.

Recall that we can write the propagator as the inverse Fourier transform of its momentum space equivalent:

$$G^{(2,0)}(x_1, x_2) = i\Delta(x_1 - x_2) = \int \frac{d^D p}{(2\pi)^D} e^{ip \cdot (x_1 - x_2)} \frac{i}{p^2 - m^2 + i\varepsilon}. \quad (23.2.13)$$

23.2.2 $V = 1$

For $V = 1$ we require the equation $2 = 2P - 3$ to hold, but there is no integer solution for this, and hence there is no nonzero contribution from $V = 1$. This is a peculiarity of φ^3 theory, not a general rule. We saw this through explicit calculations of the $V = 1$ terms in [Chapter 10](#).

23.2.3 $V = 2$

For a nonzero contribution with $V = 2$ we require the equation $2 = 2P - 6$ to hold, and it does if $P = 4$. We could go through the exercise of writing out the

$V = 2$ and $P = 4$ contribution to $Z[J]$ and then computing the functional derivative with respect to $J(x_1)$ and $J(x_2)$, but this isn't that fun. We'd have six derivatives to compute, all of which can act on any of eight factors of J , it's a mess. Instead we can realise that a $V = 2$ $P = 4$ contribution consists of a diagram with two vertices and four propagators. Before computing the derivative we must have two sources, so that our final result has no sources. So, we can simply write down all diagrams with two vertices, four propagators, and two sources. It turns out that there are two distinct diagrams satisfying these requirements.

23.2.3.1 The First Diagram

¹Note that the circle is actually two propagators, one making up the top and the other the bottom.

The following diagram has two vertices, four propagators¹, and two sources:



$$(23.2.14)$$

Note that the labels are just to assist in constructing the integral.

Let's start with the symmetry factor. We have two planes of mirror symmetry, vertical and horizontal, and no other symmetries. So the symmetry factor is $2 \cdot 2 = 4$. We then have two source factors, at z_1 and z_2 , propagators from z_1 to w_1 and w_2 to z_2 , and then two propagators from w_1 to w_2 . The integral corresponding to this diagram is

$$I = \frac{1}{4} \int_{z_1, z_2, w_1, w_2} J(z_1) i\Delta(z_1 - w_1) [i\Delta(w_1 - w_2)]^2 i\Delta(w_2 - z_2) J(z_2). \quad (23.2.15)$$

Notice that there are four factors of i giving an overall factor of 1.

We can compute the first derivative, splitting into two terms with the product rule, then getting Dirac deltas, which we use to remove one integral, and then performing a change of variables to show both terms are the same, so we can recombine them and pick up a factor of 2 giving the result

$$\frac{1}{i} \frac{\delta I}{\delta J(x_2)} = \frac{1}{2i} \int_{z_1, w_1, w_2} J(z_1) \Delta(z_1 - w_1) \Delta(w_1 - w_2)^2 \Delta(w_2 - x_2). \quad (23.2.16)$$

We can perform the next derivative, now with only one factor of J to act on, and we again get a Dirac delta which we use to remove one of the integrals for the final result

$$G_a^{(2,2)}(x_1, x_2) = \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left(\frac{1}{i} \frac{\delta I}{\delta J(x_2)} \right) \quad (23.2.17)$$

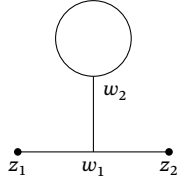
$$= -\frac{1}{2} \int_{w_1, w_2} \Delta(x_1 - w_1) \Delta(w_1 - w_2)^2 \Delta(w_2 - x_2). \quad (23.2.18)$$

The subscript a here simply labels that this is the contribution from the first diagram of this type.

23.2.3.2 The Second Diagram

²This time the circle is just a single propagator connecting the vertex to itself.

The other diagram with two vertices, four propagators², and two sources is



(23.2.19)

Again, the labels are just to help us construct the integral.

Starting with the symmetry factor we have a vertical plane of mirror symmetry, and we can also take the top propagator, forming a circle, swap the ends, and the diagram doesn't change. So the symmetry factor is $2 \cdot 2 = 4$. We then have two source factors at z_1 and z_2 , propagators from z_1 to w_1 , w_1 to w_2 , and w_1 to z_2 . We also have a propagator from w_2 back to w_2 , which will give a factor of $\Delta(0)$. The integral corresponding to this diagram is

$$I = \frac{1}{4} \int_{z_1, z_2, w_1, w_2} J(z_1) i\Delta(z_1 - w_1) i\Delta(w_1 - w_2) i\Delta(0) i\Delta(w_2 - z_2) J(z_2). \quad (23.2.20)$$

Notice that again the factors of i cancel.

We can compute the first derivative, much like for the previous diagram, and the result is

$$\frac{1}{i} \frac{\delta I}{\delta J(x_2)} = \frac{1}{2i} \int_{z_1, w_1, w_2} J(z_1) \Delta(z_1 - w_1) \Delta(w_1 - w_2) \Delta(0) \Delta(w_2 - x_2). \quad (23.2.21)$$

Computing the second derivative we have

$$G_b^{(2,2)}(x_1, x_2) = \left(\frac{1}{i} \frac{\delta}{\delta J(x_1)} \right) \left(\frac{1}{i} \frac{\delta I}{\delta J(x_2)} \right) \quad (23.2.22)$$

$$= -\frac{1}{2} \int_{w_1, w_2} \Delta(x_1 - w_1) \Delta(w_1 - w_2) \Delta(0) \Delta(w_2 - x_2). \quad (23.2.23)$$

Again, the subscript b simply labels that this is the contribution from the second diagram of this type.

23.3 Momentum Space

Its often easier to do computations in Fourier space. One example of this is the correlator, which is given by

$$\Delta(x) = \int \frac{d^D p}{(2\pi)^D} \frac{e^{-ip \cdot x}}{p^2 - m^2 + i\epsilon} \quad (23.3.1)$$

in real space, but in Momentum space is simply

$$\tilde{\Delta}(p) = \frac{1}{p^2 - m^2 + i\epsilon}. \quad (23.3.2)$$

We can Fourier transform the correlators we have already calculated to work out the correlator in momentum space, but first, consider the Fourier transform of the two-point correlator, $G^{(n)}$, in a general theory. This is given by

$$\tilde{G}^{(2)}(p_1, p_2) = \int d^D x_1 d^D x_2 e^{ip_1 \cdot x_1} e^{ip_2 \cdot x_2} G^{(2)}(x_1, x_2). \quad (23.3.3)$$

Now perform a change of variables for the x_1 integral, defining $\xi = x_1 - x_2$. The measure and limits are invariant under translations, as is $G^{(2)}(x_1, x_2)$. This means we can rewrite $G^{(2)}$ as a function of $x_1 - x_2 = \xi$, so $G^{(2)}(x_1 - x_2) = G^{(2)}(\xi)$. Doing so we have

$$\tilde{G}^{(2)}(p_1, p_2) = \int d^D \xi d^D x_2 e^{ip_1 \cdot (\xi + x_2)} e^{ip_2 \cdot x_2} G^{(2)}(\xi) \quad (23.3.4)$$

$$= \int d^D \xi d^D x_2 e^{ip_1 \cdot \xi} e^{i(p_1 + p_2) \cdot x_2} G^{(2)}(\xi) \quad (23.3.5)$$

$$= \int d^D \xi e^{ip_1 \cdot \xi} (2\pi)^D \delta(p_1 + p_2) G^{(2)}(\xi) \quad (23.3.6)$$

$$= (2\pi)^D \delta(p_1 + p_2) \tilde{G}^{(2)}(p_1), \quad (23.3.7)$$

where in the last step we recognise $\tilde{G}^{(2)}(p_1)$ as the Fourier transform of the single variable function depending on $x_1 - x_2 = \xi$.

What we see here is that the 2-point correlator depends on only 1 momentum value, the other being fixed by the Dirac delta, so that all momenta sum to zero. This is a general fact. In momentum space the n -point correlator will depend on $n - 1$ momenta values.

23.3.1 $V = 0$

We already know that $G^{(2)}(x_1, x_2) = i\Delta(x_1 - x_2)$, so we trivially have

$$\tilde{G}^{(2,0)}(p) = \frac{i}{p^2 - m^2 + i\epsilon}. \quad (23.3.8)$$

We represent this diagrammatically as

$$\begin{array}{c} p \\ \longrightarrow \\ \hline \end{array} \quad (23.3.9)$$

We use arrows to denote the direction of the momentum.

23.3.2 $V = 2$

23.3.2.1 The First Diagram

We can similarly compute the contribution from the first diagram in momentum space by computing the Fourier transform of the known result:

$$G_a^{(2,0)}(x_1, x_2) = -\frac{1}{2} \int_{w_1, w_2} \Delta(x_1 - w_1) \Delta(w_1 - w_2)^2 \Delta(w_2 - x_2). \quad (23.3.10)$$

We start by writing the propagators as inverse Fourier transforms of the momentum space propagator, so

$$\begin{aligned} G_a^{(2,0)}(x_1, x_2) = & -\frac{1}{2} \int_{w_1, w_2} \int_{k_1, k_2, k_3, k_4} \frac{e^{-ik_1 \cdot (x_1 - w_1)}}{k_1^2 - m^2 + i\epsilon} \frac{e^{-ik_2 \cdot (w_1 - w_2)}}{k_2^2 - m^2 + i\epsilon} \\ & \times \frac{e^{-ik_3 \cdot (w_1 - w_2)}}{k_3^2 - m^2 + i\epsilon} \frac{e^{-ik_4 \cdot (w_2 - x_2)}}{k_4^2 - m^2 + i\epsilon}. \end{aligned} \quad (23.3.11)$$

Now comes the actual Fourier transform, we have to Fourier transform both x_1 and x_2 , giving

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} \int_{x_1, x_2} \int_{w_1, w_2} \int_{k_1, k_2, k_3, k_4} e^{-ip_1 \cdot x_1} e^{-ip_2 \cdot x_2} \frac{e^{-ik_1 \cdot (x_1 - w_1)}}{k_1^2 - m^2 + i\varepsilon} \\ & \times \frac{e^{-ik_2 \cdot (w_1 - w_2)}}{k_2^2 - m^2 + i\varepsilon} \frac{e^{-ik_3 \cdot (w_1 - w_2)}}{k_3^2 - m^2 + i\varepsilon} \frac{e^{-ik_4 \cdot (w_2 - x_2)}}{k_4^2 - m^2 + i\varepsilon} \end{aligned} \quad (23.3.12)$$

where the Fourier transform is

$$\tilde{f}(p) = \int_x e^{-ip \cdot x} f(x). \quad (23.3.13)$$

Now we can rewrite the exponentials collecting x_1 and x_2 together:

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} \int_{x_1, x_2} \int_{w_1, w_2} \int_{k_1, k_2, k_3, k_4} e^{-i(p_1 + k_1) \cdot x_1} e^{-i(p_2 - k_4) \cdot x_2} \\ & \times \frac{e^{ik_1 \cdot w_1}}{k_1^2 - m^2 + i\varepsilon} \frac{e^{-ik_2 \cdot (w_1 - w_2)}}{k_2^2 - m^2 + i\varepsilon} \frac{e^{-ik_3 \cdot (w_1 - w_2)}}{k_3^2 - m^2 + i\varepsilon} \frac{e^{-ik_4 \cdot w_2}}{k_4^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.14)$$

We can now perform the integrals over x_1 and x_2 , and we'll get Dirac deltas from these first two exponentials:

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} \int_{w_1, w_2} \int_{k_1, k_2, k_3, k_4} (2\pi)^D \delta(p_1 + k_1) (2\pi)^D \delta(p_2 - k_4) \\ & \times \frac{e^{ik_1 \cdot w_1}}{k_1^2 - m^2 + i\varepsilon} \frac{e^{-ik_2 \cdot (w_1 - w_2)}}{k_2^2 - m^2 + i\varepsilon} \frac{e^{-ik_3 \cdot (w_1 - w_2)}}{k_3^2 - m^2 + i\varepsilon} \frac{e^{-ik_4 \cdot w_2}}{k_4^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.15)$$

The Dirac deltas allow us to easily perform two of the k integrals, setting $k_1 = -p_1$ and $k_4 = p_2$. Note that the k integrals, being momentum integrals, come with implicit factors of $1/(2\pi)^D$, and so the k_1 and k_4 integrals cancel the factor of $(2\pi)^{2D}$:

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} \int_{w_1, w_2} \int_{k_2, k_3} \frac{e^{-ip_2 \cdot w_1}}{p_2^2 - m^2 + i\varepsilon} \frac{e^{-ik_2 \cdot (w_1 - w_2)}}{k_2^2 - m^2 + i\varepsilon} \\ & \times \frac{e^{-ik_3 \cdot (w_1 - w_2)}}{k_3^2 - m^2 + i\varepsilon} \frac{e^{-ip_1 \cdot w_2}}{p_1^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.16)$$

Next rewrite the exponentials again, this time collecting factors of w_1 and w_2 . We get

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} \int_{w_1, w_2} \int_{k_2, k_3} e^{-i(p_2 + k_2 + k_3) \cdot w_1} e^{-i(-k_2 - k_3 + p_1) \cdot w_2} \\ & \times \frac{1}{p_2^2 - m^2 + i\varepsilon} \frac{1}{k_2^2 - m^2 + i\varepsilon} \frac{1}{k_3^2 - m^2 + i\varepsilon} \frac{1}{p_1^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.17)$$

We can then perform the w_1 and w_2 integrals, and we get another two Dirac deltas:

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} \int_{k_2, k_3} (2\pi)^D \delta(p_2 + k_2 + k_3) (2\pi)^D \delta(-k_2 - k_3 + p_1) \\ & \times \frac{1}{p_2^2 - m^2 + i\varepsilon} \frac{1}{k_2^2 - m^2 + i\varepsilon} \frac{1}{k_3^2 - m^2 + i\varepsilon} \frac{1}{p_1^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.18)$$

We can then perform the k_2 integral using the first Dirac delta setting $k_2 = -p_2 - k_3$, and again cancelling one of the $(2\pi)^D$ factors with one implicit in the k_2 integral, giving

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} \int_{k_3} (2\pi)^D \delta(p_2 + k_3 - k_3 + p_1) \frac{1}{p_2^2 - m^2 + i\varepsilon} \\ & \times \frac{1}{(p_2 + k_3)^2 - m^2 + i\varepsilon} \frac{1}{k_3^2 - m^2 + i\varepsilon} \frac{1}{p_1^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.19)$$

Note that we use the square to get rid of the negatives in the second denominator. We can rename $k_3 \rightarrow k$ and rewrite this as

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p_1, p_2) = & -\frac{1}{2} (2\pi)^D \delta(p_1 + p_2) \frac{1}{p_2^2 - m^2 + i\varepsilon} \\ & \times \left\{ \int_k \frac{1}{(p_2 + k)^2 - m^2 + i\varepsilon} \frac{1}{k^2 - m^2 + i\varepsilon} \right\} \frac{1}{p_1^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.20)$$

Notice that the Dirac delta means we can set $p_2 = -p_1$ and then we have $(-p_1 + k)^2 = (p_1 - k)^2$ in the first denominator in the integral. Then the result can be written as a function of $p = p_1$ alone, with the only p_2 term being in the Dirac delta:

$$\begin{aligned} \tilde{G}_a^{(2,0)}(p) = & -\frac{1}{2} (2\pi)^D \delta(p + p_2) \frac{1}{p^2 - m^2 + i\varepsilon} \\ & \times \left\{ \int \frac{d^D k}{(2\pi)^D} \frac{1}{(p - k)^2 - m^2 + i\varepsilon} \frac{1}{k^2 - m^2 + i\varepsilon} \right\} \frac{1}{p^2 - m^2 + i\varepsilon}. \end{aligned} \quad (23.3.21)$$

We can also express this in terms of a momentum space diagram:



$$(23.3.22)$$

Notice that at each vertex the total momentum is conserved, that is the sum of the momenta on the arrows pointing towards the vertex is equal to the sum of the momenta on the arrows pointing out of the vertex.

Appendices

A

Pre-Course Revision

The content of this appendix is based on the notes released before the course began to recap key ideas from previous courses. In particular, it covers material from Quantum Theory and Classical Electrodynamics.

A.1 Special Relativity

A.1.1 Four Vectors

Consider some event occurring at the spatial position \mathbf{x} and at time t . The coordinates of this event, in four-dimensional Minkowski space, form a **contravariant** four-vector, which we write with upper indices:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, \mathbf{x}). \quad (\text{A.1.1})$$

The **covariant** four-vector has lower indices:

$$x_\mu = (x_0, x_1, x_2, x_3) = (ct, -\mathbf{x}). \quad (\text{A.1.2})$$

Similarly a general four-vector has a contravariant and covariant form:

$$a^\mu = (a^0, a^1, a^2, a^3) = (a^0, \mathbf{a}), \quad \text{and} \quad a_\mu = (a_0, a_1, a_2, a_3) = (a^0, -\mathbf{a}). \quad (\text{A.1.3})$$

In particular, $a^0 = a_0$ and $a^i = -a_i$, where we follow the convention that Greek indices run from 0 to 3 and Latin indices run from 1 to 3.

We move between upper and lower indices using the **metric tensor**, η , also denoted g :

$$a^\mu = \eta^{\mu\nu} a_\nu, \quad \text{and} \quad a_\mu = \eta_{\mu\nu} a^\nu, \quad (\text{A.1.4})$$

where we apply the summation convention, summing over repeated indices with one up and one down. The covariant and contravariant components of the Minkowski metric are given by

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (\text{A.1.5})$$

and the mixed components are given by

$$\eta_{\mu}^{\nu} = \eta^{\nu}_{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \delta_{\mu}^{\nu} = \delta^{\nu}_{\mu}. \quad (\text{A.1.6})$$

Here δ^{μ}_{ν} is the Kronecker delta, defined to be 1 if $\mu = \nu$ and 0 otherwise. Note that we have made a choice of signs here.

! Sometimes the Minkowski metric is instead chosen to be

$$\eta^{\mu\nu} \stackrel{!}{=} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (\text{A.1.7})$$

These two conventions are known as $(+---)$ and $(-+++)$ for short, where the signs refer to the signs along the diagonal.

The **scalar product** of two four-vectors, a and b , in Minkowski space is defined to be

$$a \cdot b := a^{\mu} b_{\mu} = a_{\mu} b^{\mu} = a_{\mu} b_{\nu} \eta^{\mu\nu} = a^{\mu} b^{\nu} \eta_{\mu\nu} = a^0 b^0 - \mathbf{a} \cdot \mathbf{b} \quad (\text{A.1.8})$$

where $\mathbf{a} \cdot \mathbf{b} = a^1 b^1 + a^2 b^2 + a^3 b^3$ is the normal three-dimensional scalar product.

A.1.2 Lorentz Transformations

Lorentz transformations are linear transformations on four-vectors which leave the scalar product invariant. If Λ is a Lorentz transformation then the four-vector a transforms as

$$a^{\mu} \rightarrow a'^{\mu} = \Lambda^{\mu}_{\nu} a^{\nu}. \quad (\text{A.1.9})$$

These are the **homogeneous Lorentz transformations**, and they form a group called the **homogeneous Lorentz group**, or just the Lorentz group, denoted $O(1, 3)$, where $(1, 3)$ is the signature of the metric, one time component and three spatial components. Including translations as well we get the Poincaré group, $\mathbb{R}^{1,3} \rtimes O(1, 3)$, where $\mathbb{R}^{1,3}$ denotes translations in Minkowski space¹.

The “standard” Lorentz transformation is a **boost** along the x -axis, where the boosted frame moves at a constant speed, v , along the x -axis:

$$\Lambda^{\mu}_{\nu} = \begin{pmatrix} \cosh \omega & -\sinh \omega & 0 & 0 \\ -\sinh \omega & \cosh \omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{A.1.10})$$

where

$$\tanh \omega = \beta := \frac{v}{c}, \quad (\text{A.1.11})$$

$$\cosh \omega = \gamma := \frac{1}{\sqrt{1 - \beta^2}} = \frac{1}{\sqrt{1 - v^2/c^2}}, \quad (\text{A.1.12})$$

$$\sinh \omega = \gamma \beta. \quad (\text{A.1.13})$$

¹for more details on groups see Symmetries of Quantum Mechanics or Symmetries of Particles and Fields

In terms of coordinates we have

$$ct' = \gamma \left(ct - \frac{v}{c} x \right), \quad (\text{A.1.14})$$

$$x' = \gamma(x - vt), \quad (\text{A.1.15})$$

$$y' = y, \quad (\text{A.1.16})$$

$$z' = z. \quad (\text{A.1.17})$$

Note that the Lorentz group contains ordinary three-dimensional rotations as a subgroup, $O(3)$, for example, the following Lorentz transformation is a clockwise rotation by φ about the z -axis, leaving time unaffected:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi & 0 \\ 0 & -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (\text{A.1.18})$$

By definition the scalar product is invariant under the action of the Lorentz group. This means that for some Lorentz transformation Λ and four-vectors a and b we have

$$b'^\mu b'^\nu \eta_{\mu\nu} = \Lambda^\mu{}_\rho a^\rho \Lambda^\nu{}_\sigma b^\sigma \eta_{\mu\nu} = a^\rho b^\sigma \eta_{\rho\sigma}. \quad (\text{A.1.19})$$

The last equality is given by computing the scalar product after the transformation. Since a and b are arbitrary this means we must have

$$\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta_{\mu\nu} = \eta_{\rho\sigma}. \quad (\text{A.1.20})$$

We can rewrite this as

$$(\Lambda^\top)_\rho{}^\mu \eta_{\mu\nu} \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}. \quad (\text{A.1.21})$$

In matrix notation we then have $\Lambda^\top \eta \Lambda = \eta$. Taking the determinant we have

$$\det(\Lambda^\top \eta \Lambda) = \det(\Lambda^\top) \det(\eta) \det(\Lambda) = \det(\eta) \implies (\det \Lambda)^2 = 1, \quad (\text{A.1.22})$$

where we've used $\det(AB) = \det(A) \det(B)$, and $\det(A^\top) = \det(A)$. Hence, we have

$$\det \Lambda = \pm 1 \quad (\text{A.1.23})$$

for all $\Lambda \in O(1, 3)$. Transformations with $\det \Lambda = +1$ are called **proper Lorentz transformations**, and form a subgroup, $SO(1, 3)$. Lorentz transformations with $\det \Lambda = -1$ are called **improper Lorentz transformations**.

Now set the free indices, ρ and σ , to 0 in Equation (A.1.20) and we get

$$\Lambda^\mu{}_0 \Lambda^\nu{}_0 \eta_{\mu\nu} = \eta_{00} = 1. \quad (\text{A.1.24})$$

Writing out the components this becomes

$$\Lambda^0{}_0 \Lambda^0{}_0 - \Lambda^1{}_0 \Lambda^1{}_0 - \Lambda^2{}_0 \Lambda^2{}_0 - \Lambda^3{}_0 \Lambda^3{}_0 = 1 \quad (\text{A.1.25})$$

which gives

$$(\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^3 (\Lambda^i{}_0)^2 \geq 1, \quad (\text{A.1.26})$$

where the inequality follows from the fact that Λ^i_0 is real, so its square is nonnegative. We can conclude from this that either

$$\Lambda^0_0 \geq 1 \quad \text{or} \quad \Lambda^0_0 \leq -1. \quad (\text{A.1.27})$$

We call transformations with $\Lambda^0_0 \geq 1$ **orthochronous Lorentz transformations**, these form a subgroup, $O^+(1, 3)$. There is also a subgroup of proper orthochronous Lorentz transformations, $SO^+(1, 3)$.

We can divide homogeneous Lorentz transformations into four classes, or in more technical language, the Lie group $O(1, 3)$ has four connected components:

$$\begin{array}{ccc} \det \Lambda = +1 & \det \Lambda = -1 & \\ \Lambda^0_0 \geq +1 & \text{I} & \text{II} \\ \Lambda^0_0 \leq -1 & \text{III} & \text{IV} \end{array} \quad (\text{A.1.28})$$

It is component I, the proper orthochronous Lorentz transformations, which contains the identity and so is the most important. In particular, it is the Lie group generated by the Lie algebra $\mathfrak{so}(1, 3)$.

Any transformation of type II, III, or IV can be generated by combining a proper orthochronous transformation of type I with one of the following:

- **Spatial inversion:** The parity operation $x^0 \rightarrow x^0$ and $\mathbf{x} \rightarrow -\mathbf{x}$:

$$\Lambda^\mu_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (\text{A.1.29})$$

- **Time reversal:** the operation $x^0 \rightarrow -x^0$ and $\mathbf{x} \rightarrow \mathbf{x}$:

$$\Lambda^\mu_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (\text{A.1.30})$$

- **Spacetime inversion:** The operation $x^\mu \rightarrow -x^\mu$:

$$\Lambda^\mu_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (\text{A.1.31})$$

If x and y are four vectors then the **spacetime interval** $s^2 := (x - y)^\mu (x - y)_\mu$ is invariant under Lorentz transformations, since its just a scalar product of some four-vector $x - y$.

A.1.3 Classification of Four-Vectors

We classify four-vectors, a , by whether their scalar product with themselves is positive, negative, or zero:

- if $a^2 > 0$ we call a **time-like**,
- if $a^2 = 0$ we call a **light-like**,
- if $a^2 < 0$ we call a **space-like**.

A.1.4 Differential Operators

The gradient operator in normal \mathbb{R}^n consists of spatial derivatives at each component. The generalisation to Minkowski space includes a time derivative, and a factor of c for dimensional consistency:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) = \left(\frac{1}{c} \frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right). \quad (\text{A.1.32})$$

Note that the derivative with respect to a contravariant vector gives a covariant operator. As with normal four-vectors we can raise the index to get the operator

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right) = \left(\frac{1}{c} \frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z} \right). \quad (\text{A.1.33})$$

The Laplacian, $\nabla^2 = \nabla \cdot \nabla$, can also be generalised to Minkowski space as the scalar product of ∂ with itself, we call the result the **d'Alembertian**:

$$\partial^2 := \partial_\mu \partial^\mu = \partial^\mu \partial_\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}. \quad (\text{A.1.34})$$

This is also denoted \square^2 or \square . Note that the wave equation for some field, φ , can then be written as $\partial^2 \varphi = 0$.

A.1.5 Momentum and Energy

The four-momentum, p , is a four-vector given by

$$p^\mu = \left(\frac{E}{c}, \mathbf{p} \right) \quad (\text{A.1.35})$$

where E is the energy of the particle and \mathbf{p} is the relativistic three-momentum of the particle.

For a free particle of mass m we have that

$$p^2 = \frac{E^2}{c^2} - \mathbf{p} \cdot \mathbf{p} = m^2 c^2. \quad (\text{A.1.36})$$

This can be rewritten as

$$E^2 = |\mathbf{p}|^2 c^2 + m^2 c^4, \quad (\text{A.1.37})$$

which is known as the **energy-momentum relation**, or the **mass-shell condition**.

Particles for which $p^\mu p_\mu = m^2 c^2$ are said to be on their **mass-shell**.

A.2 Electromagnetic Theory

A.2.1 Maxwell's Equations

We use **Heaviside-Lorentz units**, where $\epsilon_0 = 1$. In these units **Maxwell's equations** are

$$\nabla \cdot \mathbf{B} = 0, \quad (\text{ME1})$$

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad (\text{ME2})$$

$$\nabla \times \mathbf{B} = \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}, \quad (\text{ME3})$$

$$\nabla \cdot \mathbf{E} = \rho. \quad (\text{ME4})$$

Here \mathbf{E} is the electric field, \mathbf{B} is the magnetic field, \mathbf{j} is the current density, and ρ is the charge density.

Any vector field with vanishing divergence, including \mathbf{B} , can be written as the curl of another field, say \mathbf{A} , then $\nabla \cdot \mathbf{B} = 0$ follows from the identity that $\nabla \cdot (\nabla \times \mathbf{A}) = 0$ for any vector field \mathbf{A} . Hence, we can rewrite the first equation as

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (\text{ME1}')$$

Similarly we can write \mathbf{E} as

$$\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \nabla \Phi \quad (\text{ME2}')$$

since $\nabla \times (\nabla \Phi) = \mathbf{0}$ for all scalar fields Φ and the curl of the first term just gets us back to \mathbf{B} , so this equation can replace the second of Maxwell's equations.

Combining the third and fourth equation it is possible to derive the continuity equation

$$\nabla \cdot \mathbf{j} + \frac{1}{c} \frac{\partial \rho}{\partial t} = 0. \quad (\text{A.2.1})$$

Defining the electromagnetic **four-current**, $j_{\text{em}}^\mu := (\rho, \mathbf{j})$ then we can write this as

$$\partial_\mu j_{\text{em}}^\mu = 0. \quad (\text{A.2.2})$$

This is a statement of charge conservation.

A.2.2 Four-Vector Potential

We can define the **four-potential** as

$$A^\mu := (\Phi, \mathbf{A}). \quad (\text{A.2.3})$$

Recall the identity

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}. \quad (\text{A.2.4})$$

Substitute the modified Maxwell equations into the third Maxwell equation, giving

$$\nabla \times (\nabla \times \mathbf{A}) = \mathbf{j} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \frac{1}{c} \nabla \frac{\partial \Phi}{\partial t}. \quad (\text{A.2.5})$$

On the left we can use the identity to get

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \mathbf{j} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \frac{1}{c} \nabla \frac{\partial \Phi}{\partial t}. \quad (\text{A.2.6})$$

Rearranging this gives

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} + \nabla \left[\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} \right] = \partial^2 \mathbf{A} + \nabla \left[\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} \right] = \mathbf{j}. \quad (\text{A.2.7})$$

Doing something similar with the fourth Maxwell equation we get

$$-\frac{1}{c} \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) - \nabla^2 \Phi = \rho. \quad (\text{A.2.8})$$

We can combine these two equations into a covariant form:

$$\partial^2 A^\mu - \partial^\mu \partial_\nu A^\nu = j_{\text{em}}^\mu. \quad (\text{A.2.9})$$

A.2.3 Gauge Transformations

We can define two new potentials, $\tilde{\mathbf{A}}$ and $\tilde{\Phi}$, via the **gauge transformation**

$$\tilde{\mathbf{A}} = \mathbf{A} - \nabla\chi \quad \text{and} \quad \tilde{\Phi} = \Phi + \frac{1}{c} \frac{\partial\chi}{\partial t}, \quad (\text{A.2.10})$$

or in covariant form,

$$\tilde{A}^\mu = A^\mu + \partial^\mu\chi. \quad (\text{A.2.11})$$

Here χ is any scalar field (with second derivatives). Importantly the fields \mathbf{E} and \mathbf{B} don't change if we make this transformation. We can use this to choose potentials satisfying the **Lorenz gauge condition**

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial\Phi}{\partial t} = 0, \quad (\text{A.2.12})$$

or in covariant form,

$$\partial_\nu A^\nu = 0. \quad (\text{A.2.13})$$

In this case Maxwell's equations reduce to two decoupled equations,

$$\partial^2 \mathbf{A} = \mathbf{j}, \quad \text{and} \quad \partial^2 \Phi = \rho, \quad (\text{A.2.14})$$

or in covariant form,

$$\partial^2 A^\mu = j_{\text{em}}^\mu. \quad (\text{A.2.15})$$

Note that the Lorenz gauge doesn't fix the potentials uniquely, since

$$\nabla \cdot \tilde{\mathbf{A}} + \frac{1}{c} \frac{\partial\tilde{\Phi}}{\partial t} = \nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial\Phi}{\partial t} - \nabla^2\chi + \frac{1}{c^2} \frac{\partial^2\chi}{\partial t^2}. \quad (\text{A.2.16})$$

So if the original potential, A^μ , satisfies $\partial_\mu A^\mu = 0$ then we can choose to have $\partial_\mu \tilde{A}^\mu = 0$ so long as χ is such that $\partial^2\chi = 0$.

A.2.4 Covariant Form of Maxwell's Equations

The **Maxwell tensor**, or **electromagnetic field strength tensor** is defined to be

$$F^{\mu\nu} := \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (\text{A.2.17})$$



Note that there is a different sign convention used, for example, by Mandl and Shaw, where

$$F^{\mu\nu} \stackrel{!}{=} \partial^\nu A^\mu - \partial^\mu A^\nu. \quad (\text{A.2.18})$$

We can think of the electromagnetic field strength tensor as the four-dimensional “curl” of the four-potential²

The electromagnetic field strength tensor is manifestly gauge invariant, since any gauge transformation cancels between the two terms.

²this becomes precise in the language of geometric algebra, where we replace curl with the wedge product, and the wedge product of two vectors gives a bivector, which for our purposes is a two index tensor.

Considering the space-time and space-space components of F we can fairly simply show that they correspond to the normal electric and magnetic fields:

$$F^{i0} = \partial^i A^0 - \partial^0 A^i = \frac{\partial \Phi}{\partial x^i} - \frac{1}{c} \frac{\partial A^i}{\partial t} = E^i, \quad (\text{A.2.19})$$

$$F^{ij} = \partial^i A^j - \partial^j A^i = -\frac{\partial A^i}{\partial x^j} + \frac{\partial A^j}{\partial x^i} = -\epsilon^{ijk} B^k. \quad (\text{A.2.20})$$

Here ϵ^{ijk} is the Levi-Civita symbol, which is defined to be 1 if ijk is an even permutation of 123, -1 if it's an odd permutation, and 0 otherwise.

We can write F as a 4×4 matrix:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \quad (\text{A.2.21})$$

We can then write Equation (A.2.9) as

$$\partial_\mu F^{\mu\nu} = j_{\text{em}}^\nu. \quad (\text{A.2.22})$$

Conservation of the electromagnetic four-current is then very obvious, since

$$\partial_\nu j_{\text{em}}^\nu = \partial_\nu \partial_\mu F^{\mu\nu} = 0, \quad (\text{A.2.23})$$

since $F^{\mu\nu}$ is antisymmetric in μ and ν and $\partial_\nu \partial_\mu$ is symmetric in μ and ν .

A.2.5 Interaction of Matter with EM Fields

The classical Hamiltonian for a free, non-relativistic particle is

$$H = \frac{|\mathbf{p}|^2}{2m}. \quad (\text{A.2.24})$$

For a particle with charge q in an electromagnetic field we modify the Hamiltonian through the **minimal coupling prescription**,

$$H \rightarrow H - q\Phi, \quad \text{and} \quad \mathbf{p} \rightarrow \mathbf{p} - \frac{q}{c} \mathbf{A}, \quad (\text{A.2.25})$$

giving the classical Hamiltonian for a non-relativistic particle in an electromagnetic field:

$$H = \frac{|\mathbf{p} - q\mathbf{A}/c|^2}{2m} + q\Phi. \quad (\text{A.2.26})$$

Note that this equation is in Heaviside–Lorentz units, in SI units there is no factor of c .

This Hamiltonian is chosen such that the force on a particle due to the electromagnetic field is

$$\mathbf{F} = q \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right), \quad (\text{A.2.27})$$

which is the Lorentz force in Heaviside–Lorentz units, again the factor of c is absent in SI units.

A.3 Relativistic Quantum Mechanics

A.3.1 Klein–Gordon Equation

The **Schrödinger equation** for a free particle of mass m is

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t). \quad (\text{A.3.1})$$

One way to obtain this is from the non-relativistic energy-momentum relation,

$$E = \frac{|\mathbf{p}|^2}{2m} = H, \quad (\text{A.3.2})$$

by substituting in the relevant operators:

$$E \rightarrow i\hbar\frac{\partial}{\partial t}, \quad \text{and} \quad \mathbf{p} \rightarrow -i\hbar\nabla. \quad (\text{A.3.3})$$

To obtain a relativistic equation then we should use the relativistic energy-momentum relation,

$$E^2 = |\mathbf{p}|^2c^2 + m^2c^4. \quad (\text{A.3.4})$$

Then after the operator substitution we have

$$-\hbar^2\frac{\partial^2}{\partial t^2}\varphi(\mathbf{r},t) = -\hbar^2c^2\nabla^2\varphi(\mathbf{r},t) + m^2c^4\varphi(\mathbf{r},t). \quad (\text{A.3.5})$$

This is the **Klein–Gordon equation**.

We can write the Klein–Gordon equation in a manifestly covariant form as

$$\left(\partial^2 + \frac{m^2c^2}{\hbar^2}\right)\varphi(x) = 0 \quad \text{or} \quad (\partial^2 + \mu^2)\varphi(x) = 0 \quad (\text{A.3.6})$$

where $\mu = mc/\hbar$. We can jump straight to this covariant form by instead using the operator prescription

$$p^\mu \rightarrow i\hbar\frac{\partial}{\partial x^\mu} = i\hbar\left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right). \quad (\text{A.3.7})$$

Note that if $m = 0$ then the Klein–Gordon equation reduces to a wave equation, $\partial^2\varphi(x) = 0$.

The Klein–Gordon equation is valid for any spin 0 particle.

A.3.1.1 Free Particle Solutions

The Klein–Gordon equation has plane-wave solutions,

$$\varphi(\mathbf{r},t) = \exp[i\mathbf{k} \cdot \mathbf{r} - i\omega t]. \quad (\text{A.3.8})$$

Substituting this into the Klein–Gordon equation we see it is a solution as long as ω , \mathbf{k} , and m are related by

$$\hbar^2\omega^2 = \hbar^2c^2|\mathbf{k}|^2 + m^2c^4. \quad (\text{A.3.9})$$

Taking the square root we have

$$\hbar\omega = \pm\sqrt{\hbar^2c^2|\mathbf{k}|^2 + m^2c^4}. \quad (\text{A.3.10})$$

These solutions are eigenfunctions of the momentum operator and the energy operator, with eigenvalues $\mathbf{p} = \hbar\mathbf{k}$ and $E = \hbar\omega$ respectively. Interpreting $\hbar\omega$ as the allowed energy of the free particle there is ambiguity in the sign of the total energy. We have both positive and negative energy solutions. Initially this was viewed as a problem, but later people realised that both sorts of solutions were important, with negative energy solutions corresponding to antiparticles.

Define the four-vector $k^\mu = (\omega/c, \mathbf{k})$. Then the solutions to the Klein–Gordon equation can be written in a covariant form as

$$\varphi(x) = \exp[-ik \cdot x] = \exp[-ik^\mu x_\mu] = \exp[-ip^\mu x_\mu/\hbar]. \quad (\text{A.3.11})$$

This allows us to interpret the four-momentum as $p^\mu = \hbar k^\mu$.

If φ is a solution of the Klein–Gordon equation then so is its complex conjugate, φ^* . The general solution is a superposition of plane-wave solutions:

$$\varphi(x) = \int \frac{d^3x}{2E_p(2\pi)^3} [f(p) \exp[-ip \cdot x/\hbar] + g(p) \exp[ip \cdot x/\hbar]]. \quad (\text{A.3.12})$$

Here the factors in the denominator of the measure are included for future convenience, and

$$E_p := +\sqrt{|\mathbf{p}|^2c^2 + m^2c^4} \quad (\text{A.3.13})$$

is defined to be positive.

A.3.2 Continuity Equation and Probability Interpretation

Denote the Schrödinger equation by (SE) and its conjugate by (SE)*, that is

$$(\text{SE}) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\mathbf{r}, t) - i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = 0. \quad (\text{A.3.14})$$

Then by considering

$$\Psi^*(\text{SE}) - \Psi(\text{SE}) = 0 \quad (\text{A.3.15})$$

we can derive a continuity equation

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0, \quad (\text{A.3.16})$$

where

$$\rho := \Psi^* \Psi, \quad \text{and} \quad \mathbf{j} := -\frac{i\hbar}{2m} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*). \quad (\text{A.3.17})$$

These are the **probability density** and **probability current** respectively.

We can do the same for the Klein–Gordon equation and obtain the same continuity equation, but now with

$$\rho := \frac{i\hbar}{2mc^2} \left(\varphi^* \frac{\partial \varphi}{\partial t} - \varphi \frac{\partial \varphi^*}{\partial t} \right), \quad \text{and} \quad \mathbf{j} := -\frac{i\hbar}{2m} (\varphi^* \nabla \varphi - \varphi \nabla \varphi^*). \quad (\text{A.3.18})$$

Here we chose to make \mathbf{j} identical to the non-relativistic Schrödinger current. We can show that in the non-relativistic limit ρ reduces to $\varphi^* \varphi$. There is a problem however, the candidate for the probability density, $\rho(x)$, is no longer positive definite, the negative energy solutions will have $\rho(x) < 0$. This means that there is no simple probability density interpretation for ρ from the Klein–Gordon equation.

A.3.3 Dirac Equation

Dirac, aiming to avoid the problems with negative energy/probability solutions, looked for a relativistic equation linear in $\partial/\partial t$. He then proceeded to argue that the equation must be linear in spatial derivatives, since in relativity we have to treat space and time on an equal footing. We therefore start with an equation of the form

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \hat{H} \psi(\mathbf{r}, t) \quad (\text{A.3.19})$$

where \hat{H} is linear in space and time derivatives. These requirements mean that we can write out a candidate Hamiltonian with some unknown coefficients:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = -i\hbar c \left(\alpha^1 \frac{\partial}{\partial x^1} + \alpha^2 \frac{\partial}{\partial x^2} + \alpha^3 \frac{\partial}{\partial x^3} \right) \psi(\mathbf{r}, t) + \beta mc^2 \psi(\mathbf{r}, t). \quad (\text{A.3.20})$$

We choose to have α^i and β be dimensionless by including the required factors of c , \hbar , and m , we also include a conventional factor of $-i$. Using the operator prescription $\hat{\mathbf{p}} = -i\hbar \nabla$ we can write this more compactly as

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = (c\boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta mc^2) \psi(\mathbf{r}, t). \quad (\text{A.3.21})$$

This is the **Dirac equation** for a free particle. We assume that the Hamiltonian is independent of position and time, which is what we would want for free particle solutions. Hence α^i and β must be independent of position and time, so commute with the derivatives, such as $\hat{\mathbf{p}}$, but not necessarily each other.

Next, we impose the relativistic energy-momentum equation:

$$\hat{H}^2 \psi(\mathbf{r}, t) = (c^2 |\hat{\mathbf{p}}|^2 + m^2 c^4) \psi(\mathbf{r}, t). \quad (\text{A.3.22})$$

We know that

$$\hat{H}^2 \psi(\mathbf{r}, t) = (c\boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta mc^2)(c\boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta mc^2) \psi(\mathbf{r}, t). \quad (\text{A.3.23})$$

Expanding the right hand side, while being careful about the ordering of terms, we get

$$\hat{H}^2 \psi(\mathbf{r}, t) = \{c^2 [(\alpha^1)^2 (\hat{p}^1)^2 + (\alpha^2)^2 (\hat{p}^2)^2 + (\alpha^3)^2 (\hat{p}^3)^2] + m^2 c^4 \beta^2\} \psi(\mathbf{r}, t) \quad (\text{A.3.24})$$

$$+ \{(\alpha^1 \alpha^2 + \alpha^2 \alpha^1) \hat{p}^1 \hat{p}^2 + (\alpha^2 \alpha^3 + \alpha^3 \alpha^2) \hat{p}^2 \hat{p}^3\} \psi(\mathbf{r}, t) \quad (\text{A.3.25})$$

$$+ mc^3 \{(\alpha^1 \beta + \beta \alpha^1) \hat{p}^1 + (\alpha^2 \beta + \beta \alpha^2) \hat{p}^2 + (\alpha^3 \beta + \beta \alpha^3) \hat{p}^3\} \psi(\mathbf{r}, t). \quad (\text{A.3.26})$$

The relativistic energy-momentum relation is satisfied if

$$(\alpha^1)^2 = (\alpha^2)^2 = (\alpha^3)^2 = \beta^2 = 1, \quad (\text{A.3.27})$$

$$\alpha^i \alpha^j + \alpha^j \alpha^i = 0, \quad (\text{A.3.28})$$

$$\alpha^i \beta + \beta \alpha^i = 0, \quad (\text{A.3.29})$$

which can be written more compactly as

$$\{\alpha^i, \alpha^j\} = 2\delta^{ij}, \quad \{\alpha^i, \beta\} = 0, \quad \text{and} \quad \beta^2 = 1, \quad (\text{A.3.30})$$

where $\{A, B\} = AB + BA$ is the **anticommutator**.

Clearly α^i and β cannot just be numbers for these requirements to hold. Instead, we assume they are matrices. Since \hat{H} is Hermitian they must be Hermitian, and hence are square matrices. This means they must have real eigenvalues. Since these matrices square to the identity their squares have eigenvalues whose product is 1, since the determinant is the product of the eigenvalues and the determinant of the unit matrix is 1. Since the eigenvalues are real the eigenvalues of α^i and β must be ± 1 .

Lemma A.3.31 We have that $\text{tr } \alpha^i = \text{tr } \beta = 0$.

Proof. Consider $\text{tr } \alpha^i$. Since $\beta^2 = 1$ we can insert β^2 into the argument of the trace and then use the cyclic property to get

$$\text{tr } \alpha^i = \text{tr}(\beta^2 \alpha^i) = \text{tr}(\beta \alpha^i \beta). \quad (\text{A.3.32})$$

Now using the anticommutation relations we can replace $\alpha^i \beta$ with $-\beta \alpha^i$ and then take the negative outside the trace to get

$$\text{tr } \alpha^i = \text{tr}(\beta \alpha^i \beta) = \text{tr}(\beta(-\beta \alpha^i)) = -\text{tr}(\beta^2 \alpha^i) = -\text{tr}(\alpha^i) \quad (\text{A.3.33})$$

since $\beta^2 = 1$, and so $\text{tr } \alpha^i = -\text{tr } \alpha^i$, which must mean $\text{tr } \alpha^i = 0$. Similarly, consider $\text{tr } \beta$, we have

$$\text{tr } \beta = \text{tr}((\alpha^i)^2 \beta) = \text{tr}(\alpha^i \beta \alpha^i) = -\text{tr}((\alpha^i)^2 \beta) = -\text{tr } \beta, \quad (\text{A.3.34})$$

and so $\text{tr } \beta = 0$. \square

So, if α^i and β are $n \times n$ matrices with eigenvalues ± 1 since the trace is the sum of the eigenvalues and the trace vanishes n must be even so that the eigenvalues can cancel pairwise. There is no 2×2 set of Hermitian traceless 2×2 matrices satisfying the anticommutation relations, although the Pauli matrices come close, satisfying

$$\{\sigma^i, \sigma^j\} = 2\delta^{ij}, \quad (\text{A.3.35})$$

but there is no viable candidate for β .

The simplest representation is 4×4 . The **standard representation** is

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \alpha^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{A.3.36})$$

$$\alpha^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \quad \alpha^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \quad (\text{A.3.37})$$

These can be written in a 2×2 block form as

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad \alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix} \quad (\text{A.3.38})$$

where $\sigma = (\sigma^1, \sigma^2, \sigma^3)$ are the Pauli matrices. We can also write these as $\alpha^i = \sigma^1 \otimes \sigma^i$. It's not too difficult to check that these matrices have the desired properties.

Since the Hamiltonian is a 4×4 matrix the wave function, $\psi(\mathbf{r}, t)$, must be a four component column vector:

$$\psi(\mathbf{r}, t) = \begin{pmatrix} \psi_1(\mathbf{r}, t) \\ \psi_2(\mathbf{r}, t) \\ \psi_3(\mathbf{r}, t) \\ \psi_4(\mathbf{r}, t) \end{pmatrix}. \quad (\text{A.3.39})$$

A.3.4 Probability Density

The Dirac equation for a free particle is

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = (-i\hbar c \boldsymbol{\alpha} \cdot \vec{\nabla} + \beta mc^2) \psi(\mathbf{r}, t). \quad (\text{A.3.40})$$

Here $\vec{\nabla}$ is the normal gradient operator acting to the right. From this we can construct a probability density. Take the Hermitian conjugate of this equation:

$$-i\hbar \frac{\partial}{\partial t} \psi^\dagger(\mathbf{r}, t) = \psi^\dagger(\mathbf{r}, t) (i\hbar c \boldsymbol{\alpha} \cdot \vec{\nabla} + \beta mc^2). \quad (\text{A.3.41})$$

Here $\vec{\nabla}$ denotes a gradient operator acting to the left. Note that ψ^\dagger is a row vector. Multiplying the first equation by ψ^\dagger on the left, and the second by ψ on the right we obtain

$$i\hbar \frac{\partial}{\partial t} (\psi^\dagger \psi) = -i\hbar c (\psi^\dagger \boldsymbol{\alpha} \cdot \vec{\nabla} \psi + \psi^\dagger \boldsymbol{\alpha} \cdot \vec{\nabla} \psi) \quad (\text{A.3.42})$$

$$= -i\hbar c (\psi^\dagger \boldsymbol{\alpha} \cdot \nabla \psi + (\nabla \psi^\dagger) \cdot \boldsymbol{\alpha} \psi) \quad (\text{A.3.43})$$

$$= -i\hbar c \nabla \cdot (\psi^\dagger \boldsymbol{\alpha} \psi). \quad (\text{A.3.44})$$

We can write this as a continuity equation,

$$\nabla \cdot \mathbf{j} + \frac{1}{c} \frac{\partial \rho}{\partial t} = 0, \quad (\text{A.3.45})$$

where

$$\rho := \psi^\dagger \psi, \quad \text{and} \quad \mathbf{j} = \psi^\dagger \boldsymbol{\alpha} \psi. \quad (\text{A.3.46})$$

We can write this in a covariant form:

$$\partial_\mu j^\mu = 0, \quad (\text{A.3.47})$$

where $j = (\rho, \mathbf{j})$ is the **four-probability current**.

This tells us that $\psi^\dagger \psi$ transforms as the time component of a four-vector, and $\psi^\dagger \boldsymbol{\alpha} \psi$ transforms as the corresponding space part.

Note that $\rho = |\psi|^2$ is positive definite, solving one of our earlier problems with the Klein–Gordon equation.

A.3.5 Free Particle Solutions

The Dirac equation has plane-wave solutions:

$$\psi(\mathbf{r}, t) = \psi(x) = \exp[-ik_\mu x^\mu] u(\mathbf{p}) = \exp\left[-\frac{i}{\hbar}(cp_0 t - \mathbf{p} \cdot \mathbf{r})\right] u(\mathbf{p}). \quad (\text{A.3.48})$$

Here $u(\mathbf{p})$ is a four component column vector.

Substituting this into the Dirac equation we get

$$p^0 u = (\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc) u. \quad (\text{A.3.49})$$

We interpret this solution as representing a particle of energy cp^0 and momentum \mathbf{p} . Putting in the matrices α^i and β this becomes

$$\begin{pmatrix} -p^0 + mc & 0 & p^3 & p^1 - ip^2 \\ 0 & -p^0 + mc & p^1 + ip^2 & -p^3 \\ p^3 & p^1 - ip^2 & -(p^0 + mc) & 0 \\ p^1 + ip^2 & -p^3 & 0 & -(p^0 + mc) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0. \quad (\text{A.3.50})$$

The condition for a nontrivial solution for u_i to exist is that the determinant of the matrix vanishes. Computing the determinant we find that the determinant vanishes exactly when

$$[m^2 c^2 + |\mathbf{p}|^2 - (p^0)^2]^2 = 0, \quad (\text{A.3.51})$$

which is exactly the energy-momentum relation. Taking the square root we have

$$p^0 = \pm \sqrt{m^2 c^2 + |\mathbf{p}|^2}, \quad (\text{A.3.52})$$

so we still have negative energy solutions.

A.3.5.1 Positive Energy Solutions

First suppose $p^0 > 0$, that is

$$p^0 = +\sqrt{m^2 c^2 + |\mathbf{p}|^2} =: p_+^0 = \frac{E}{c}, \quad (\text{A.3.53})$$

where we choose to take $E > 0$ as the magnitude of the energy. Substituting into the first two rows of Equation (A.3.50) we get

$$0 = (-p_+^0 + mc)u_1 + p_3 u_3 + (p^1 - ip^2)u_4, \quad (\text{A.3.54})$$

$$0 = (-p_+^0 + mc)u_2 + (p^1 + ip^2)u_3 - p^3 u_4. \quad (\text{A.3.55})$$

Only two of the four u_i are linearly independent, so we are free to fix two values. The conventional choice is either $u_1 = 1$ and $u_2 = 0$, or $u_1 = 0$ and $u_2 = 1$. The first choice gives

$$u_3 = \frac{p^3}{p_+^0 + mc}, \quad \text{and} \quad u_4 = \frac{p^1 + ip^2}{p_+^0 + mc}, \quad (\text{A.3.56})$$

whereas the second gives

$$u_3 = \frac{p^1 - ip^2}{p_+^0 + mc}, \quad \text{and} \quad u_4 = \frac{-p^3}{p_+^0 + mc}. \quad (\text{A.3.57})$$

A.3.5.2 Negative Energy Solutions

Now suppose $p^0 < 0$, that is

$$p^0 = -\sqrt{m^2c^2 + |\mathbf{p}|^2} =: p_-^0 = -\frac{E}{c}, \quad (\text{A.3.58})$$

where again, we choose to have $E > 0$. We can follow a similar process to the positive energy solutions, but for the lower two rows of Equation (A.3.50). Again, we make an arbitrary choice for two components, the conventions being either $u_3 = 1$ and $u_4 = 0$, or $u_3 = 0$ and $u_4 = 1$. The first choice leads to

$$u_1 = \frac{p^3}{p_-^0 - mc}, \quad \text{and} \quad u_2 = \frac{p^1 + ip^2}{p_-^0 - mc}, \quad (\text{A.3.59})$$

and the second gives

$$u_1 = \frac{p^1 - ip^2}{p_-^0 - mc}, \quad \text{and} \quad u_2 = \frac{-p^3}{p_-^0 - mc}. \quad (\text{A.3.60})$$

A.3.5.3 Summary

The positive energy solutions with four-momentum $p_+^\mu = (p_+^0, \mathbf{p})$ are

$$\omega^1(\mathbf{p}) = \begin{pmatrix} 1 \\ 0 \\ \frac{p^3}{p_+^0 + mc} \\ \frac{p^1 + ip^2}{p_+^0 + mc} \end{pmatrix}, \quad \text{and} \quad \omega^2(\mathbf{p}) = \omega^1(\mathbf{p}) = \begin{pmatrix} 0 \\ 1 \\ \frac{p^1 - ip^2}{p_+^0 + mc} \\ \frac{-p^3}{p_+^0 + mc} \end{pmatrix}. \quad (\text{A.3.61})$$

The negative energy solutions with four-momentum $p_-^\mu = (p_-^0, -\mathbf{p})$ are

$$\omega^1(\mathbf{p}) = \begin{pmatrix} -\frac{p^1 - ip^2}{p_+^0 + mc} \\ \frac{p^3}{p_+^0 + mc} \\ 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad \omega^2(\mathbf{p}) = \omega^1(\mathbf{p}) = \begin{pmatrix} -\frac{p^3}{p_+^0 + mc} \\ \frac{p^1 + ip^2}{p_+^0 + mc} \\ 1 \\ 0 \end{pmatrix}. \quad (\text{A.3.62})$$

Note that there are varying conventions on how the ω^μ are labelled.

A.3.5.4 Rest Frame Solutions

When the particle is at rest, that is $\mathbf{p} = \mathbf{0}$, we have

$$\omega^1(\mathbf{0}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \omega^2(\mathbf{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \omega^3(\mathbf{0}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad \omega^4(\mathbf{0}) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \quad (\text{A.3.63})$$

These are such that the solutions

$$\psi^{(1)} = \exp\left[-\frac{imc^2t}{\hbar}\right] \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \psi^{(2)} = \exp\left[-\frac{imc^2t}{\hbar}\right] \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad (\text{A.3.64})$$

are degenerate in energy. There should therefore be another operator which commutes with the Hamiltonian and whose eigenvalues label the states. One such operator is

$$\Sigma^3 := \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (\text{A.3.65})$$

The rest frame four-component spinors, $\omega^\mu(\mathbf{0})$, are eigenvectors of Σ^3 with eigenvalues ± 1 . This suggests that we can interpret the Dirac equation as describing a spin 1/2 particle, and the eigenvalues of Σ^3 are interpreted as the 3rd component of the spin in the rest frame.

The notation Σ^3 suggests we should define

$$\Sigma := \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}, \quad \text{or} \quad \Sigma^i := \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}. \quad (\text{A.3.66})$$

We then have

$$\left(\frac{1}{2}\hbar\Sigma\right) \cdot \left(\frac{1}{2}\hbar\Sigma\right) = \frac{3}{4}\hbar^2\hat{1} \quad (\text{A.3.67})$$

and $\hbar\Sigma^3/2$ has eigenvalues of $\pm\hbar/2$ suggesting that the spin operator is

$$\hat{\mathbf{s}} = \frac{1}{2}\hbar\Sigma. \quad (\text{A.3.68})$$

Note that Σ does not commute with the Hamiltonian in frames other than the rest frame, when $\mathbf{p} = \mathbf{0}$. The operator $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ also doesn't commute with the Hamiltonian. However, the operator

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{s}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} + \frac{1}{2}\hbar\Sigma \quad (\text{A.3.69})$$

does commute with the Hamiltonian. This suggests we can interpret $\hat{\mathbf{J}}$ as the total angular momentum.

When $\mathbf{p} \neq \mathbf{0}$ there are two independent states for any fixed four-momentum. Therefore there exists some operator which commutes with $c\boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta mc^2$ and labels the states. One option is the **helicity operator**

$$\hat{h}(\mathbf{p}) = \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} & 0 \\ 0 & \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} \end{pmatrix}. \quad (\text{A.3.70})$$

This commutes with $c\boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta mc^2$ and has eigenvalues ± 1 . We can therefore choose general plane-wave states with $\mathbf{p} \neq \mathbf{0}$ to be helicity states.

B

Hamiltonian Mode Expansion

The Hamiltonian for the free Klein–Gordon field is

$$H = \frac{1}{2} \int d^3x \{(\pi(x))^2 + (\nabla\varphi(x))^2 + m^2(\varphi(x))^2\} \equiv H_1 + H_2 + H_3. \quad (\text{B.0.1})$$

In [Section 6.3](#) we claimed that the time independent contribution from each term is

$$H_1 \rightarrow \frac{1}{8} \int \frac{d^3p}{(2\pi)^3} \{a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})\}, \quad (\text{B.0.2})$$

$$H_2 \rightarrow \frac{1}{8} \int \frac{d^3p}{(2\pi)^3} \frac{\mathbf{p}^2}{\omega(\mathbf{p})^2} \{a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})\}, \quad (\text{B.0.3})$$

$$H_3 \rightarrow \frac{1}{8} \int \frac{d^3p}{(2\pi)^3} \frac{m^2}{\omega(\mathbf{p})^2} \{a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})\}. \quad (\text{B.0.4})$$

We also claimed that the time dependent parts cancel in $H_1 + H_2 + H_3$. In this section we complete the full calculation showing this to be true.

Start with the mode expansion of $\varphi(x)$:

$$\varphi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega(\mathbf{p})} [a(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (\text{B.0.5})$$

Taking the time derivative we get

$$\pi(x) = \dot{\varphi}(x) = -\frac{i}{2} \int \frac{d^3p}{(2\pi)^3} [a(\mathbf{p})e^{-ip \cdot x} - a^\dagger(\mathbf{p})e^{ip \cdot x}]. \quad (\text{B.0.6})$$

We can then compute π^2 :

$$\begin{aligned} \pi(x)^2 = -\frac{1}{4} \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3p'}{(2\pi)^3} [a(\mathbf{p})e^{-ip \cdot x} - a^\dagger(\mathbf{p})e^{ip \cdot x}] \\ [a(\mathbf{p}')e^{-ip' \cdot x} - a^\dagger(\mathbf{p}')e^{ip' \cdot x}] \end{aligned} \quad (\text{B.0.7})$$

Expand the integrand to get

$$\begin{aligned} a(\mathbf{p})a(\mathbf{p}')e^{-i(p+p') \cdot x} - a(\mathbf{p})a^\dagger(\mathbf{p}')e^{-i(p-p') \cdot x} \\ - a^\dagger(\mathbf{p})a(\mathbf{p}')e^{i(p-p') \cdot x} + a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')e^{i(p+p') \cdot x} \end{aligned} \quad (\text{B.0.8})$$

Now consider the identities

$$\int d^3x e^{\pm i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} = (2\pi)^3 \delta(\mathbf{p} + \mathbf{p}') e^{\pm 2i\omega(\mathbf{p})t} \quad (\text{B.0.9})$$

$$\int d^3x e^{\pm i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} = (2\pi)^3 \delta(\mathbf{p} - \mathbf{p}'). \quad (\text{B.0.10})$$

With these we can perform the x integral to get

$$\begin{aligned} H_1 = & -\frac{1}{8} \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3p'}{(2\pi)^3} [a(\mathbf{p})a(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} + \mathbf{p}') e^{-2i\omega(\mathbf{p})t} \\ & - a(\mathbf{p})a^\dagger(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') - a^\dagger(\mathbf{p})a(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') \\ & + a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} + \mathbf{p}') e^{2i\omega(\mathbf{p})t}]. \end{aligned} \quad (\text{B.0.11})$$

Now performing the p' integral we get

$$\begin{aligned} H_1 = & -\frac{1}{8} \int \frac{d^3p}{(2\pi)^3} [a(\mathbf{p})a(-\mathbf{p})e^{-2i\omega(\mathbf{p})} - a(\mathbf{p})a^\dagger(\mathbf{p}) \\ & + a^\dagger(\mathbf{p})a(\mathbf{p}) + a^\dagger(\mathbf{p})a^\dagger(-\mathbf{p})e^{2i\omega(\mathbf{p})t}]. \end{aligned} \quad (\text{B.0.12})$$

The time independent part of this is

$$H_1 \rightarrow \frac{1}{8} \int \frac{d^3p}{(2\pi)^3} [a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})]. \quad (\text{B.0.13})$$

The time dependent part is

$$H_1^t = -\frac{1}{8} \int \frac{d^3p}{(2\pi)^3} [a(\mathbf{p})a(-\mathbf{p})e^{-2i\omega(\mathbf{p})} + a^\dagger(\mathbf{p})a^\dagger(-\mathbf{p})e^{2i\omega(\mathbf{p})t}]. \quad (\text{B.0.14})$$

This should (hopefully) cancel out with the other two terms.

Now consider the second term. We start by computing $\nabla\varphi$ in terms of the mode operators, which gives

$$\nabla\varphi(x) = \frac{i}{2} \int \frac{d^3p}{(2\pi)^3} \frac{\mathbf{p}}{\omega(\mathbf{p})} [a(\mathbf{p})e^{-ip\cdot x} - a^\dagger(\mathbf{p})e^{ip\cdot x}]. \quad (\text{B.0.15})$$

We can then compute $(\nabla\varphi)^2$:

$$\begin{aligned} (\nabla\varphi(x))^2 = & -\frac{1}{4} \int \frac{d^3p}{(2\pi)^3} \frac{\mathbf{p} \cdot \mathbf{p}'}{\omega(\mathbf{p})\omega(\mathbf{p}')} [a(\mathbf{p})e^{-ip\cdot x} - a^\dagger(\mathbf{p})e^{ip\cdot x}] \\ & [a(\mathbf{p}')e^{-ip'\cdot x} - a^\dagger(\mathbf{p}')e^{ip'\cdot x}]. \end{aligned} \quad (\text{B.0.16})$$

Expanding the integrand we get

$$\begin{aligned} & a(\mathbf{p})a(\mathbf{p}')e^{-i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} - a(\mathbf{p})a^\dagger(\mathbf{p}')e^{-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} \\ & - a^\dagger(\mathbf{p})a(\mathbf{p}')e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} + a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')e^{i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}}. \end{aligned} \quad (\text{B.0.17})$$

We can then perform the integral over x and we'll get Dirac deltas:

$$\begin{aligned} H_2 = & -\frac{1}{8} \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3p'}{(2\pi)^3} \frac{\mathbf{p} \cdot \mathbf{p}'}{\omega(\mathbf{p})\omega(\mathbf{p}')} [a(\mathbf{p})a(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} + \mathbf{p}') e^{-2i\omega(\mathbf{p})t} \\ & - a(\mathbf{p})a^\dagger(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') - a^\dagger(\mathbf{p})a(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') \\ & + a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} + \mathbf{p}') e^{2i\omega(\mathbf{p})t}] \end{aligned} \quad (\text{B.0.18})$$

Performing the p' integral we get

$$H_2 = -\frac{1}{8} \int \frac{d^3 p}{(2\pi)^3} \frac{|\mathbf{p}|^2}{\omega(\mathbf{p})^2} [-a(\mathbf{p})a(-\mathbf{p})e^{-2i\omega(\mathbf{p})t} - a(\mathbf{p})a^\dagger(\mathbf{p}) - a^\dagger(\mathbf{p})a(\mathbf{p}) - a^\dagger(\mathbf{p})a^\dagger(-\mathbf{p})e^{2i\omega(\mathbf{p})t}]. \quad (\text{B.0.19})$$

Note the extra negative sign from setting $\mathbf{p} = -\mathbf{p}$ in the factor out the front for the time dependent parts. The time independent part of this is

$$H_2 \rightarrow \int \frac{d^3 p}{(2\pi)^3} \frac{|\mathbf{p}|^2}{\omega(\mathbf{p})^2} [a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})]. \quad (\text{B.0.20})$$

The time dependent part is

$$H_2^t = \frac{1}{8} \int \frac{d^3 p}{(2\pi)^3} \frac{|\mathbf{p}|^2}{\omega(\mathbf{p})^2} [a(\mathbf{p})a(-\mathbf{p})e^{-2i\omega(\mathbf{p})t} + a^\dagger(\mathbf{p})a^\dagger(-\mathbf{p})e^{2i\omega(\mathbf{p})t}]. \quad (\text{B.0.21})$$

Finally, consider the φ^2 term:

$$\varphi(x)^2 = \frac{1}{4} \int \frac{d^3 p}{(2\pi)^3} \int \frac{d^3 p'}{(2\pi)^3} \frac{1}{\omega(\mathbf{p})\omega(\mathbf{p}')} [a(\mathbf{p})e^{-ip \cdot x} + a^\dagger(\mathbf{p})e^{ip \cdot x}] [a(\mathbf{p}')e^{-ip' \cdot x} + a^\dagger(\mathbf{p}')e^{ip' \cdot x}]. \quad (\text{B.0.22})$$

Expand the integrand to get

$$a(\mathbf{p})a^\dagger(\mathbf{p}')e^{-i(p+p') \cdot x} + a(\mathbf{p})a^\dagger(\mathbf{p}')e^{-i(p-p') \cdot x} + a^\dagger(\mathbf{p})a(\mathbf{p}')e^{i(p-p') \cdot x} + a^\dagger(\mathbf{p})a(\mathbf{p}')e^{i(p+p') \cdot x}. \quad (\text{B.0.23})$$

We can then perform the x integral to get Dirac deltas:

$$H_3 = \frac{1}{8} \int \frac{d^3 p}{(2\pi)^3} \int \frac{d^3 p'}{(2\pi)^3} \frac{m^2}{\omega(\mathbf{p})\omega(\mathbf{p}')} [a(\mathbf{p})a(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} + \mathbf{p}')e^{-2i\omega(\mathbf{p})t} + a(\mathbf{p})a^\dagger(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') + a^\dagger(\mathbf{p})a(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}') + a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')(2\pi)^3 \delta(\mathbf{p} + \mathbf{p}')e^{2i\omega(\mathbf{p})t}]. \quad (\text{B.0.24})$$

Performing the p' integral we get

$$H_3 = \frac{1}{8} \int \frac{d^3 p}{(2\pi)^3} \frac{m^2}{\omega(\mathbf{p})^2} [a(\mathbf{p})a(-\mathbf{p})e^{-2i\omega(\mathbf{p})t} + a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p}) + a^\dagger(\mathbf{p})a^\dagger(-\mathbf{p})e^{2i\omega(\mathbf{p})t}]. \quad (\text{B.0.25})$$

The time independent part of this is

$$H_3 \rightarrow \frac{1}{8} \int \frac{d^3 p}{(2\pi)^3} \frac{m^2}{\omega(\mathbf{p})^2} [a(\mathbf{p})a^\dagger(\mathbf{p}) + a^\dagger(\mathbf{p})a(\mathbf{p})]. \quad (\text{B.0.26})$$

The time dependent part is

$$H_3^t = \frac{1}{8} \int \frac{d^3 p}{(2\pi)^3} \frac{m^2}{\omega(\mathbf{p})^2} [a(\mathbf{p})a(-\mathbf{p})e^{-2i\omega(\mathbf{p})t} + a^\dagger(\mathbf{p})a^\dagger(-\mathbf{p})e^{2i\omega(\mathbf{p})t}]. \quad (\text{B.0.27})$$

Now we consider the time dependent parts. First notice that each has a factor of

$$\frac{1}{8} \int \frac{d^3 p}{(2\pi)^3} [a(\mathbf{p})a(-\mathbf{p})e^{-2i\omega(\mathbf{p})t} + a^\dagger(\mathbf{p})a^\dagger(-\mathbf{p})e^{2i\omega(\mathbf{p})t}], \quad (\text{B.0.28})$$

so we focus on the terms not present here:

$$-1 + \frac{|\mathbf{p}|^2}{\omega(\mathbf{p})^2} + \frac{m^2}{\omega(\mathbf{p})^2} = \frac{|\mathbf{p}|^2 + m^2}{\omega(\mathbf{p})^2} - 1 = 1 - 1 = 0 \quad (\text{B.0.29})$$

since by definition $\omega(\mathbf{p})^2 = |\mathbf{p}|^2 + m^2$.

Acronyms

C

CCR: Canonical Commutation Relation [24](#)

E

ETCR: Equal Time Commutation Relation [29](#)

Q

QCD: Quantum Chromodynamics [1](#)

QED: Quantum Electrodynamics [1](#)

QFT: Quantum Field Theory [1](#)

Index

Symbols

η , *see* metric tensor
 θ , Heaviside step function, 40
 ∂_μ , 182
 φ^3 theory, 56
 5th gamma matrix, *see* fifth gamma matrix

A

action, 11
 annihilation operator, 37, 46
 anticommutator, 111, 189
 antiunitary operator, 106

B

boost, 179

C

canonical commutation relation, 24
 canonical momentum, 13, 16
 charge conjugation, 19
 Clifford algebra, 113
 commutator, 25
 contravariant, 178
 correlation function, 134
 coupling constant, 56
 covariant, 178
 creation operator, 37, 46

D

d'Alembert operator, 6
 d'Alembertian, 182
 Dirac adjoint, 119
 Dirac bilinear, 120
 Dirac equation, 110, 188
 Dirac Hamiltonian, 110
 Dirac matrices, 112
 Dirac picture, *see* interaction picture

Dirac representation, 112
 Dyson series, 59, 62

E

Ehrenfest's theorem, 28
 electromagnetic field, 9
 electromagnetic field strength, 9
 electromagnetic field strength tensor, 184
 electroweak, 1
 energy-momentum relation, 6, 182
 energy-momentum tensor, 23
 equal time commutation relation, 29
 Euler-Lagrange equation, 12, 16
 expectation value
 with respect to a measure, 133

F

Feynman amplitude, 84
 Feynman propagator, 72, 161
 Feynman rules, 87
 φ^3 theory, 87
 fifth gamma matrix, 113
 first quantisation, 37
 Fock space, 46
 four-current, 183
 four-potential, 183
 four-probability current, 190
 free field Lagrangian, 56
 functional, 150

G

gamma matrices, 112
 gauge transformation, 184
 Gaussian
 correlation function, 135
 generating function, 134
 integral, 130

measure, 134
 generalised
 coordinate, 11
 velocity, 11
 generating function, 133

H

Hamilton's equations, 14, 16
 Hamilton's principle, 11
 Hamiltonian, 13
 Hamiltonian density, 16
 Heaviside step function, 40
 Heaviside–Lorentz units, 182
 Heisenberg
 equation, 26
 picture, 25
 helicity, 127
 helicity operator, 193
 Hermitian
 conjugate, 24
 operator, 24
 homogeneous Lorentz group, 179
 homogeneous Lorentz transformations, 179

I

idempotent, 104
 improper Lorentz transformation, 180
 interaction picture, 58

K

Klein–Gordon equation, 8, 186

L

Lagrangian, 11
 density, 15
 light-like, 181
 loop diagram, 80
 Lorentz group, 5
 Lorentz transformation, 179
 Lorenz gauge, 9
 Lorenz gauge condition, 184

M

Mandelstam invariants, 85
 mass-shell, 182
 mass-shell condition, 182
 Maxwell tensor, *see* electromagnetic field strength tensor
 Maxwell's equations, 182
 metric tensor, 178

minimal coupling prescription, 185
 mode expansion, 40

N

Noether's theorem, 20
 normal ordering, 48
 number density operator, 47

O

$O^+(1, 3)$, *see also* orthochronous Lorentz group
 on-shell, 8
 orthochronous Lorentz group, 181
 orthochronous Lorentz transformation, 181

P

parity, 19
 path integral, 146
 phase space measure, 89
 Poisson bracket, 14
 polarisation vector, 10
 probability current, 187
 probability density, 187
 proper Lorentz transformation, 180
 proper orthochronous Lorentz group, 5, 181
 proper orthochronous Lorentz transformation, 5
 pseudoscalar, 105

Q

quantum chromodynamics, 1
 quantum electrodynamics, 1
 quantum field theory, 1

S

S matrix, 57
 s -channel, 79
 scalar product, 179
 Schrödinger
 equation, 7, 25, 186
 picture, 25
 second quantisation, 37
 slash notation, 112
 $SO(1, 3)$, *see also* proper Lorentz transformation
 $SO^+(1, 3)$, *see also* proper orthochronous Lorentz group
 space-like, 181
 spacetime interval, 181

spacetime inversion, [181](#)
spatial inversion, [181](#)
spinor, [112](#)
standard representation, [189](#)
stress-energy tensor, *see* energy-momentum
tensor

T

t -channel, [80](#)
time dependent Schrödinger equation,
[25](#)
time ordering, [61](#)
time reversal, [19](#), [181](#)
time-like, [181](#)
transverse, [10](#)
tree diagram, [80](#)

V

vacuum state, [46](#)
virtual particle, [63](#)

W

wave function, [7](#)
Weyl ordering, [153](#)
Wick contraction, [67](#)
Wick's theorem, [70](#), [136](#)

Y

Yukawa interaction, [102](#)