A One-Semester Course on Quantum Field Theory

Thomas DeGrand

July 10, 2025

Contents

1 Introduction **7**

2 A reading list for quantum field theory **13**

3 An overview of the course **21**

4 Lagrangians, Hamiltonians, quantum field theories **27**

4.1 The oscillator 29

4.2 Many uncoupled oscillators 32

4.3 Our first quantum field theory 33

5 Quantum field theory for many-body problems **41**

5.1 Particles to waves to particles again 42

5.2 The free electron gas at zero temperature 46

5.3 Particles that interact among themselves 50

5.4 Excitation in a Bose-Einstein condensate 52

6 Quantum field theory from a Hamiltonian point of view **57**

6.1 Time dependent perturbation theory 58

6.2 The S-matrix 60

6.2.1 Unitarity 60

6.2.2 Energy conservation . . . 61
6.2.3 Momentum conservation . . . 61
6.2.4 Lorentz invariance . . . 62
6.3 Relativistic spin-zero quantum fields . . . 63
6.4 Example 1: ϕ^4 field theory . . . 66
6.5 Example 2: Discovering the propagator . . . 70
6.6 Example 3: Another amplitude . . . 74
6.7 Wick's theorem . . . 74
6.8 A glance at higher orders . . . 77
6.9 Feynman rules . . . 80
6.10 Conclusion . . . 81

7 Path integrals in quantum mechanics and quantum field theories 83
7.1 Path integrals in quantum mechanics . . . 84
7.2 Path integrals in quantum field theory (for bosons) . . . 87
7.3 The propagator and Feynman rules, again . . . 90

8 Spin-$\frac{1}{2}$: foundations, generalities, practicalities 101
8.1 The Dirac equation . . . 102
8.2 The Lorentz group and relativistic quantum field theories . . . 102
8.3 Technology for spinors . . . 108
8.4 Technical details about solutions of the Dirac equation . . . 112
8.5 The spin-statistics theorem ($j = 0$ and $j = 1/2$ only) . . . 120
8.6 The fermion propagator . . . 122
8.7 A sample problem – an electron in an external field . . . 122

9 Quantum electrodynamics, for itself and as a example of a gauge theory 129

9.1 Symmetries and conservation laws 130
9.2 Beginning to think about quantum electrodynamics 133
9.3 Quantizing gauge theories 136
9.4 Gaussian integrals with zero eigenvalues 140
9.5 More about gauge theories 141
9.6 A well behaved path integral 146
9.7 Feynman rules for QED 150
9.8 Scattering processes in lowest order – e^+e^- annihilation 153
9.9 Scattering processes in lowest order – Compton scattering and related 158
9.10 Bremsstrahlung 159

10 Renormalization and the renormalization group 163

10.1 Regularization 164
10.2 Dimensional analysis and engineering dimensions 168
10.3 Renormalization in scalar field theory 171
10.4 Towards the renormalization group 175
10.5 Renormalization prescriptions 184
10.6 More general couplings 186
10.7 Where do Lagrangians come from, or the story of effective field theories . . . 187
10.8 A few words about renormalization in QED 194

11 Closing thoughts 203

A Conventions for fields and scattering amplitudes 205

A.1 Kinematics conventions 206
A.2 Particles in a box 207
A.3 Delta function normalization 209

A.4 The normalization everybody uses 210

Chapter 1

Introduction

Over the years I have occasionally taught the graduate course on quantum field theory at the University of Colorado. The book you are holding is my attempt to make my notes (somewhat) legible. It descends from notes prepared for my 2017 course, and presumably I will add to it as time goes on.

There are two ways that quantum field theory is presented in a graduate course. One can start either with a quantum Hamiltonian, or with path integrals. I chose to start with Hamiltonians, and then to introduce path integrals later. The reason I did that is that most students at Colorado are experimentalists working with atoms or molecules, and my experience is that they are more familiar with Hamiltonians given what they have learned about nonrelativistic quantum mechanics.

David Tong's lectures [1] are close to what I taught. There are a lot of useful links to other material on his web page.

Sidney Coleman's lecture notes [2] also basically parallel what I am doing. They are different because his audience was Harvard theoretical physics graduate students, and because he was much more creative than I am.

There are videos of both these teachers' courses out there, but I am not a video learner so you are on your own with all that.

In addition, there are bits of the following books which come close to these lectures: Schwartz 7.2-7.4, Srednicki Ch. 3, Peskin and Schroeder Sec. 2.4. Weinberg Volume 1, chapter 6 parallels a lot of what I did in these notes.

And now a glance at what is to come:

Chapter 2 is a suggested reading list for the interested student. Chapter 3 is a look at the "big picture"before we actually get down to work.

In Chapter 4 I introduce the story of "canonical quantization," the passage from Lagrangian to Hamiltonian, and present our first quantum field theory, the quantum mechanical string.

Chapter 5 is a discussion of the quantum field theoretic approach to nonrelativistic many-body problems. It is the mostly the same chapter from my quantum mechanics notes.

Chapter 6 tells the story of perturbative quantum field theory for scalar fields, starting with a Hamiltonian and ending with the Feynman rules for scattering problems.

I then repeat the physics story (perturbation theory for scalar field theory), but from the

point of view of path integrals. This is Chapter 7.

We then look at the interplay of spin one-half (and other spins) with relativity, in Ch. 8.

Chapter 9 pulls the preceding chapters together to look at quantum electrodynamics.

Chapter 10 describes the physics of renormalization, introducing the concept of the renormalization group and illustrating the subject with calculations in scalar field theory.

Appendix A is a long discussion about the conventions for defining field operators in relativistic field theory. This is a subject which is impossible to teach (everyone falls asleep) but absolutely necessary to write down somewhere, since so much of the discussion depends on the choice of conventions.

These notes are full of typos. Please report them to me.

Bibliography

[1] David Tong: Lectures on Quantum Field Theory `https://www.damtp.cam.ac.uk/user/tong/qft.html`

[2] S. Coleman, "Notes from Sidney Coleman's Physics 253a: Quantum Field Theory," [arXiv:1110.5013 [physics.ed-ph]].

Chapter 2

A reading list for quantum field theory

Here are a collection of books and papers I have found useful for describing various aspects of quantum field theory. The preprints with "arXiv" labels come from the data base `http://arXiv.org/` (which is the standard data base for most of theoretical physics).

As I said in the Introduction, David Tong's [1] and Sidney Coleman's [2] lecture notes are probably the closest courses to the way I will do things.

The two books I debated between as a text in 2017 (not that I ever followed anything very closely) were Zee [3] and Srednicki [4]. They are quite complementary (with identical advice for the reader). I used Zee the penultimate time I taught this course and the students didn't care for it so I tried Srednicki in 2017. But it's not really close to how I present things. Two very particle-physics oriented books are Schwartz [5] and Peskin and Schroeder [6]. Ramond [7] has things other books don't, specifically the use of generating functionals for perturbation theory and a nice calculation of beta functions. Ryder [8] is concise but complete; I often put it on reserve when I teach the formal parts of electrodynamics. Weinberg's books [9, 10] are very complete, but maybe they aren't really textbooks. (My first QFT course was from Weinberg.) Finally we are far in the past with Bjorken and Drell [11, 12]. In their day they were quite influential and for a few special subjects (solving the Dirac equation, doing tree-level QED) they are still worth consulting. I've put Refs. [3, 4, 5, 6, 7, 8] on reserve in the Engineering Library, which is in the basement of the math building. Try to violate social norms and go in the stacks there to browse the QFT book section. You will find a lot of books, including (hopefully) the one for you.

I added Martin's supersymmetry primer [13] for its Sec. 2, which has a terse description of Dirac, Weyl and Majorana fermions.

The many - body physics book I own is Fetter and Walecka [14]: very old but still full of useful things.

I included some statistical mechanics oriented texts, Refs. [15, 16, 17, 18, 19, 20] on my list. I have a problem with condensed matter field theory books in that I find their treatment of path integrals and perturbation theory somewhat unsystematic, but then I am a particle physicist. Maybe I am missing something! The classic review article on the renormalization group by Wilson and Kogut [21] is not to be missed. Polyakov [22] had interesting things to say. Finally I include my own book [23] (not to be selfish, but bits of this course washed up there) and for BEC aficionados, Pethick and Smith[24].

James Thompson suggested three quantum optics books [25, 26, 27].

Various "inspirational" articles [28, 29, 30, 31, 32] and some introductions to effective field theory [33, 34, 35, 36, 37] round out the list.

Remember, you can't learn quantum field theory from any one book or any one course. Bohr once said that "The opposite of truth is falsehood but the opposite of a great truth is another great truth." Quantum field theory contains many great truths.

Bibliography

[1] David Tong: Lectures on Quantum Field Theory `https://www.damtp.cam.ac.uk/user/tong/qft.html`

[2] S. Coleman, "Notes from Sidney Coleman's Physics 253a: Quantum Field Theory," [arXiv:1110.5013 [physics.ed-ph]].

[3] A. Zee, "Quantum field theory in a nutshell," *Princeton, UK: Princeton Univ. Pr. (2003)*

[4] M. Srednicki, "Quantum field theory," *Cambridge, UK: Univ. Pr. (2007)*

[5] M. D. Schwartz, "Quantum Field Theory and the Standard Model," *Cambridge, UK: Univ. Pr. (2014)*

[6] M. E. Peskin and D. V. Schroeder, "An Introduction To Quantum Field Theory," *Reading, USA: Addison-Wesley (1995)*

[7] P. Ramond, "FIELD THEORY: A MODERN PRIMER," Front. Phys. **74**, 1 (1989).

[8] L. H. Ryder, "QUANTUM FIELD THEORY," *Cambridge, Uk: Univ. Pr. (1985)*

[9] S. Weinberg, "The Quantum theory of fields. Vol. 1: Foundations," *Cambridge, UK: Univ. Pr. (1995)*

[10] S. Weinberg, "The quantum theory of fields. Vol. 2: Modern applications," *Cambridge, UK: Univ. Pr. (1996)*

[11] J. D. Bjorken and S. D. Drell, "Relativistic quantum mechanics," *McGraw-Hill (1964)*

[12] J. D. Bjorken and S. D. Drell, "Relativistic quantum fields," *McGraw-Hill (1965)*

[13] S. P. Martin, "A Supersymmetry primer," Adv. Ser. Direct. High Energy Phys. **21**, 1 (2010) [Adv. Ser. Direct. High Energy Phys. **18**, 1 (1998)] [hep-ph/9709356].

[14] A. Fetter and D. Walecka, "Quantum theory of many-particle systems," *McGraw-Hill (1971)*

[15] C. Itzykson and J. M. Drouffe, "STATISTICAL FIELD THEORY. VOL. 1: FROM BROWNIAN MOTION TO RENORMALIZATION AND LATTICE GAUGE THEORY," *Cambridge, UK: Univ. Pr. (1989)*

[16] C. Itzykson and J. M. Drouffe, "STATISTICAL FIELD THEORY. VOL. 2: STRONG COUPLING, MONTE CARLO METHODS, CONFORMAL FIELD THEORY, AND RANDOM SYSTEMS," *Cambridge, UK: Univ. Pr. (1989) 405-810*

[17] J. L. Cardy, "Scaling And Renormalization In Statistical Physics," *Cambridge, UK: Univ. Pr. (1996). (Cambridge lecture notes in physics: 3)*

[18] D. J. Amit, "Field Theory, The Renormalization Group, And Critical Phenomena," *Singapore, Singapore: World Scientific (1984)*

[19] S. K. Ma, "Modern theory of critical phenomena," *Benjamin (1976)*

[20] M. Kardar, "Statistical physics of fields," *Cambridge (2007).*

[21] K. G. Wilson and J. B. Kogut, "The Renormalization group and the epsilon expansion," Phys. Rept. **12**, 75 (1974).

[22] A. M. Polyakov, "GAUGE FIELDS AND STRINGS," *CHUR, SWITZERLAND: HARWOOD (1987) 301 P. (CONTEMPORARY CONCEPTS IN PHYSICS, 3)*

[23] T. DeGrand and C. E. Detar, "Lattice methods for quantum chromodynamics," *New Jersey, USA: World Scientific (2006) 345 p*

[24] C. J. Pethick and H. Smith, "Bose - Einstein condensation in dilute gases," *Cambridge (2008)*

[25] Rodney Loudon, "The Quantum Theory of Light"

[26] Marlan O. Scully and M. Suhail Zubairy, "Quantum Optics"

[27] C. Gardiner and P. Zoller, "Quantum Noise"

[28] D. J. Gross, "The triumph and limitations of quantum field theory," arXiv:hep-th/9704139.

[29] F. Wilczek, "Quantum field theory," Rev. Mod. Phys. **71**, S85 (1999) [arXiv:hep-th/9803075].

[30] G. 't Hooft, "Gauge Theory And Renormalization," arXiv:hep-th/9410038.

[31] S. Weinberg, "What is quantum field theory, and what did we think it was?," arXiv:hep-th/9702027.

[32] B. L. Ioffe, "Landau's theoretical minimum, Landau's Seminar, ITEP in the beginning of the 1950's," arXiv:hep-ph/0204295.

[33] G. P. Lepage, "What is Renormalization?," arXiv:hep-ph/0506330.

[34] G. P. Lepage, "How to renormalize the Schrodinger equation," arXiv:nucl-th/9706029.

[35] J. Polchinski, "Effective field theory and the Fermi surface," arXiv:hep-th/9210046.

[36] R. Shenkar, "Renormalization group approach to interacting fermions," Rev. Mod. Phys. **66**, 129 (1994).

[37] J. D. Wells, "When effective theories predict: The Inevitability of Mercury's anomalous perihelion precession," arXiv:1106.1568 [physics.hist-ph].

Chapter 3

An overview of the course

This course is intended to give you an introduction to quantum field theory. So – what is a quantum field theory? A field theory is a dynamical system with degrees of freedom defined at every space-time point. In a quantum field the degrees of freedom are quantum variables. Quantum fields are interesting and important for two big reasons:

First, three of the four fundamental forces in Nature can certainly be described in terms of quantum fields: electromagnetism (quantum electrodynamics or QED), the weak interactions (which combine with QED to give electroweak theory) and the strong interactions (quantum chromodynamics or QCD). These three interactions form the Standard Model of elementary particle physics. In "grand unification" models these three forces are also combined in some way. Gravity is a classical field theory (general relativity); quantum gravity is not part of the course and so we'll move on.

In fact, all four of these theories form a sub-class of quantum field theories called "gauge theories." They all possess hidden local symmetries or invariances.

Second, problems in physics are often many body problems. Quantum field theories are quantum systems where the number of degrees of freedom is very large. If we are interested in systems with a large number of degrees of freedom, the techniques of quantum field theory allow us to study them in the most efficient way.

Often, complicated physical systems can be modeled from the start in terms of "effective field theories" and analyzing their behavior takes us back to the technical aspects of this course.

In a quantum field theory, the concept of a "particle" is a derived quantity. The field is the fundamental object. Quantum field theory allows a natural description of particle creation and annihilation, as is seen in Nature. There is no particular a priori reason for particle number to be conserved. Conserved quantities are built of the field variables. They are currents and charges and their existence is related to symmetries in the underlying field Lagrangian. (Charge, not electron number, is conserved.)

The idea of particle creation and annihilation is also useful in situations where the underlying particle number (for example, the number of electrons in an atomic system) is conserved. Often it is more convenient to think about scattering in terms of objects which are created and destroyed, rather than following the particle itself through space-time. An example is the use of particles and holes in semiconductors in lieu of explicitly considering many electrons in filled and unfilled bands.

Because of particle creation and annihilation, relativistic quantum mechanics – replacing the Schr̈odinger equation by the Dirac equation or the Klein-Gordon equation – is an incomplete description of dynamics. Relativistic quantum systems are quantum field theories.

Quantum field theory for nonrelativistic systems has much in common with relativistic quantum field theory. However, Lorentz invariance puts strong constraints on relativistic systems, which are not present in the nonrelativistic case. An immediate example is that the symmetry group structure of Lorentz transformations constrains the intrinsic spin of a particle, so that only spin 0, $\frac{1}{2}$, 1, ... are allowed. Relativity also constrains dynamics. Imagine the process of Compton scattering, $\gamma + e^- \to \gamma + e^-$. There are two ways the process can occur (in second order perturbation theory): emission can follow absorption, or absorption can follow emission. (See Fig. 3.1.) In a nonrelativistic system, the intermediate states for the two time orderings, $|i_1\rangle$ and $|i_2\rangle$, can be very different. But now think relativistically: if the two spacetime events (absorption and emission) are separated by a spacelike interval, the earlier time in one frame can be the later time in another frame. This says $|i_1\rangle$ and $|i_2\rangle$ must be related. If we label the particles going into each interaction as in the figure, in a frame where $t_1 < t_2$, we'd say particle A converts to particle B which is later absorbed by particle C. But in a frame where $t_1 > t_2$, we would say that C emits a $\bar{B}$ which is later absorbed by A. It's easy to convince yourself the B and $\bar{B}$ have "opposite" quantum numbers. In fact, we will see that they are antiparticles of each other. (The existence of antiparticles has nothing to do with the presence of negative energy solutions to the Dirac equation; it is a general feature of relativistic quantum field theories.)

When we actually get down to business, you'll see that there are two ways to formulate quantum field theories, with Hamiltonians or with path integrals. Hamiltonian methods for quantum field theory are pretty much just a specialized version of what you already know from quantum mechanics. There is a Hilbert space, states (which are nearly always number states, $|n_{k_1}, n_{k_2}, \ldots\rangle$ with the n's integer-valued), and there are operators (the fields) acting on the states. A new item is that a field operator $\phi(x,t)$ combines c-number spatial dependence with quantum operators that typically raise and lower the n's in the states. These are the $a^\dagger(k)$ and $a(k)$ in a field operator

$$\phi(x,t) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}[e^{i(\vec{k}\cdot\vec{x}-\omega t)}a(k) + e^{-i(\vec{k}\cdot\vec{x}-\omega t)}a^\dagger(k)]. \tag{3.1}$$

Typically the "k" in $|n_k \ldots\rangle$ is continuous. Time is both "what the system evolves in" –

$$i\hbar\frac{\partial\,|\psi\rangle}{\partial t} = \hat{H}\,|\psi\rangle \tag{3.2}$$

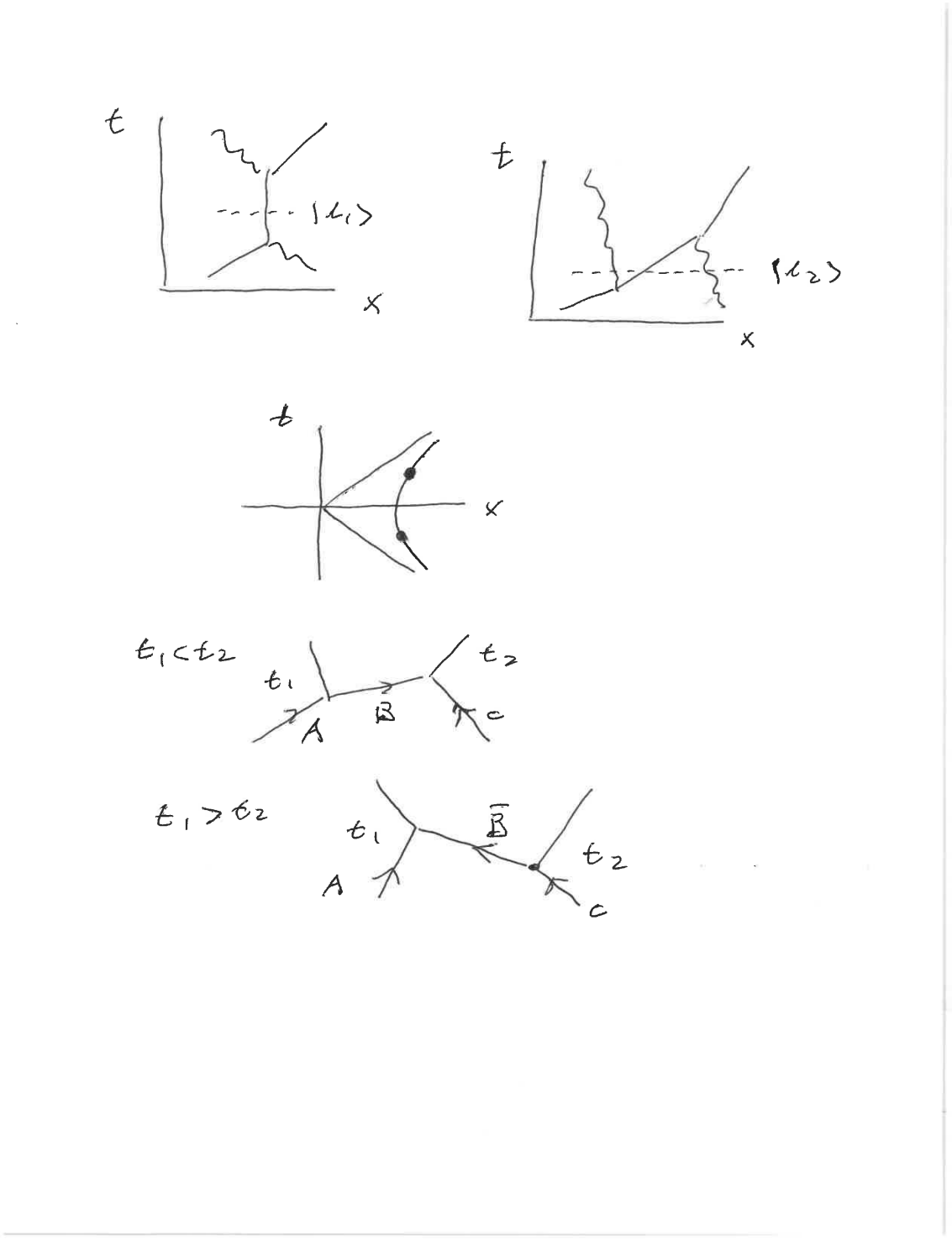

Figure 3.1: The two time orderings for Compton scattering, moving around under a Lorentz transformation.

and time is also a component of the space-time four vector x_μ.

Path integrals for quantum field theories look superficially like partition function in statistical mechanics. The idea is that there is a field variable $\phi(x,t)$ at every space-time point. For a bosonic system, these are classical variables. Everything begins with the "generating functional for Green's functions," which is basically the partition function, which itself depends on the classical Lagrange density,

$$Z = \int \prod_{x,t} d\phi(x,t) \exp(\frac{i}{\hbar} \int d^3x dt \mathcal{L}(\phi, \partial_\mu \phi)). \tag{3.3}$$

Of course, one can derive one formalism from the other.

In either case, the (derived) concept of a particle makes sense only if its persistence probability is large and the probability of its disappearance is small. In that case, we could treat the interactions as small corrections, and describe them using perturbation theory. This leads us (either from Hamiltonians or from path integrals) to the funny pictures you see in every quantum field theory book, Feynman diagrams. These are pictorial representations of Green's functions. Each part of the diagram (vertex or line) corresponds to a particular mathematical expression, linked together in some rigid way to give a perturbative formula for some probability amplitude. A big chunk of the course is learning how to construct – and evaluate – Feynman diagrams to produce scattering amplitudes. At the end, horrible calculations are reduced to exercises with lego blocks.

When we do an actual calculation, we encounter a collection of computational and interpretational problems which are labelled by the word "renormalization." This is a superposition of the following physical problems:

First, suppose you have a theory which you can't solve exactly, so you try to solve it approximately. You write $H = H_0 + \lambda V$ and hope for solutions of the form

$$\begin{aligned} F &= E_0 + \lambda E_1 + \lambda^2 F_2 + \dots \\ \psi &= \psi_0 + \lambda \psi_1 + \lambda^2 \psi_2 + \dots . \end{aligned} \tag{3.4}$$

Experiments measure observables of the full H. Theory is based on H_0, and the properties of H_0 often can't be compared directly to experiment. (An example: an electron coupled to an electromagnetic field, H_0 is the Hamiltonian for a free electron by itself – an electron which does not interact with the electromagnetic field (?!)) H_0 and λV depend on various

"theorist's parameters," such as m_0 and e_0, the mass of our decoupled electron, the coupling of our decoupled electron to the decoupled photon.

To make predictions, we must first use our theory to calculate a set of derived quantities, which can be measured. For example, we compute the amplitude for Rutherford scattering and interpret the coefficient of $1/q^2$ as the physical electric charge e. A second calculation can give us the physical electron mass m. These quantities are functions of the theorist (or "bare") parameters e_0 and m_0: $e(e_0, m_0)$, $m(m_0, e_0)$. Generally $e = e_0 + Ce_0^3 + \ldots$. Once we have made this determination, we can go on to make predictions for additional observables. We would like to express them in terms of e and m, not in terms of e_0 and m_0.

There is an additional complication. Quantum systems become very "rough" at short distance. For example, one can compute the size of fluctuations of the quantum electric field in a sphere of radius "a" (which could be the size of a probe). You would discover that

$$\left\langle 0|(\vec{E}(x+a,t) - \vec{E}(x,t))^2|0 \right\rangle \sim \frac{1}{a^4}. \tag{3.5}$$

As the size of the probe shrinks to zero, fluctuations in the field diverge. Physically, fluctuations arise because you can't probe E at a point without creating photons. This causes a problem because, at least in principle, all the high energy states can contribute to physical processes. The second order formula shows this explicitly:

$$\Delta E = \sum_j \frac{|\langle i|V|j\rangle|^2}{E_i - E_j} \tag{3.6}$$

It's hard to keep track of what is going on when processes at arbitrarily short distance become arbitrarily large.

I could keep going, but at this point in the course, the explanation would be too cryptic. Suffice it to say that there is a rather deep story which ultimately comes down to the statement that in real life everything is an approximation. We will discover the issues as we follow our nose and do what we think are straightforward calculations only to discover that they give us absurd - looking answers. The part you may have heard about already is the notion of a "scale - dependent" coupling, that is, the value of parameters in the theory which are used to make comparisons with experiment are not constants ("coupling constants"); they depend on the energy scale at which the experiment is being done.

Chapter 4

Lagrangians, Hamiltonians, quantum field theories

Like everything else in the study of dynamics, quantum field theory begins with a Lagrangian. Presumably this is because the connections between symmetries and conservation laws, and the consequences of those connections, is most simply achieved through the Lagrangian.

A *symmetry* is a transformation of the coordinates which leaves the Lagrangian unchanged. Symmetries imply conservation laws. (This connection is called Noether's theorem.) That is easiest to see by imagining a Lagrangian $L(q, \dot{q})$ and a transformation $q \to q + \delta q$. If the transformation is a symmetry, the Lagrangian is unchanged,

$$L(q + \delta q, \dot{q}) = L(q, \dot{q}) \tag{4.1}$$

but if the transformation is small, we can write

$$L(q + \delta q, \dot{q}) = L((q, \dot{q}) + \frac{\partial L}{\partial q}\delta q + \dots \tag{4.2}$$

so it must be that

$$\frac{\partial L}{\partial q} = 0. \tag{4.3}$$

The Lagrange equations of motion

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \tag{4.4}$$

then say that $\frac{\partial L}{\partial \dot{q}}$ is time independent. In Hamiltonian dynamics, this is the momentum conjugate to q, and the presence of the symmetry means that it is conserved.

We encode conservation laws using classical Lagrangians. However, if we are interested in quantum dynamics, presumably we need to construct a quantum Hamiltonian. Let us defer the connection between Lagrangians, symmetries, and conservation laws for a while, and just look at how we can get from a classical Lagrangian to a quantum Hamiltonian. The construction of the quantum H from the classical L is called "canonical quantization." It is a bit illogical to go this way –quantum is more fundamental – but the procedure is at least reasonably well defined. (The alternative to what I am describing is to jump immediately to the path integral.) Here is the menu:

1. Begin with $L(q_i, \dot{q}_i)$, $i = 1$ to N for N degrees of freedom

2. Derive the Lagrange equations of motion

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \tag{4.5}$$

3. Pass to the classical Hamiltonian by finding the canonical momenta $p_j = \frac{\partial L}{\partial \dot{q}_j}$

4. Construct the Hamiltonian
$$H = \sum_j p_j \dot{q}_j - L \tag{4.6}$$

5. The p's and q's obey Poisson bracket equations of motion
$$\{H, q_j\} = \frac{\partial H}{\partial p_j} = \dot{q}_j; \qquad \{H, p_j\} = -\frac{\partial H}{\partial q_j} = \dot{p}_j \tag{4.7}$$

6. The quantum system whose quantum Hamiltonian corresponds to H is found by replacing H and all the p's and q's by operators and replacing the Poisson brackets by commutators,
$$[p_i, q_j] = -i\delta_{ij} \tag{4.8}$$
$$[H, q_i] = -i\frac{\partial q_i}{\partial t} \tag{4.9}$$
$$[H, p_i] = -i\frac{\partial p_i}{\partial t} \tag{4.10}$$
and in fact, for any operator O,
$$i[H, O] = \frac{\partial O}{\partial t} \tag{4.11}$$
We are obviously working in Heisenberg basis.

Let's do a series of examples, building up to our first quantum field theory.

4.1 The oscillator

What else? The classical Lagrangian is
$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2 \tag{4.12}$$
and the classical Hamiltonian is
$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \tag{4.13}$$
A useful way to describe even a classical oscillator is to work in terms of the classical variables
$$a = \sqrt{\frac{m\omega}{2}}\left(x + \frac{ip}{m\omega}\right) \qquad a^* = \sqrt{\frac{m\omega}{2}}\left(x - \frac{ip}{m\omega}\right). \tag{4.14}$$

We can rewrite the classical equations of motion for p and q in terms of a and a^*:

$$\dot{a} = -i\omega a; \qquad \dot{a}^* = i\omega a^* \tag{4.15}$$

The classical Hamiltonian can be rewritten as

$$H = \omega a^* a. \tag{4.16}$$

These two equations define a classical oscillator as well as anything else.

We pass to quantum mechanics by interpreting p and q, or a and a^* as operators $\hat{p}$, $\hat{q}$, $\hat{a}$ and $\hat{a}^\dagger$. The Poisson bracket expression involving the classical a's is $\{a, a^*\} = 1$. It means that the operators obey a commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$. The usual passage of steps then gives us the quantum Hamiltonian

$$H = \omega(\hat{a}^\dagger \hat{a} + \frac{1}{2}) \equiv \omega(\hat{N} + \frac{1}{2}) \tag{4.17}$$

where $\hat{N} = \hat{a}^\dagger \hat{a}$ is the number operator. The Heisenberg equations of motion for the operators $\hat{a}$ and $\hat{a}^\dagger$ are identical to the classical equations of motion for a and a^*, Eq. 4.15.

The algebra of the commutation relation between a and $a^\dagger$ plus the form of the Hamiltonian lead immediately to the familiar features of the quantum oscillator: eigenstates of the Hamiltonian are labelled by an integer $n \geq 0$; the operator a acts on these states to lower n by one unit, and $a^\dagger$ acts to raise the integer by one unit. For completeness, let's review:

Obviously, eigenstates of the Hamiltonian will also be eigenstates of N, so let's solve the equation $\hat{N}\,|n\rangle = n\,|n\rangle$. Using the commutation relations

$$[\hat{N}, \hat{a}] = [\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a}^\dagger [\hat{a}, \hat{a}] + [\hat{a}^\dagger, \hat{a}]\hat{a} = -\hat{a} \tag{4.18}$$

and

$$[\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger. \tag{4.19}$$

What do raising and lowering operators do to eigenstates of $\hat{N}$? Consider

$$\begin{aligned} \hat{N}\hat{a}^\dagger\,|n\rangle &= ([\hat{N}, \hat{a}^\dagger] + \hat{a}^\dagger \hat{N})\,|n\rangle \\ &= (\hat{a}^\dagger + \hat{a}^\dagger \hat{N})\,|n\rangle \\ &= (n+1)\hat{a}^\dagger\,|n\rangle \end{aligned} \tag{4.20}$$

This shows us that $a^\dagger |n\rangle$ is an un-normalized eigenstate of $\hat{N}$ with eigenvalue $n+1$,

$$\hat{a}^\dagger |n\rangle = C_{n+1}^{(+)} |n+1\rangle . \tag{4.21}$$

The lowering operator behaves similarly:

$$\hat{N}\hat{a} |n\rangle = ([\hat{N}, \hat{a}] + \hat{a}\hat{N}) |n\rangle = (n-1)\hat{a} |n\rangle \tag{4.22}$$

so

$$\hat{a} |n\rangle = C_{n-1}^{(-)} |n-1\rangle . \tag{4.23}$$

To find the constants, consider the norm of the state $\hat{a} |n\rangle$:

$$\begin{aligned} \langle n|\hat{a}^\dagger\hat{a}|n\rangle &= (\langle n| \hat{a}^\dagger)(\hat{a} |n\rangle) = |C_{n-1}^{(-)}|^2 \langle n-1|n-1\rangle \\ &= n \langle n|n\rangle = n. \end{aligned} \tag{4.24}$$

Therefore, (choosing a phase) the normalization factor is

$$C_{n-1}^{(-)} = \sqrt{n}. \tag{4.25}$$

An analogous calculation for the raising operator gives

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle \tag{4.26}$$

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle . \tag{4.27}$$

Observe that we cannot lower the state forever, because the norm of the state (the inner product of bra and ket) must always be greater than 0. From Eq. 4.25, this norm is n. Therefore, the minimum n must then be 0, and all the n's must be integers (otherwise one could lower states to negative values of n).

Familiar stuff! the important ingredients are

1. We started with a classical Lagrangian or Hamiltonian which is quadratic in the degrees of freedom. Note, though, that while we could have begin in the standard way by writing everything in terms of p's and q's, Eqs. 4.15 and 4.16 are an equally valid way to characterize a classical oscillator.

2. The classical variables p, q, or a and a^*, become operators. The algebra of the commutation relation $[a, a^\dagger] = 1$ gives us an energy spectrum labelled by an integer. There are equally spaced energy levels $E = n\omega +$ constant, with $n = 0, 1, 2, \ldots$. The operators a and $a^\dagger$ act on these states to lower and raise the integer by one unit.

.

All of perturbative quantum field theory lurks in these two ingredients.

4.2 Many uncoupled oscillators

This is trivial, just add an index to what we have already done! The classical Hamiltonian is

$$H = \sum_i H_i = \sum_i \frac{p_i^2}{2m_i} + \frac{1}{2} m_i \omega_i^2 \tag{4.28}$$

or

$$H = \sum_i H_i = \sum_i \omega_i a_i^* a_i \tag{4.29}$$

Clearly, we can introduce the linear combinations (classical or quantum, it doesn't matter)

$$a_i = \sqrt{\frac{m_i \omega_i}{2}} \left(\hat{x_i} + \frac{i \hat{p_i}}{m_i \omega_i} \right) \qquad a_i^\dagger = \sqrt{\frac{m_i \omega_i}{2}} \left(\hat{x_i} - \frac{i \hat{p_i}}{m_i \omega_i} \right). \tag{4.30}$$

The quantum commutation relations descend from $[p_i, x_j] = -i\delta_{ij}$, so $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$. The Hamiltonian is

$$H = \sum_i \omega_i [\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2}] \tag{4.31}$$

The individual a' and $a^\dagger$'s obey the same equation of motion as Eq. 4.15.

Energy eigenstates are simply product states

$$|\{n\}\rangle = |n_1, n_2, \ldots\rangle = \prod_i |n_i\rangle \tag{4.32}$$

with

$$H |\{n\}\rangle = \sum_i \omega_i [n_i + \frac{1}{2}] |\{n\}\rangle . \tag{4.33}$$

Note the sum of terms, each is an integer times a characteristic frequency. And of course

$$a_j |\{n\}\rangle = \sqrt{n_j} |n_1, n_2, \ldots, n_j - 1, \ldots\rangle \tag{4.34}$$

and so on.

4.3 Our first quantum field theory

You have now been set up for our first quantum field theory, a set of coupled oscillators, which you can think of as a set of masses and springs along a line, i. e. a one dimensional solid. The Lagrangian is

$$L = \sum_{n=0}^{N-1} \frac{1}{2}\dot{y}_n^2 - \frac{1}{2}\mu^2 y_n^2 - \frac{1}{2a^2}(y_n - y_{n+1})^2 \tag{4.35}$$

where y_n is the displacement of the coordinate of the nth oscillator from its equilibrium position, the second term is the local restoring potential (think of the individual oscillator as a pendulum) and the last term is a coupling between the individual oscillators (imagine a spring connecting successive mass points). The Lagrange equation of motion for each oscillator is

$$\ddot{y}_n = -\mu^2 y_n - \frac{1}{a^2}(2y_n - y_{n-1} - y_{n+1}) \tag{4.36}$$

This is a coupled system, but it can be decoupled by Fourier series; write

$$y_n = \frac{1}{N}\sum_k e^{-ikna} y_k \tag{4.37}$$

where $k = \frac{2\pi}{Na} n_k$ and $n_k = 0$ to $N-1$ if we impose periodic boundary conditions. The inverse relation is

$$y_k = \sum_{n=0}^{N-1} e^{ikna} y_n. \tag{4.38}$$

Noticing that

$$\sum_n e^{ikna} y_{n+1} = e^{-ika} y_k, \tag{4.39}$$

we have the equation of motion

$$\ddot{y}_k + \mu^2 y_k + \frac{1}{a^2}(2 - e^{ika} - e^{-ika}) y_k = 0 \tag{4.40}$$

which for each Fourier mode is just

$$\ddot{y}_k + \omega^2 y_k = 0 \tag{4.41}$$

where

$$\omega_k^2 = \mu^2 + \frac{2}{a^2}(1 - \cos ka). \tag{4.42}$$

That is, viewed in Fourier space, the mass-spring combination is just a set of uncoupled oscillators. We can see this directly by writing L in terms of the y_k's:

$$L = \frac{1}{N^2}\sum_n \sum_{k.k'} e^{iknа} e^{ik'na} [\frac{1}{2}\dot{y}_{k'}\dot{y}_k - \frac{1}{2}\mu^2 y_{k'} y_k - \frac{1}{2a^2} 2(1-\cos ka) y_{k'} y_k] \tag{4.43}$$

The $\sum_n$ gives a $\delta(k+k')$ which eats an N, leaving

$$L = \frac{1}{N}\sum_k \frac{1}{2}\dot{y}_{-k}\dot{y}_k - \frac{1}{2} y_{-k} y_k [\mu^2 + 2(1-\cos ka)]. \tag{4.44}$$

If the y_n's are real, then $y_{-k} = y_k$ and

$$L = \frac{1}{N}\sum_k \frac{1}{2}\dot{y}_k^2 - \frac{1}{2}\omega_k^2 y_k^2 \tag{4.45}$$

which is again a set of uncoupled oscillators.

Clearly, we can perform the same transformation on H. We could define $p_n = \partial L/\partial \dot{y}_n$ in coordinate space or $p_k = \partial L/\partial \dot{y}_k = \dot{y}_k/N$ in Fourier space. With the latter choice

$$H = \frac{1}{N}\sum_k \frac{1}{2} p_k^2 + \frac{1}{2}\omega_k^2 y_k^2 \tag{4.46}$$

which is yet again a set of uncoupled oscillators. The momentum variables p_n and p_k are related by Fourier transform, of course.

Before looking at the quantization problem, let's pass to the limit of an infinite number of degrees of freedom. We can do that by removing the lattice spacing a, $y_n \to y(x_n = na)$. The nearest neighbor coupling is

$$\frac{1}{2}(\frac{y_n - y_{n+1}}{a})^2 = \frac{1}{2}[\frac{y(x,t) - (y(x,t) + a\frac{\partial y}{\partial x})}{a}]^2 = \frac{1}{2}(\frac{\partial y}{\partial x})^2 \tag{4.47}$$

And with $\sum_n = \sum_n \Delta n = \sum_x \frac{\Delta x}{a} = \frac{1}{a}\int dx$,

$$L = \frac{1}{a}\int dx [\frac{1}{2}\dot{y}(x,t)^2 - \frac{1}{2}\mu^2 y(x,t)^2 - \frac{1}{2}(\frac{\partial y(x,t)}{\partial x})^2]. \tag{4.48}$$

Finally define rescaled field variables

$$\phi(x,t) = \frac{y(x,t)}{\sqrt{a}} \tag{4.49}$$

and we have the Lagrangian in its continuum form

$$L = \int dx [\frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\mu^2\phi^2 - \frac{1}{2}(\frac{\partial \phi}{\partial x})^2]. \tag{4.50}$$

Obviously, a three dimensional analog is

$$L = \int d^3x [\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}\mu^2\phi^2] = \int d^3x \mathcal{L} \tag{4.51}$$

where

$$(\partial_\mu \phi)^2 = (\frac{\partial \phi}{\partial t})^2 - (\vec{\nabla}\phi)^2 \tag{4.52}$$

and $\mathcal{L}$ is called the Lagrange density.

This is the Lagrangian for a classical field. With our choice of scale factors, the derivative term ∂_μ transforms as a four-vector and $(\partial_\mu \phi)^2$ is Lorentz invariant. The field $\phi(x,t)$ will be a scalar field (scalar in the Lorentz transformation sense) if we argue that under a Lorentz transform, only its arguments (x,t) change. (More on this later.)

We can play the $a \to 0$ game on the Lagrange equations of motion, too. They are (see Goldstein ....)

$$\partial_\mu [\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}] - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{4.53}$$

or

$$\Box\phi + \mu^2\phi = 0 \tag{4.54}$$

where I've introduced the $(c = 1)$ shorthand

$$\Box = \frac{\partial^2}{\partial t^2} - (\vec{\nabla})^2. \tag{4.55}$$

This is called the Klein - Gordon equation; it is the continuum analog of the equation of motion for a collection of masses and springs.

So far, this is all classical – what is the quantum analog of this system? Let's carry out canonical quantization in parallel for the lattice and continuum theories. Lagrangians:

Lattice:

$$L = \sum_n \mathcal{L}_n; \qquad \mathcal{L}_n = \frac{1}{2}\dot{\phi}_n^2 - \frac{1}{2}\mu^2\phi_n^2 - \frac{1}{2a^2}(\phi_n - \phi_{n+1})^2 \tag{4.56}$$

and $p_n = \frac{\partial L}{\partial \dot{\phi}_n} = \dot{\phi}_n$.

Continuum:

$$L = \int d^3x \mathcal{L}; \qquad \mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}\mu^2\phi^2 \tag{4.57}$$

with

$$\pi(x,t) = \frac{\partial L}{\partial(\partial_0 \phi)} = \dot{\phi}(x,t). \tag{4.58}$$

Hamiltonians:

Lattice:

$$H = \sum_n \mathcal{H}_n; \qquad \mathcal{H}_n = p_n\dot{\phi}_n - \mathcal{L}_n = \frac{1}{2}p_n^2 + \frac{1}{2}\mu^2\phi_n^2 + \frac{1}{2a^2}(\phi_n - \phi_{n+1})^2 \tag{4.59}$$

Continuum:

$$H = \int d^3x\mathcal{H}; \qquad \mathcal{H}(x,t) = \frac{1}{2}\pi(x,t)^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}\mu^2\phi^2 \tag{4.60}$$

Quantization:

Lattice: $[p_i, \phi_j] = -i\delta_{ij}$; continuum $[\pi(x,t], \phi(x',t)] = -i\delta^3(x-x')$.

Hmm ..., this is all true but perhaps working with $\phi(x,t)$ and $\pi(x,t)$ is not maximally efficient? Perhaps it might be better to begin in a basis in which the oscillators are uncoupled. Return to the continuum Lagrange equation of motion

$$(\frac{\partial^2}{\partial t^2} - \nabla^2)\phi(x,t) + \mu^2\phi(x,t) = 0. \tag{4.61}$$

We write a $\phi(x,t)$ in terms of a superposition of normal modes (with $\omega_k = \sqrt{k^2+\mu^2}$)

$$\phi(x,t) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}[a(k,t)e^{-ik\cdot x} + a^*(k,t)e^{-ik\cdot x}]. \tag{4.62}$$

The integration measure seems arbitrary, but it will be convenient to adopt it later. The quantities $a(k,d)$ and $a^*(k,t)$ are classical Fourier coefficients. Hitting ϕ in Eq. 4.62 with Eq. 4.61 gives an equation of motion for the Fourier coefficients

$$\frac{d^2a(k,t)}{dt^2} + \omega_k^2 a(k,t) = 0. \tag{4.63}$$

The $a(k,t)$'s are obviously oscillator variables.

We have been writing down first order equations for the oscillator and this is a second order equation. But there is a "2" we have ignored: if ϕ is real, then keeping an independent a and an a^* for each k mode is an overcounting by a factor of two. We can take care of this over-counting if we clip out half the k modes, perhaps by restricting $k_z > 0$ in the sum.

The solution to Eq. 4.63 is, of course,

$$a(k,t) = a(k,0)^{(1)}e^{-i\omega_k t} + a(k,0)^{(2)}e^{i\omega_k t}. \tag{4.64}$$

The annoying factor of two now returns, usefully. We can remove the restriction $k_z > 0$ if we define $a(k) = a(k,0)^{(1)}$ if $k_z > 0$ and $a(k,0)^{(2)}$ if $k_z < 0$. This amounts to just keeping the solution

$$a(k,t) = a(k,0)e^{-i\omega_k t} \tag{4.65}$$

or

$$\frac{da(k,t)}{dt} = -i\omega_k a(k,t). \tag{4.66}$$

This gives us a set of classical variables with the desired first-order equation of motion.

Incidentally, we'll see later that a fully relativistic theory requires the presence of both a and a^*.

To construct the Hamiltonian, we need the momentum conjugate to $\phi(x,t)$. It is

$$\pi(x,t) = \dot{\phi}(x,t) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}[i\omega_k][a(k,t)e^{-ik\cdot x} - a^*(k,t)e^{-ik\cdot x}]. \tag{4.67}$$

Now we make our system quantum by thinking of ϕ and π as operators and imposing the quantization condition that $[\pi(x,t],\phi(x',t)] = -i\delta^3(x-x')$. We impose the condition by replacing the classical Fourier coefficients by operators $\hat{a}(k)$ and $\hat{a}^\dagger(k)$. The commutator is

$$\begin{aligned} -i\delta^3(x-x') &= \int \frac{d^3k d^3k'}{(2\pi)^3\sqrt{4\omega_k\omega_{k'}}}[i\omega_k] \\ &\quad ([a^\dagger(k,t), a^\dagger(k',t)] \times \ldots \\ &\quad +[a(k,t), a(k',t)] \times \ldots \\ &\quad +2[a^\dagger(k,t), a(k',t)]e^{ik\cdot x}e^{-ik'\cdot x'}). \end{aligned} \tag{4.68}$$

This says (yet again!) that in order to enforce the quantization condition $[\pi(x,t],\phi(x',t)] = -i\delta^3(x-x')$, the a and $a^\dagger$ operators must obey the commutation relations

$$[a^\dagger(k,t), a^\dagger(k',t)] = 0 \tag{4.69}$$

$$[a(k,t), a(k',t)] = 0 \tag{4.70}$$

$$[a(k.t), a^\dagger(k',t)] = \delta^3(k-k') \tag{4.71}$$

for then (just substituting back)

$$\begin{aligned} -i\delta^3(x-x') &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}[-2i\omega_k]e^{ik(x-x')} \\ &= -i\int \frac{d^3k}{(2\pi)^3}e^{ik(x-x')}. \end{aligned} \tag{4.72}$$

The commutation relations are identical to what we saw for the simple harmonic oscillator. The rest of the oscillator story follows immediately: energy eigenstates, which are eigenstates of the number operator, are labelled by a set of integers, one for each value of k. The classical Fourier coefficients become operators which raise and lower the n_k's, just like the raising and lowering operators of the oscillator. Let's fill in the details.

The Hamiltonian operator is

$$H = \int d^3x [\frac{1}{2}\pi(x,t)^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}\mu^2\phi^2 \tag{4.73}$$

which we can also evaluate in terms of creation and annihilation operators. A tedious calculation gives

$$H = \int d^3k \frac{1}{2}\omega_k [a^\dagger(k)a(k) + a(k)a^\dagger(k)]. \tag{4.74}$$

This makes perfect sense! A set of many harmonic oscillators would have $H = \sum_j \omega_j (a_j^\dagger a_j + \frac{1}{2})$ and $\frac{1}{2} = \frac{1}{2}(a_j a_j^\dagger - a_j^\dagger a_j)$, so $H = \sum_j \frac{1}{2}\omega_j (a_j^\dagger a_j + a_j a_j^\dagger)$.

Our theory is diagonal in k space, so our states will be number states labeled by k.

$$|\psi\rangle \sim |n_{k_1}, n_{k_2}, \ldots\rangle \,. \tag{4.75}$$

Before we continue, let's deal with an annoyance (and stick with sums rather than integrals). Obviously

$$H = \sum_j \omega_j (a^\dagger a_j + \frac{1}{2}). \tag{4.76}$$

The $\frac{1}{2}$ is the annoyance. There is a zero point energy for each oscillator. This energy is proportional to the number of oscillators, and in this case we have an infinite number of oscillators, so we have a infinite contribution to the zero point energy,

$$H\,|0,0,0,0,0\ldots\rangle = (\sum_j \frac{1}{2}\omega_j)\,|0,0,0,0,0\ldots\rangle \,. \tag{4.77}$$

However, the gap, the splitting of any state from the ground state $|0,0,0,0,0\ldots\rangle$ is finite. And gaps are all we ever measure in quantum mechanics. So let's "define the zero point energy away." (There is an issue with quantum gravity at the point which I will ignore.) A practical Hamiltonian is

$$H = \sum_j \omega_j a_j^\dagger a_j \to \int d^3k \omega_k a^\dagger(k) a(k). \tag{4.78}$$

There is a formalism to deal with practicality: define "normal ordering" for any operator O as taking the operator, moving all the $a^\dagger$'s in the operator to the left and all a's to the right, and discarding all resulting c-numbers which arise from commutation. The label for a normal ordered operator is $:O:$ with colons. Then

$$:H:=\int d^3k\omega_k a^\dagger(k)a(k). \tag{4.79}$$

I wil drop the :: in what follows.

It's all downhill from here. Each individual k state forming the product state is an eigenstate of energy $n_k\omega_k$ with $\omega_k = \sqrt{k^2+\mu^2}$. The integer n_k counts the number of particles in the state. As in the case of the coupled oscillator, a complete specification of the state is as a set of integers, one for every k state.

Let's look at the lowest lying states in the spectrum:

1. $|0\rangle = |0_{k_1}, 0_{k_2}, 0_{k_3}, \ldots\rangle$. This state is called the "vacuum." It has zero energy. Clearly it is the state with zero particle content. If we try to remove a particle from it, we annihilate the state: $a(k)\,|0\rangle = 0$ for any k. It is also clearly the analog of the ground state for a collection of oscillators.

2. One particle states are $a^\dagger(k)\,|0\rangle = C\,|\vec{k}\rangle$. $:H:(a^\dagger(k)\,|0\rangle) = \omega_k(a^\dagger(k)\,|0\rangle)$, so these states have energy ω_k, the energy-momentum dispersion relation for a massive relativistic particle with mass μ and momentum k.

3. Two particle states can be $a^\dagger(k_1)a^\dagger(k_2)\,|0\rangle$ with $E=\omega_{k_1}+\omega_{k_2}$ or $1/\sqrt{2}(a^\dagger(k))^2\,|0\rangle$ with energy $2\omega_k$

and so on.

Congratulations, you have just solved your first quantum field theory. This procedure can be repeated over and over. Let's summarize. We began with classical fields and a classical Lagrange density $\mathcal{L}(\phi, \partial_\mu\phi)$. Notice that $\mathcal{L}$ only depends on ϕ and its derivatives; its space-time dependence is all through ϕ and derivatives of ϕ. (We haven't really discussed where Lagrangians come from, yet.) We then found the classical field momentum $\pi(x,t)$ and the classical Hamiltonian density. Under quantization, $\phi(x,t)$ and $\pi(x,t)$ become noncommuting quantum operators acting on an abstract Hilbert space.

To go farther, it's better to get at particle creation and annihilation operators $a^\dagger(k)$ and $a(k)$ in a round-about way. Write solutions to the classical field equations associated with

the Lagrangian in terms of its normal modes,

$$D\phi(x,t) = 0 \rightarrow \phi(x,t) = \sum_n c_n(x) a_n(t) + c.c., \tag{4.80}$$

that is, the $c_n(x)$'s are solutions to the differential equation and $a_n(t)$ and $a_n^*(t)$ are classical expansion coefficients which obey oscillator-like equations of motion. Then enforce the commutation relations by regarding the a's and a^*'s as operators. You discover that $[a_n, a_m^\dagger] \propto \delta_{mn}$, that is, the a's and a^*'s are number operators. The algebra immediately gives equally spaced energy levels; the spacing between the levels is the energy-momentum relation for the excitations of the system and the integer counts particle number.

Our description has mixture of particle and wave attributes: Waves – from the $c_n(x,t)$'s, because different modes can interfere spatially. And the energy-momentum relation for the excitations comes from the classical wave equation, too. Particles –in that energies come quantized in integers, $E_{k,n_k} = n_k \omega_k$.

Because the allowed values of any n_j are integers, $n_j = 0,, 1, 2, \ldots$ with no restrictions on an upper limit, we have a theory of bosons. (Fermions, with their restriction to $n_j = 0$ or 1, will be dealt with shortly.) Because $\omega_k = \sqrt{k^2 + \mu^2}$, our excitations have the dispersion relation of relativistic particles of rest mass μ.

Everything else is just technical details.

But wait –we don't yet have interactions. However, it is easy to see how to build an interacting quantum field theory. Only $\mathcal{L}$'s which are bilinear in the field variables can be solved as we have done, to produce a tower of noninteracting states. Anything else,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}\phi^2 + \mathcal{L}_I \tag{4.81}$$

will give an interacting theory, a system in which the free particles associated with the quadratic part of $\mathcal{L}$ are coupled together by $\mathcal{L}_I$. And $\mathcal{L}_I$, like all Lagrangians, will be built of ϕ's and derivative terms. Typical $\mathcal{L}_I$'s will be polynomials in the ϕ's, like $\phi(x,t)^4$. Because $\phi(x,t)$ has an $a^\dagger$ piece and an a piece, it both creates and annihilates particles. So the interactions will involve the creation and annihilation of multiple particles at a single point in space-time.

Chapter 5

Quantum field theory for many-body problems

5.1 Particles to waves to particles again

Recall the ingredients of our sample quantum field theory: a classical wave equation and classical Fourier coefficients which are reinterpreted by quantization to become a set of quantum creation and annihilation operators. This resulted in a system with integer quantum numbers, easily interpreted as a system of an integer number of particles.

Let's think about some analog situations in Nature. In particular, consider a collection of non-interacting particles. If they have a set of energy levels ϵ_n, then the energy of a many-particle system is

$$E = \sum_{j=1}^{n} \epsilon_j n_j \tag{5.1}$$

where n_j is the number of particles in energy state j. Note the resemblance to the field theoretic formula.

Next, think more closely about conservation laws. Generally, quantities like charge are conserved, but the number of particles might not be conserved. For example, in a gas of electrons and positrons, the number of electrons or positrons is not conserved because they can annihilate. Only the difference in their number, the net charge, is conserved. Also, it is often useful to think about particle creation and annihilation in a more extended context: imagine the situation when an electron in an atom leaves the 1S state to go to the 2P state: $\Delta n_{1S} = -1$, $\Delta n_{2P} = +1$. The total particle number is conserved, the number of particles in any given state is not.

Finally, why are all electrons alike? A crazy but consistent answer: they are all states created by some field operator $\phi(x,t)$, and there is only one operator $\phi(x,t)$.

A quantum field theory formulation of many body problems might have some useful features. Let's formulate quantum field theory for Schrödinger particles. This will (hopefully) give us a formalism which is simpler than working directly with a many-body Schrödinger equation and wave function.

The idea is to imagine thinking of the Schrödinger equation as a differential equation (∇^2 is just a differential operator, nothing more)

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \psi + V(x)\psi. \tag{5.2}$$

Let $\psi_n(x)$ solve

$$[-\frac{\hbar^2}{2m}\nabla^2 + V(x)]\psi_n(x) = E_n \psi_n(x) \tag{5.3}$$

and write the general solution to Eq. 5.2 as

$$\psi(x,t) = \sum_n b_n(t)\psi_n(x) \tag{5.4}$$

where

$$i\hbar\frac{\partial b_n}{\partial t} = E_n b_n \tag{5.5}$$

or

$$\frac{\partial b_n}{\partial t} = -i\omega_n b_n. \tag{5.6}$$

This is just a "classical" differential equation, but we appear to have two thirds (the wave equation, a time dependent b) of the field theoretic ingredients in hand. If we could write

$$H = \sum_n \epsilon_n b_n^* b_n \tag{5.7}$$

we could re-interpret the expansion coefficient as an operator in an abstract Hilbert space. So we need to find an H which can be used to give Eq. 5.6 as an equation of motion. A natural guess for H is

$$H = \int d^3x \psi^*(x,t)[-\frac{\hbar^2}{2m}\nabla^2 + V(x)]\psi(x,t) \tag{5.8}$$

which reduces to

$$H = \sum_n E_n b_n^* b_n, \tag{5.9}$$

just from orthogonality. Now boldly interpret b_n as an operator with commutation relations

$$[b_n, b_m^\dagger] = \delta_{nm} \qquad [b_n, b_m] = 0 \qquad [b_n^\dagger, b_m^\dagger] = 0 \tag{5.10}$$

(so that different levels n don't interact). The Heisenberg equation of motion is

$$i\hbar\frac{\partial b_n}{\partial t} = [b_n, H] = \sum_j [b_n, b_j^\dagger b_j]E_j = E_n b_n. \tag{5.11}$$

We have a theory of "field quanta of the Schrödinger equation." Our Hilbert space is a number - operator - diagonal space

$$|\{n\}\rangle = |n_1, n_2, n_3, \ldots\rangle \tag{5.12}$$

where the subscript labels levels of the Schrödinger differential operator. The quantity

$$\psi(x,t) = \sum_n \psi_n(x) b_n(t) \tag{5.13}$$

is an operator, a sum of a product of a c-number function of the spatial coordinates times an (annihilation) operator acting on the space of states. In contrast to photons, for Schrödinger particles, different operators create or annihilate particles:

$$\psi^\dagger(x,t) = \sum_n \psi_n^*(x) b_n^\dagger(t) \tag{5.14}$$

Note that the b's act on states like harmonic oscillator operators: $b_n^\dagger b_n \left|n\right\rangle = n \left|n\right\rangle$. The quantity n is an integer $n = 0,\ 1,\ 2, \ldots$. Also $b_n^\dagger \left|n\right\rangle = \sqrt{n+1}\left|n+1\right\rangle$, $b_n \left|n\right\rangle = \sqrt{n}\left|n-1\right\rangle$. This follows immediately from the commutation relations. As we can have any number of particles in the state, we have just invented a quantum field theory for bosons.

What about fermions? We want to keep

$$H = \sum_n E_n b_n^\dagger b_n, \tag{5.15}$$

but we want to restrict n to be only zero or unity. It turns out that we can do that – and pick up the minus signs needed in many - fermion wave functions – if we replace the commutation relation for the b's by an anti-commutation relation

$$\begin{aligned} [b_n, b_m^\dagger]_+ &= b_n b_m^\dagger + b_m^\dagger b_n = \delta_{nm} \\ [b_n, b_m]_+ &= b_n b_m + b_m b_n = 0 \\ [b_n^\dagger, b_m^\dagger]_+ &= b_n^\dagger b_m^\dagger + b_m^\dagger b_n^\dagger = 0 \end{aligned} \tag{5.16}$$

Then

$$\begin{aligned} i\hbar \frac{\partial b_n}{\partial t} &= \sum_m E_m [b_n, b_m^\dagger b_m] \\ &= \sum_m E_m (b_n b_m^\dagger b_m - b_m^\dagger b_m b_n) \\ &= \sum_m E_m ((\delta_{nm} - b_m^\dagger b_n) b_m - b_m^\dagger b_m b_n) \\ &= \sum_m E_m (\delta_{nm} b_m + b_m^\dagger b_m b_n - b_m^\dagger b_m b_n) \\ &= E_n b_n. \end{aligned} \tag{5.17}$$

The desired equation of motion is preserved.

Now we want eigenstates of $b_n^\dagger b_n$. Note that $(b_n^\dagger b_n)(b_n^\dagger b_n) = b_n^\dagger(1 - b_n^\dagger b_n)b_n = b_n^\dagger b_n - b_n^\dagger b_n^\dagger b_n b_n$. The second term is zero applied to any state because $b_n b_n = -b_n b_n$. So if $|\lambda\rangle$ is an eigenstate of $b_n^\dagger b_n$,

$$\begin{aligned} b_n^\dagger b_n |\lambda\rangle &= \lambda |\lambda\rangle \\ (b_n^\dagger b_n)(b_n^\dagger b_n) |\lambda\rangle &= \lambda^2 |\lambda\rangle = \lambda |\lambda\rangle \end{aligned} \tag{5.18}$$

or $\lambda^2 = \lambda$, meaning that $\lambda = 0, 1$. Only zero or one particle can occupy a state.

Finally, we need matrix elements of b_n and $b_n^\dagger$. We have

$$\begin{aligned} b_n^\dagger b^n |N_n\rangle &= N_n |N_n\rangle ; \qquad N_n = 0, 1 \\ (b_n^\dagger b_n) b_n^\dagger |N_n\rangle &= b_n^\dagger (1 - b_n^\dagger b_n) |N_n\rangle = (1 - N_n) |N_n\rangle \end{aligned} \tag{5.19}$$

so $b_n^\dagger |N_n\rangle = C_n |1 - N_n\rangle$. To find C_n, square this:

$$|C_n|^2 = \langle N_n | b_n b_n^\dagger | N_n \rangle = \langle N_n | (1 - b_n^\dagger b_n | N_n \rangle = 1 - N_n \tag{5.20}$$

Thus $C_n = \theta_n \sqrt{1 - N_n}$ where θ_n is a phase factor. Similarly,

$$b_n |N_n\rangle = \theta_n \sqrt{N_n} |1 - N_n\rangle \tag{5.21}$$

For most applications, the value of the phase factor is irrelevant.

Although the formalism looks different, everything is the same as in ordinary quantum mechanics. Let's see how that works out. Note that

$$\psi(x,t) = \sum_n b_n(t) \psi_n(x) \tag{5.22}$$

is again an operator which annihilates particles in states n. Let's look at the (anti)commutation relations for ψ and $\psi^\dagger$:

$$[\psi(x,t)\psi^\dagger(x',t)]_\pm = \sum_n \sum_m \psi_n(x)\psi_m(x')[b_n, b_m^\dagger]_\pm = \delta^3(x - x') \tag{5.23}$$

from a combination of the commutation relations for the b's and completeness for the ψ's. Similarly, $[\psi(x,t)\psi(x',t)]_\pm = [\psi^\dagger(x,t)\psi^\dagger(x',t)]_\pm = 0$. You can show that when the Hamiltonian is

$$H = \int d^3x \psi^\dagger(x,t)[-\frac{\hbar^2}{2m}\nabla^2 + V]\psi(x,t), \tag{5.24}$$

the Heisenberg operator equation

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = [\psi(x,t), H] \tag{5.25}$$

gives the time-dependent Schrödinger equation for ψ. The number density operator is $n(x,t) = \psi^\dagger(x,t)\psi(x,t)$ and the total particle number is

$$N = \int d^3x n(x,t). \tag{5.26}$$

We can build states beginning with the "vacuum state" $|0\rangle$, which has no particles in it. It obeys the relation $\psi(x,t)|0\rangle = 0$. Following this line, $\psi^\dagger(x,t)|0\rangle$ should be a state with one particle at x. Is it? Look at the application of the number density operator

$$\begin{aligned} n(x',t)\psi^\dagger(x,t)|0\rangle &= \psi^\dagger(x',t)\psi(x',t)\psi^\dagger(x,t)|0\rangle \\ &= \psi^\dagger(x',t)(\delta^3(x'-x) \mp \psi^\dagger(x,t)\psi(x',t))|0\rangle \\ &= \delta^3(x'-x)\psi^\dagger(x',t))|0\rangle\,. \end{aligned} \tag{5.27}$$

The state is an eigenfunction of the number density operator with an eigenvalue which is a delta function. This is the mathematical equivalent of the statement that there is a particle located at x. Integrating, we find that the state is an eigenstate of the number operator,

$$N\psi^\dagger(x,t)|0\rangle = \int d^3x' n(x',t)\psi^\dagger(x,t)|0\rangle = \psi^\dagger(x,t)|0\rangle\,, \tag{5.28}$$

so the state does have one particle in it. Similarly

$$\psi^\dagger(x_1,t)\psi^\dagger(x_2,t)|0\rangle \tag{5.29}$$

is a state with two particles in it, one at x_1, the other at x_2.

5.2 The free electron gas at zero temperature

Let's use the language of quantum field theory to calculate some of the properties of a gas of non-interacting electrons at zero temperature. The ground state wave function has one filled state for every value of momentum up to the Fermi momentum p_F, and then all states are empty. This makes the momentum space properties of the Fermi gas pretty simple. However, the coordinate space properties are nontrivial, and those are our goal.

We should spend a few moments trying to make all the factors correct. The ground state wave function $|\Phi\rangle$ must be such that

$$\langle\Phi|a^\dagger_{ps}a_{ps}|\Phi\rangle = 1 \qquad |p| < p_F \tag{5.30}$$

and zero otherwise, where

$$[a_{ps}, a^\dagger_{p's'}]_+ = \delta^3(p-p')\delta_{ss'}. \tag{5.31}$$

We will have a lot of momentum integrals to do, and it is convenient to simplify notation. Put the system in a box of volume V. Then the momentum spectrum is discrete and in fact there are V distinct p's in all. Replacing the Dirac delta function by a Kronecker delta,

$$[a_{ps}, a^\dagger_{p's'}]_+ = \delta_{p,p'}\delta_{ss'}, \tag{5.32}$$

a properly normalized field variable is

$$\psi(r) = \frac{1}{\sqrt{V}}\sum_p e^{ipr}a_p. \tag{5.33}$$

To check (suppressing the spin label):

$$\begin{aligned}[\psi(r), \psi^\dagger(r')]_+ &= \frac{1}{V}\sum_{pp'} e^{ipr}e^{ip'r'}[a_p, a^\dagger_{p'}]_+ \\ &= \frac{1}{V}\sum_p e^{ip(r-r')} \\ &= \frac{V}{V}\delta_{rr'}\end{aligned} \tag{5.34}$$

The number density operator is $n(x) = \psi^\dagger(x)\psi(x)$ and its expectation value is

$$\begin{aligned}\langle\Phi|\psi^\dagger(x)\psi(x)|\phi\rangle &= \frac{1}{V}\sum_{pp'}\langle\Phi|a^\dagger_p a_p|\Phi\rangle \\ &= \frac{1}{V}\sum_p n_p \\ &= = \frac{N}{V}.\end{aligned} \tag{5.35}$$

The gas has uniform density – no surprise. (Incidentally, to make contact with the usual statistical mechanics story,

$$\sum_p = V\int\frac{d^3p}{(2\pi)^3} \tag{5.36}$$

per spin state.) The Fermi momentum is defined via the number density,

$$\frac{N}{V} = \sum_s \int_0^{p_F} \frac{4\pi p^2 dp}{(2\pi)^3} = \frac{p_F^3}{3\pi^2} \tag{5.37}$$

With our confidence high we move to more interesting observables. Consider the Green's function, a generalization of Eq. 5.35,

$$G_s(x - x') = \langle \Phi | \psi_s^\dagger(x) \psi_s(x') | \Phi \rangle \tag{5.38}$$

It is the amplitude to remove a particle at location x' and put it back in at location x. From the calculation of the density, it is

$$\begin{aligned} \langle \Phi | \psi_s^\dagger(x) \psi_s(x) | \Phi \rangle & \\ &= \frac{1}{V} \sum_{ps} n_{ps} p e^{ip(r-r')} \\ &= \frac{N}{V} \frac{\sum_p n_{ps} e^{ip(r-r')}}{\sum_p n_{ps}} \end{aligned} \tag{5.39}$$

The last step is done to pull in an overall factor of the density, $n = N/V$. Our technical problem is the integral

$$I(p) = \int_0^{p_F} d^3 p e^{i\vec{p}\cdot(\vec{r}-\vec{r}')}. \tag{5.40}$$

We do this in spherical coordinates, picking the $\hat{z}$ axis along the $\vec{x} - \vec{x}'$ direction and calling $|\vec{x} - \vec{x}'| = R$. This gives

$$\begin{aligned} I(p) &= 2\pi \int_0^{p_F} p^2 dp dcos\theta e^{ipR\cos\theta} \\ &= 2\pi \int_0^{p_F} p^2 dp \frac{\sin pR}{pR} \\ &= 2\pi \int_0^{p_F} p^2 dp j_0(pR) \end{aligned} \tag{5.41}$$

where $j_0(x)$ is the spherical Bessel function. Dropping this result into Eq. 5.39 and grooming it a bit gives

$$G(r - r') = \frac{3}{2} n \frac{j_1(p_F R)}{p_F R} \tag{5.42}$$

and the limiting value of this expression at $R = 0$ is just $n/2$, as we expect. (Half the particles have one of the spins.)

A more interesting question: what is the probability to find a particle at location x', given that there is one at location x? One way to answer this question is to remove a particle of spin s at x, leaving behind $N-1$ particles, and then measure the density of particles of spin s' in the new state, $|\Phi'(r,s)\rangle$:

$$\begin{aligned}\langle\Phi'(r,s)|\psi^\dagger_{s'}(r')\psi_{s'}(r')\Phi'(r,s)\rangle &= \langle\Phi|\psi^\dagger_s(r)\psi^\dagger_{s'}(r')\psi_{s'}(r')\psi_s(r)\Phi\rangle \\ &\equiv (\frac{n}{2})^2 g_{ss'}(r-r') \\ &= \sum_{pp'qq'} e^{-i(p-p')r}e^{-i(q-q')r'}\,\langle\Phi|a^\dagger_{ps}a^\dagger_{qs'}a_{q's'}a_{p's}|\Phi\rangle\end{aligned} \tag{5.43}$$

There are two cases. First, suppose $s \neq s'$. Then it must be that $p = p'$ and $q = q'$ to get a non-vanishing contraction of the a's and $a^\dagger$'s. Then

$$\begin{aligned}(\frac{n}{2})^2 g_{ss'}(r-r') &= \sum_{pq}\langle\Phi|n_{ps}n_{qs'}|\Phi\rangle \\ &= n_s n_{s'}.\end{aligned} \tag{5.44}$$

This means $g_{ss'}(r-r') = 1$. The different spins do not know about each other.

If $s = s'$ then either $p = p'$, $q = q'$ or $p = q'$,$q = p'$. The first term will contain a factor

$$\langle\Phi|\, a^\dagger_{ps}a^\dagger_{qs}a_{qs}a_{ps}\,|\Phi\rangle = \langle\Phi|\, n_{ps}n_{qs}\,|\Phi\rangle \tag{5.45}$$

after anti-commuting the operators through. The second term has a factor

$$\langle\Phi|\, a^\dagger_{ps}a^\dagger_{qs}a_{ps}a_{qs}\,|\Phi\rangle = -\,\langle\Phi|\, n_{ps}n_{qs}\,|\Phi\rangle \tag{5.46}$$

Putting the pieces together,

$$\begin{aligned}C(r,r') &= \frac{1}{V^2}\sum_{pq} n_{ps}n_{qs}(1-e^{i(p-q)(r-r')}) \\ &= n^2 - |\frac{1}{V}\sum_p n_{ps}e^{i(r-r')})|^2 \\ &= = n^2 - n^2 g(r-r')^2\end{aligned} \tag{5.47}$$

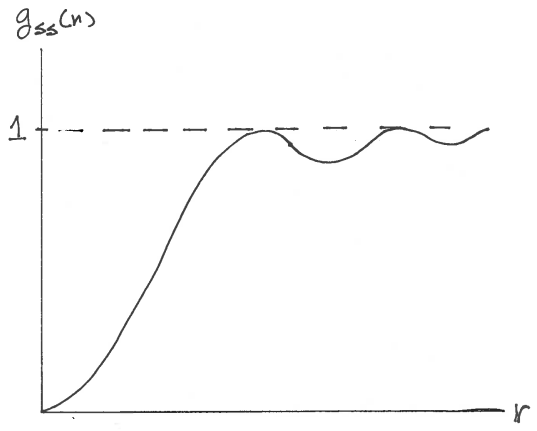

Figure 5.1: pair correlation function for the non-interacting Fermi gas.

recalling the expression for the Green's function. We can call $C(r, r') = n^2\mathcal{G}(r - r'))$ where $\mathcal{G}(r)$ parametrizes the interesting physics. From the explicit functional form of $g(r)$,

$$\mathcal{G}(r) = 1 - (\frac{3}{p_F r} j_1(p_F r))^2. \tag{5.48}$$

Note that $\mathcal{G}(r)$ vanishes at $r = 0$. This is a consequence of Fermi statistics: if there was a particle already at a location r, there will not be a second one there, too.

5.3 Particles that interact among themselves

So far we have discussed "free particles" – particles in a potential $V(x)$, but otherwise not interacting with each other. How can we introduce two-body interactions? This is easy –

add to our one-body Hamiltonian

$$H_1 = \int d^3x \psi^\dagger(x,t)[-\frac{\hbar^2}{2m}\nabla^2 + V(x)]\psi(x,t) \tag{5.49}$$

a two-body term

$$H_2 = \int d^3x d^3x' \psi^\dagger(x,t)\psi^\dagger(x',t)v(x,x')\psi(x,t)\psi(x',t). \tag{5.50}$$

Note that we have chosen the same ordering of operators as we used for the pair correlation function. Here $v(x)$ is the potential between the particles. Now suppose that $v(x,x')$ is, in some sense, small. If it were zero, it would be natural to expand in plane wave states,

$$\psi(x,t) = \frac{1}{\sqrt{V}}\sum_k b_k(t)e^{ikx}. \tag{5.51}$$

If $v(x,x')$ is actually a function of the relative separation of the particles, $v(x-x')$, then we can write the Hamiltonian in a plane wave basis as

$$\begin{aligned} H &= \sum_{k_1 k_2} b^\dagger_{k_1} b_{k_2} \int \frac{d^3x}{V} e^{-ik(_1-k_2)x} \frac{\hbar^2 k^2}{2m} \\ &+ \sum_{k_1}\sum_{k_2}\sum_{k_3}\sum_{k_4} b^\dagger_{k_1} b^\dagger_{k_2} b_{k_3} b_{k_4} \int \frac{d^3x_1}{V}\frac{d^3x_2}{V} e^{i(k_3-k_1)x'} e^{i(k_4-k_2)x} v(x_1-x_2). \end{aligned} \tag{5.52}$$

The first term is

$$\sum_k \frac{\hbar^2 k^2}{2m} b^\dagger_{k_1} b_{k_1} \tag{5.53}$$

(from the delta function in the integral) and the second term can also be condensed: define $r - x' - x$, change variables from x, x' to x, r, and use

$$\int \frac{d^3x_1}{V}\frac{d^3x_2}{V} e^{i(k_3-k_1)r} v(r) e^{i(k_4-k_2+k_3-k_1)x} = \delta^3(k_4-k_2+k_3-k_1)v(q) \tag{5.54}$$

where $\vec{q} = \vec{k}_3 - \vec{k}_1$ and

$$v(q) = \int \frac{d^3x}{V} e^{iqr} v(r). \tag{5.55}$$

Then the two-body Hamiltonian is

$$H_2 = \sum_{k_1}\sum_{k_2}\sum_q v(q) b^\dagger_{k_1+q} b^\dagger_{k_2-q} b_{k_2} b_{k_1}, \tag{5.56}$$

that is, the interaction scatters particles of wave number k_1 and k_2 into wave number $k_1 + q$ and $k_2 - q$ (Or, it annihilates initial state particles and creates final state ones.) We can compute the lowest order T-matrix by specifying initial and final states,

$$\begin{aligned} |i\rangle &= b^\dagger_{p_1} b^\dagger_{p_2} |0\rangle \\ |f\rangle &= b^\dagger_{p_3} b^\dagger_{p_4} |0\rangle \end{aligned} \tag{5.57}$$

and then the lowest order T-matrix element is

$$\langle f| H_2 |i\rangle = v(q)\delta^3(p_3 + p_4 - p_1 - p_2) \tag{5.58}$$

No surprise again, but we can do so much more.

For an off-beat example of the use of this formalism, consider a weakly interacting collection of bosons near absolute zero.

5.4 Excitation in a Bose-Einstein condensate

Recall your statistical mechanics for bosons: the number of particles in a state $\vec{k}$ is given by

$$N(k) = \frac{1}{Ce^{E(k)/T} - 1} \tag{5.59}$$

where T is the temperature (in energy units, $k = 1$) and C is related to the fugacity or chemical potential. In a normal system, C and the particle number are related:

$$\frac{N}{V} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{Ce^{E(k)/T} - 1}. \tag{5.60}$$

However, as T falls, there is no C which can solve the equation. You have to split off the zero energy state,

$$N = \frac{1}{C - 1} + V \int \frac{d^3k}{(2\pi)^3} \frac{1}{Ce^{E(k)/T} - 1} \tag{5.61}$$

and it contributes a finite amount to the right hand side. The mathematics is telling us that the ($\vec{k} = 0$) ground state is macroscopically occupied – there is a condensate. This is in contrast to the microscopic occupation of each remaining phase space differential volume element d^3k.

Let's think about the condensed system in the language of second quantization. The ground state has order N particles in it,

$$b_0^\dagger b_0 \left|\psi\right\rangle = N_0 \left|\psi\right\rangle . \tag{5.62}$$

In addition, there is some occupation at $k \neq 0$

$$b_k^\dagger b_k \left|\psi\right\rangle = N(k) \tag{5.63}$$

given by Eq. 5.59 above.

To make life interesting, let's assume that we do not have an ideal Bose gas, but suppose that there is some interaction among the bosons. We can describe it by some potential $v(r)$. In fact, let's assume that $v(r) = v\delta^3(r)$ so that $v(q) = v$ is a constant. (This is old-fashioned language. Looking back at Eq. 5.58, we see that the T-matrix is a constant, and recalling the chapter about scattering, we are replacing the T-matrix by its scattering length approximation.)

Now for approximations. The ground state operators do not commute, $[b_0, b_0^\dagger] = 1$, but $b_0^\dagger b_0 \left|\psi\right\rangle = N_0 \left|\psi\right\rangle$ where $N_0 \gg 1$. In this sense, b_0 and $b_0^\dagger$ "almost" commute. Let's treat them as classical objects, whose size is about $\sqrt{N}$. Then, in H_2, there is a natural hierarchy of terms:

1. $b_0^\dagger b_0^\dagger b_0 b_0$, the scattering of condensate particles, has a size roughly $v b_0^4 \sim vN^2$
2. Terms like $b_k^\dagger b_0^\dagger b_0 b_0$ vanish – they do not conserve momentum.
3. Order N terms: $b_k^\dagger b_{-k}^\dagger b_0 b_0 + b_0^\dagger b_0^\dagger b_k b_{-k}$ for two particles either leaving or entering the condensate, and $2^2 b_k^\dagger b_k b_0^\dagger b_0$ which controls the scattering of a normal particle off the condensate
4. Ordinary particle scattering is an order (1) effect.

The Hamiltonian of item (3) is

$$H_I^{(3)} = b_0^2 \sum_{k\neq 0} (b_k^\dagger b_{-k}^\dagger + b_k b_{-k} + 4 b_k^\dagger b_k). \tag{5.64}$$

We must be slightly careful with item (1): the total number of particles is

$$N = b_0^2 + \sum_{k\neq 0} b_k^\dagger b_k \tag{5.65}$$

and so

$$b_0^4 = N^2 - 2N\sum_{k\neq 0} b_k^\dagger b_k. \tag{5.66}$$

Then, our approximate Hamiltonian, which considers scattering into or out of the condensate, is

$$H = \sum_k \frac{\hbar^2 k^2}{2m} b_k^\dagger b_k + N^2 v + Nv\sum_{k\neq 0}(b_k^\dagger b_{-k}^\dagger + b_k b_{-k} + (4-2)b_k^\dagger b_k). \tag{5.67}$$

Now notice something important: this H is quadratic in the b's. We can make a change of variables, to a new basis, $a_k^\dagger$ and a_k, and write

$$H = N^2 v + \sum_k \epsilon(k) a_k^\dagger a_k; \tag{5.68}$$

that is, the system is described by a set of excitations of "quasi-particles" annihilated by a_k. These are the normal modes of the system. The transformation will be

$$a_k = \frac{b_k + L_k b_{-k}^\dagger}{\sqrt{1-L_k^2}}; \qquad a_k^\dagger = \frac{b_k^\dagger + L_k b_{-k}}{\sqrt{1-L_k^2}} \tag{5.69}$$

or

$$b_k = \frac{a_k - L_k a_{-k}^\dagger}{\sqrt{1-L_k^2}}; \qquad b_k^\dagger = \frac{a_k^\dagger - L_k a_{-k}}{\sqrt{1-L_k^2}}. \tag{5.70}$$

Of course, we have to find L_k and we do this by making the substitutions into H and choosing it to cancel the unwanted terms. Notice that

$$[a_k, a_{k'}^\dagger] = \frac{[b_k, b_{k'}^\dagger] + L_k^2[b_{-k}^\dagger, b_{-k'}]}{1-L_k^2} = \delta_{kk'} \tag{5.71}$$

so the correct creation-annihilation operator algebra is maintained. Physically, a_k annihilates an original boson carrying momentum k (for a momentum change $= -k$) or creates a boson with momentum $-k$, which is also a momentum change $-k$. A wee bit of algebra yields

$$L_k = \frac{\epsilon_k + 2Nv - E(k)}{2Nv} \tag{5.72}$$

where $\epsilon_k = \hbar^2 k^2/(2m)$ and

$$E(k)^2 = (\frac{\hbar^2 k^2}{2m})^2 + 2Nv\frac{k^2}{m} \tag{5.73}$$

$E(k)$ is the energy of an excitation with momentum k. As the equation and Fig. 5.2 show, the high momentum quasi-particles are just the original particles. However, the low-lying, long wavelength spectrum has a linear dispersion relation, $E(k) = C_s\hbar k$ where $C_s =$

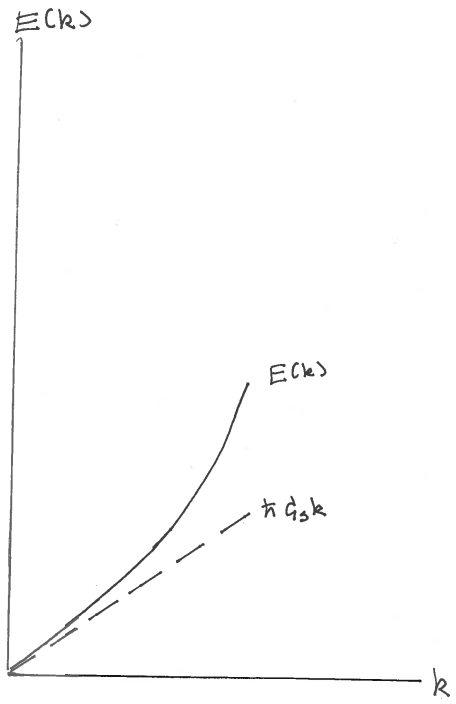

Figure 5.2: Quasi-particle (phonon) dispersion relation for the weakly-interacting condensed Bose gas.

$\sqrt{2Nv/m}$. These are sound waves – they are called "phonons," in complete analogy with the quantized vibrational modes of a crystal lattice. The non-ideal weakly-interacting condensed Bose gas is best thought of as a gas of non-interacting phonons.

Finally, suppose we have an impurity atom (of mass M) moving through the condensate. The only way it can lose energy is to create an excitation in the condensate. Suppose it has initial momentum $\hbar\vec{q}$, and also suppose it creates an excitation of momentum $\hbar\vec{k}$ at an angle θ away from its direction of motion. Conservation of energy says

$$\begin{aligned} \frac{\hbar^2 q^2}{2M} &= \frac{\hbar^2}{2M}|\vec{q}-\vec{k}|^2 + E(k) \\ &= \frac{\hbar^2}{2M}[q^2+k^2-2kq\cos\theta] + E(k) \end{aligned} \tag{5.74}$$

or

$$qk\cos\theta = \frac{k^2}{2} + \frac{ME(k)}{\hbar^2}. \tag{5.75}$$

Calling the impurity's velocity $v = \hbar q/M$, this is

$$\cos\theta = \frac{k}{q} + \frac{ME(k)}{\hbar^2 k} = \frac{\hbar k}{2Mv} + \frac{E/\hbar k}{v}. \tag{5.76}$$

For phonons, $E(k)/\hbar k > C_s$, so we need $v > C_s$ for $\cos\theta < 1$, allowing a phonon to be emitted. This means that if the impurity atom is moving too slowly, it cannot emit a phonon, and it cannot lose energy. Now look at the process from the point of view of an observer riding along with the impurity. The condensate streams past without friction (if the velocity of the condensate is low enough). This is superfluidity – we have just discovered that the condensate is a superfluid!

Chapter 6

Quantum field theory from a Hamiltonian point of view

6.1 Time dependent perturbation theory

We need approximate solutions for the time dependent Schrödinger equation. We assume that the Hamiltonian can be written as $H = H_0 + H_1$, where H_0 is "large" in some sense and can be diagonalized. We further assume that H_1 may be regarded as a perturbation. The time evolution will be developed in terms of the eigenfunctions of H_0.

It is very useful to replace the time dependent wave function by a time evolution operator, and to find matrix elements of that operator. It is defined through

$$|\psi(t)\rangle = U(t, t_0)\,|\psi(t_0)\rangle \tag{6.1}$$

and it obeys an equation of motion

$$i\hbar\frac{\partial U(t, t_0)}{\partial t} = HU(t, t_0). \tag{6.2}$$

Matrix elements of U connect initial states at some early time with final states at some later time. What they encode is called the "transition amplitude" from the initial state to the final state. It is defined as

$$U_{\beta\alpha}(t, t_0) = \langle\beta(t)|U(t, t_0)|\alpha(t_0)\rangle \tag{6.3}$$

To find a practical expression for the evolution operator, let us define the "interaction representation" perturbation

$$\hat{H}_1(t) = e^{iH_0t/\hbar}H_1(t)e^{-iH_0t/\hbar} \tag{6.4}$$

and the "interaction representation" evolution operator

$$\hat{U}(t_f, t_i) = e^{iH_0t_f/\hbar}U(t_f, t_i)e^{-iH_0t_i/\hbar}. \tag{6.5}$$

We want to solve

$$i\hbar\frac{\partial}{\partial t}U(t, t_0) = (H_0 + H_1)\,U(t, t_0). \tag{6.6}$$

In interaction representation, Eq. 6.6 becomes

$$i\hbar\frac{\partial}{\partial t}\left[e^{-iH_0t/\hbar}\hat{U}e^{iH_0t/\hbar}\right] = (H_0 + H_1)\,e^{-iH_0t/\hbar}\hat{U}e^{iH_0t_0/\hbar}. \tag{6.7}$$

The left hand side is

$$i\hbar\frac{\partial}{\partial t}\left[e^{-iH_0t/\hbar}\hat{U}e^{iH_0t_0/\hbar}\right] = \left[H_0e^{-iH_0t/\hbar}\hat{U} + e^{-iH_0t/\hbar}i\hbar\frac{\partial\hat{U}}{\partial t}\right]e^{iH_0t_0/\hbar}, \tag{6.8}$$

so

$$\left[H_0 e^{-iH_0 t/\hbar}\hat{U} + e^{-iH_0 t/\hbar} i\hbar \frac{\partial \hat{U}}{\partial t}\right] e^{iH_0 t_0/\hbar} = (H_0 + H_1) e^{-iH_0 t/\hbar}\hat{U} e^{iH_0 t_0/\hbar} \tag{6.9}$$

or

$$i\hbar \frac{\partial \hat{U}}{\partial t} = \left(e^{iH_0 t/\hbar} H_1^{-iH_0 t/\hbar}\right)\hat{U} = \hat{H}_1 \hat{U}. \tag{6.10}$$

This is just a first order differential equation. Integrating it, we find

$$i\hbar\left[\hat{U}(t,t_0) - \hat{U}(t_0,t_0)\right] = \int_{t_0}^{t} dt' \hat{H}_1(t')\hat{U}(t',t_0). \tag{6.11}$$

Because $\hat{U}(t_0,t_0) = 1$, the evolution operator satisfies the integral equation

$$\hat{U}(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt' \hat{H}_1(t')\hat{U}(t',t_0). \tag{6.12}$$

Like all Volterra integral equations, it can be solved by iteration:

$$\begin{aligned}\hat{U}(t_f,t_i) = \quad & 1 - \frac{i}{\hbar}\int_{t_i}^{t_f} dt_1 \hat{H}_1(t) + \left(-\frac{i}{\hbar}\right)^2 \int_{t_i}^{t_f} dt_2 \int_{t_i}^{t_2} dt_1 \hat{H}_1(t_2)\hat{H}_1(t_1) \\ & + \left(-\frac{i}{\hbar}\right)^3 \int_{t_i}^{t_f} dt_3 \int_{t_i}^{t_3} dt_2 \int_{t_i}^{t_2} dt_1 \hat{H}_1(t_3)\hat{H}_1(t_2)\hat{H}_1(t_1) + \cdots\end{aligned} \tag{6.13}$$

The operators $H_1(t)$ may not commute at different times, and so it is important to preserve their time ordering.

This is called the "Dyson series" for the evolution operator. Note the nested time integrals, preserving the ordering in the multiple time integrals of the temporal points where the potential acts.

We can write this even more compactly by introducing the "time ordering operator" $T(A(t_1)B(t_2)\dots)$. The time ordering operation takes the operators and evaluates them in ascending order from the right; with the operator at earliest time farthest right, then the next earliest one, and so on until the operator evaluated at the latest time sits at the far left. Then we can think of the times t_i in Eq. 6.13 as a set of dummy labels and write the evolution operator as

$$\hat{U}(t_f,t_i) = \sum_{n=0}^{\infty}\left(\frac{i}{\hbar}\right)^n \frac{1}{n!}\int_{t_i}^{t_f} dt_1 \int_{t_i}^{t_f} dt_2 \dots \int_{t_n}^{t_f} dt_n T(\hat{H}_1(t_1)\hat{H}_1(t_2)\dots\hat{H}_1(t_n)) \tag{6.14}$$

For useful insight, look at the second order term in the sum, and return to Schrödinger representation. Consider a transition from one eigenstate of H_0, $|\alpha\rangle$, to another one, $|\beta\rangle$.

Inserting complete sets of states the transition amplitude becomes

$$\begin{aligned} U_{\beta\alpha} &= \left(-\frac{i}{\hbar}\right)^2 \sum_\gamma \int_{t_i}^{t_f} dt_2 \int_{t_i}^{t_2} dt_1 e^{-iE_\beta(t_f-t_i)/\hbar} \\ &\quad \times \langle\beta|H_1(t_2)|\gamma\rangle\, e^{-iE_\gamma(t_2-t_1)/\hbar}\, \langle\gamma|H_1(t_1)|\alpha\rangle\, e^{-iE_\alpha(t_1-t_i)/\hbar}. \end{aligned} \tag{6.15}$$

Notice the structure (moving in from the right): the system begins in state $|\alpha\rangle$ and its time evolution is the usual phase factor associated with an energy eigenstate. Then the perturbation acts at one time t_1, driving the system into a superposition of eigenstates of H_0, the $|\gamma\rangle$ states. each $|\gamma\rangle$ state is also an energy eigenstate and evolves appropriately in time Lastly, the perturbation acts again at t_2 to carry the system into $|\beta\rangle$.

6.2 The S-matrix

The S-matrix is defined as the limit

$$S = \lim_{t\to\infty} \lim_{t_0\to-\infty} U(t, t_0). \tag{6.16}$$

We only have eigenstates of H_0 so we always compute matrix elements of S between them, $\langle\beta|S|\alpha\rangle$. We make a set of demands on S, and these demands constrain possible Hamiltonians. We set $\hbar = 1$ from here on.

6.2.1 Unitarity

We want $\langle\beta|S^\dagger S|\alpha\rangle = \delta_{\alpha\beta}$. This is the same thing as asking whether the evolution operator is unitary. To satisfy this requirement, H_I must be Hermitian. The proof is simple. $U(t_0, t_0) = \hat{U}(t_0, t_0) = 1$. Then

$$i\frac{d}{dt}[\hat{U}(t, t_0)^\dagger \hat{U}(t, t_0) = [-\hat{U}^\dagger \hat{H}_I^\dagger]\hat{U} + \hat{U}^\dagger[\hat{H}_I\hat{U}] = \hat{U}^\dagger[-\hat{H}_I^\dagger + \hat{H}_I]\hat{U} \tag{6.17}$$

where we have used

$$\frac{d\hat{U}}{dt} = -i\hat{H}_I\hat{U} \tag{6.18}$$

and its Hermitian conjugate. If $\hat{H}_I$ is Hermitian, $U^\dagger U$ is a constant and the constant is unity from $U(t_0, t_0) = \hat{U}(t_0, t_0) = 1$.

6.2.2 Energy conservation

Is $S_{\beta\alpha} = 0$ if $E_\alpha \neq E_\beta$? This is tricky! Recall $H_0 |\alpha\rangle = E_\alpha |\alpha\rangle$, so we might consider

$$\langle\beta|[H_0, S]|\alpha\rangle = (E_\beta - E_\alpha)S_{\beta\alpha}. \tag{6.19}$$

Is this zero? If so, we have energy conservation. But there is a problem: S depends on H_I and generally $[H_0, H_I] \neq 0$.

The escape is dirty. We have to (formally) adiabatically switch on and off H_I at early and late times, so that we can actually construct our "in" and "out" states. Then (writing formally)

$$U(t, t_0) = T \exp(i(H_o + H_I^{-\lambda|t|})t) \tag{6.20}$$

so that

$$\lim_{t\to\pm\infty}[H_0, U] \to 0. \tag{6.21}$$

This says that we can only have energy conservation between the beginning and the end of the reaction – the intermediate states can be anything. So much for being formal. The algebra will take care of energy conservation for us, modulo technical problems associated with squaring delta functions. For example

$$S^{(1)}_{\alpha\beta} = -i \int_{-\infty}^{\infty} dt e^{iE_\alpha t} \langle\alpha|H_I|\beta\rangle e^{-iE_\beta t} \propto \delta(E_\alpha - E_\beta). \tag{6.22}$$

In fact, it will turn out, depending on how time ordered products are handled, that either

1. Energy isn't conserved throughout the reaction, but all particles are on shell, $p^2 = m^2$
2. OR E is conserved, but intermediate state particles do not have $p^2 = m^2$.

6.2.3 Momentum conservation

We want $S_{\beta\alpha} = 0$ if $\sum_{i\in\alpha} p_i \neq \sum_{j\in\beta} p_j$ or more formally $[P, S] = 0$ where P is the momentum operator. We will get this if $[P, H_I] = 0$. The momentum P is not only an operator, it is the generator of translations $x \to x + a$. Then S will be invariant under space translations if H is, and momentum will be conserved. As you know, scattering in an external field does not conserve momentum, but in that case there is no translation invariance. We have to be careful about treating the momentum as a four vector and asking for time translation invariance, because in Hamiltonian quantum mechanics, time is special: it is "what we evolve in."

6.2.4 Lorentz invariance

Consider a Lorentz transformation (some combination of boosts and rotations) parametrized by a matrix Λ (with six perameters, three for rotations, three for boosts). We have observers in two frames:

- O sees state $|\alpha\rangle$
- O' sees state $|\Lambda\alpha\rangle$

In analogy with rotation matrices, states are transformed by a unitary operator $U(\Lambda)$,

$$|\Lambda\alpha\rangle = U(\Lambda)|\alpha\rangle \tag{6.23}$$

and operators must transform as

$$O \to U(\Lambda)OU^{-1}(\Lambda). \tag{6.24}$$

That the transformation is unitary means that $\langle\alpha|\beta\rangle = \langle\Lambda\alpha|\Lambda\beta\rangle = \delta_{\alpha\beta}$. Different observers must see S-matrices which do the same thing, so $S_{\alpha,\beta} = S_{\Lambda\alpha,\Lambda\beta}$, or

$$\begin{aligned} \langle\beta|S|\alpha\rangle &= \langle U(\Lambda)\beta|S|U(\Lambda)\alpha\rangle \\ &= \langle\beta|U(\Lambda)^{-1}SU(\Lambda)|\alpha\rangle\,. \end{aligned} \tag{6.25}$$

So, we want $S = U^{-1}SU$ or $[S, U(\Lambda)] = 0$.

This just looks like formalism, until we realize that the Hamiltonian perturbative S-matrix treats time in a special way – there is the time ordered product. (Go back and look at Eq. 6.14.) How can we solve this constraint? There is one possibility everyone knows (more arcane versions, for particles with higher spin, or derivative interactions, are described in Weinberg – keep it simple, now): First, write the perturbing Hamiltonian as an integral over a Hamiltonian density

$$H_I(t) \to \int d^3x \mathcal{H}(x,t) \tag{6.26}$$

where H is a Lorentz scalar. Being a scalar means that the transformation rule for $\mathcal{H}$ is

$$U(\Lambda)^{-1}\mathcal{H}(x)U(\Lambda) = \mathcal{H}(\Lambda^{-1}x). \tag{6.27}$$

(the argument changes but $\mathcal{H}$ does not change; Λ^{-1} because the order of the U's in Eq. 6.27 is reversed compared to Eq. 6.23.) Then the S-matrix is

$$S = \sum_n \frac{(-i)^n}{n!} \int d^4x_1 \dots d^4x_n T(\mathcal{H}(x_1) \dots \mathcal{H}(x_n)). \tag{6.28}$$

This looks promising: all the integrations, and $\mathcal{H}$, are invariants. Unfortunately, we are not out of the woods yet. The time ordering operator might not be Lorentz invariant. Do a Lorentz transformation:

$$U^{-1}SU = \sum_n \frac{(-i)^n}{n!} \int d^4x_1 \dots d^4x_n T_x(\mathcal{H}(\Lambda^{-1}x_1) \dots \mathcal{H}(\Lambda^{-1}x_n)). \tag{6.29}$$

I wrote T_x to remind that the time ordering is in the frame of x, NOT in $\Lambda^{-1}x$. To get this, I have inserted unity as $U(\Lambda)^{-1}U(\Lambda) = 1$ between the $\mathcal{H}$'s in the time ordered product. Now change variables, writing $x_i = \Lambda y_i$. The integration measure is invariant, $d^4x = d^4y$ so

$$U^{-1}SU = \sum_n \frac{(-i)^n}{n!} \int d^4y_1 \dots d^4y_n T_x(\mathcal{H}(y_1) \dots \mathcal{H}(y_n)). \tag{6.30}$$

T_x contains terms like

$$\theta((\Lambda^{-1}y_1)^0 - (\Lambda^{-1}y_2)^0)\mathcal{H}(y_1)\mathcal{H}(y_2) + \theta((\Lambda^{-1}y_2)^0 - (\Lambda^{-1}y_1)^0)\mathcal{H}(y_2)\mathcal{H}(y_1). \tag{6.31}$$

Now there are two possibilities. If Λ does not change the ordering of times, this expression will be the same in all frames. If the two points x_1 and x_2 are separated by a time-like interval, the time ordering will not change, and $S = U^{-1}SU$. We have Lorentz invariance.

However, if the two points are space-like separated, one can move from a frame where $t_1 > t_2$ to one where $t_1 < t_2$. To maintain Lorentz invariance, we must be able to exchange the order of the two $\mathcal{H}$'s in the time ordering. To do this, we need $\mathcal{H}(y_1)\mathcal{H}(y_2) = \mathcal{H}(y_2)\mathcal{H}(y_1)$, or

$$[\mathcal{H}(y_1), \mathcal{H}(y_2)] = 0 \; for \; (y_1 - y_2)^2 < 0. \tag{6.32}$$

We say "our Hamiltonian commutes outside the light cone." Eqs. 6.26 and 6.32 are the requirements Lorentz invariance imposes on any quantum theory.

6.3 Relativistic spin-zero quantum fields

We want to consider scattering processes in perturbation theory from a state $|\Phi_{k_1,k_2\dots}\rangle \propto a^\dagger(k_1)a^\dagger(k_2)\dots|0\rangle$ to some final state $|\Phi_{k'_1,k'_2\dots}\rangle$. We will build interaction Hamiltonians out

of fields and their derivatives, since that is all we have in a field theory. Our states are created and destroyed by field operators. In position space, a potential scalar field operator could be

$$\phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(p)}}[a(p)e^{ipx}], \tag{6.33}$$

where we temporarily assume that we only have a field annihilation operator in ϕ. The factors of π and $E(p)$ are a convenient choice (see the notes on conventions), and it lets us quantize with $[a(p), a^\dagger(p')] = \delta^3(p - p')$.

Now, we know we need two things:

- $\mathcal{H}$ must be Hermitian, so $\mathcal{H} = \mathcal{H}(\phi, \phi^\dagger)$
- $[\mathcal{H}(y_1), \mathcal{H}(y_2)] = 0$ for space-like separation, $(y_1 - y_2)^2 < 0$.

Is Eq. 6.33 good enough to satisfy these requirements? If $[\phi(x), \phi(y)^\dagger] = 0$ for $(x - y)$ space-like, we're good. Let's see:

$$[\phi(x), \phi(y)^\dagger] = \frac{1}{(2\pi)^3}\int \frac{d^3k d^3k'}{\sqrt{4E(k)E(k')}} e^{-i(kx-k'y)}[a(k), a^\dagger(k')] \tag{6.34}$$

(four dimensional dot products in the exponential, $kx = E(k)x^0 - \vec{k}\cdot\vec{x}$), or, from the commutator $[a(k), a^\dagger(k')] = \delta^3(k - k')$,

$$[\phi(x), \phi(y)^\dagger] \equiv \Delta_+(x - y) = \frac{1}{(2\pi)^3}\int \frac{d^3k}{2E(k)} e^{-ik(x-y)}. \tag{6.35}$$

This is NOT an equal time commutator ($x_0 \neq y^0$). It is a scalar function of $(x - y)$. Let's evaluate it.

(a) $x - y$ time-like: Pick a frame where $\vec{x} - \vec{y} = 0$, $x^0 - y^0 \equiv x = \pm\sqrt{(x-y)^2}$:

$$\begin{aligned} \Delta_+(x) &= \frac{4\pi}{2(2\pi)^3}\int \frac{k^2 dk}{\sqrt{k^2+m^2}} e^{\mp ix\sqrt{k^2+m^2}} \\ &= \frac{m}{8\pi x}(N_1(mx) \pm J_1(mx)) \end{aligned} \tag{6.36}$$

using tables, and the J_1 and N_1 are Bessel functions.

(b) $x-y$ space-like. Work in a frame where $t=0$, $|x| = \sqrt{-(x-y)^2}$

$$\begin{aligned}\Delta_+(x) &= \frac{1}{(2\pi)^3}\int \frac{k^2 dk d\Omega}{\sqrt{k^2+m^2}} e^{i\vec{k}\cdot\vec{x}} \\ &= \frac{1}{(2\pi)^2}\int \frac{k^2 dk}{\sqrt{k^2+m^2}} \frac{\sin kx}{kx} \\ &= \frac{m}{4\pi^2 x} K_1(mx)\end{aligned} \tag{6.37}$$

Tables again, and this is a Bessel function of imaginary argument,

$$K_1(mx) \sim \sqrt{\pi/(2mx)}\exp(-mx). \tag{6.38}$$

Although we evaluates these expressions at particular points, we can slide the results up and down the light cone by doing boosts. Eq. 6.36 is benign; the time ordering of time-like separated events is preserved by Lorentz transformations. But Eq. 6.37 is bad news: it says that $[\phi(x), \phi(y)^\dagger] \neq 0$ outside the light cone. Our guess for $\phi(x)$, Eq. 6.33, is inconsistent with special relativity.

What to do? We need something additional to cancel Δ_+ outside the light cone. Guess number two:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(p)}}[a(p)e^{ipx} + \eta a_c^\dagger(k)e^{-ipx}], \tag{6.39}$$

Now, if we assume that a and a_c commute, and we give a_c and $a_c^\dagger$ the same commutation relation as for a and $a_\dagger$, $[a_c(p), a_c(p')^\dagger] = \delta^3(p-p')$, the field commutator is

$$[\phi(x), \phi(y)^\dagger] \sim [a + a_c^\dagger, a^\dagger + a_c] = \Delta_+(x-y) - |\eta|^2\Delta_+(y-x). \tag{6.40}$$

Since Δ_+ is an even function outside the light cone, we can achieve cancellation by setting $\eta = 1$ (you can absorb phases into a_c if you want).

So (punch line) to satisfy Lorentz invariance, for every particle annihilated by ϕ there must be another particle created by ϕ, and it has to have the same mass, so that the two Δ_+ functions cancel. Hmm... what about its quantum numbers?

If some quantity is conserved, its charge operator commutes with the Hamiltonian,

$$[Q, \mathcal{H}] = 0 \tag{6.41}$$

or, if $\mathcal{H}$ is a product of fields,

$$[Q, \prod_{i\in\mathcal{H}} \phi_i(x)] = 0. \tag{6.42}$$

Now,

$$\begin{aligned}[Q, a_i(k)] |\psi\rangle &= Q[a_i(k) |\psi\rangle] - a_i(k) Q |\psi\rangle \\ &= (q_\psi - q_i) a_i |\psi\rangle - q_\psi a_i |\psi\rangle \\ &= -q_i a_i |\psi\rangle \end{aligned} \tag{6.43}$$

or $[Q, a_i] = -q_i a_i$. Similarly, $[Q, a_{ci}^\dagger] = +q_{ic} a_{ic}^\dagger$. The physical interpretation of Eq. 6.42 is that we want the total charge annihilated at any interaction to be zero. This means that the total charge never changes; it is conserved. To achieve this, it must be that

$$q_i = -q_{ic} \tag{6.44}$$

for all i. The particle annihilated by a_{ic} has to have the opposite charge as the particle annihilated by a, in addition to having the same mass. This is the antiparticle of i. The field operator creates antiparticles and annihilates particles with exactly the same strength.

Antiparticles do not exist because the Dirac equation has negative energy solutions. Antiparticles exist to make the S-matrix Lorentz invariant.

Some comments:

- a_c is called the "charge conjugate field." (we will describe charge conjugation later on).
- $\eta = 1$ for scalars, for convenience. We will discuss spinors ($j = 1/2$) later.

6.4 Example 1: ϕ^4 field theory

For our first example let us set $a_c = a$ so ϕ is Hermitian. (This is called a "real scalar field" in the literature.) Pick an interaction Hamiltonian density

$$\mathcal{H}(x) = \frac{g}{4!} \phi(x)^4. \tag{6.45}$$

Examples of such systems are the Higgs boson, or in statistical mechanics, the field theoretic analog of the Ising model. Now calculate a scattering process $k_1 + k_2 \to k_3 + k_4$. We can do this in first order perturbation theory,

$$S(k_1 + k_2 \to k_3 + k_4) = 1 - i \int d^4x \, \langle \Phi_{k_3,k_4} | \mathcal{H}(x) | \Phi_{k_1,k_2} \rangle \tag{6.46}$$

where

$$|\Phi_{k_1,k_2)} = \sqrt{2E_i 2E_j} a(k_i\rangle^\dagger \, a(k_k)^\dagger \, |0\rangle \, . \tag{6.47}$$

The "1" term does not contribute to scattering so we discard it.

Let us schematically evaluate the matrix element. Write the field operator four times using Eq. 6.42, insert it in Eq. 6.46, and inside the thicket of math symbols you will find the expression

$$\frac{1}{4!} \langle 0|a(k_3)a(k_4)[a_a + a_a^\dagger][a_b + a_b^\dagger][a_c + a_c^\dagger][a_d + a_d^\dagger]a(k_1)^\dagger a(k_2)^\dagger|0\rangle \tag{6.48}$$

Assume all the $k_a \ldots k_d$ are unequal (the equal ones form a set of measure zero). The only non-vanishing terms are those for which two of the a's, and two of the $a^\dagger$'s, in the brackets, contract against the in and out states. Make a table: there are 4 ways to contract a_d or $a_d^\dagger$ against an $a(k)$ or an $a(k)^\dagger$. This leaves 3 ways to contract the a_c or $a_c^\dagger$, two ways for the a_b (etc) and one way is left. This gives an overall 4! combinitorial factor. Each contraction has the same weight. So we are left with

$$S(k_1 + k_2 \to k_3 + k_4) = -ig \int \frac{d^4x}{(2\pi)^6} e^{i(k_1+k_2-k_3-k_4)x} \frac{\sqrt{2E_1 2E_2 2E_3 2E_4}}{\sqrt{2E_1 2E_2 2E_3 2E_4}} \tag{6.49}$$

or

$$S(k_1 + k_2 \to k_3 + k_4) = -ig \frac{(2\pi)^4}{(2\pi)^6} \delta^4(k_1 + k_2 - k_3 - k_4) \tag{6.50}$$

If you have read the notes on "Conventions for fields and scattering amplitudes," you know what comes next. With $S = 1 + \mathcal{T}$, we can strip off the delta-function and the factors of 2π, leaving the "effective T-matrix T" or "invariant amplitude M" to plug into the formula for the cross section. I'll recopy it, to save you time:

$$d\sigma = \frac{1}{4E_1 E_2 v_{rel}} |M|^2 (2\pi)^4 \delta^4(\sum k_f - \sum k_i) \prod_{j=1}^{n_f} \frac{d^3k_j}{(2\pi)^3 2E(k_j)}. \tag{6.51}$$

In this case, the invariant amplitude is very simple,

$$M = -ig. \tag{6.52}$$

What is the cross section? We have to integrate the momenta over the delta function. We can get this with (yet) another trick: writing $E = p^0$ we convert the three dimensional integral to a four dimensional one

$$\frac{d^3p}{2E} = d^4p \delta^+(p^2 - m^2) \tag{6.53}$$

The δ^+ means to take only the positive square root. Then the phase space integral is

$$d^4k_3d^4k_4\delta^4(k_3+k_4-k_1-k_2)\delta^+(k_3^2-m^2)\delta^+(k_4^2-m^2) = \frac{d^3k_3}{2E_3}\delta^+((k_1+k_2-k_3)^2-m^2) \quad (6.54)$$

Let's evaluate this in the center of mass frame:

$$\begin{aligned} k_1 &= (E,\vec{k}) \\ k_2 &= (E,-\vec{k}) \\ k_3 &= (E_3,\vec{k}) \end{aligned} \quad (6.55)$$

where $k_3^2+m^2 = E_3^2$. The delta function is $\delta((k_1+k_2)^2-2(k_1+k_2)\cdot k_3) = \delta(s-4EE_3)$ where s is called a "Mandelstam invariant': it is a generalization of the squared center of mass energy, $s = (k_1+k_2)^2 = 4E_{CM}^2$. Incidentally, the relative velocity is $v_{rel} = 2k/E$. Introducing the angular dependence through $d^3k = k^2dkd\Omega$, a few steps of algebra give

$$\frac{d\sigma}{d\Omega} = \frac{g^2}{8\pi^2 s} \quad (6.56)$$

(isotropic scattering) and the total cross section is

$$\sigma = \frac{g^2}{2\pi s}. \quad (6.57)$$

Many ingredients in this calculation will be repeated over and over. To keep track of these ingredients, it is convenient to introduce a diagrammatic language – Feynman diagrams and Feynman rules. Our little exercise gives us some proto-rules:

1) For this theory, each order in perturbation theory contributes a weight $-ig/4!$. The $-i$ is from the Dyson series, the 4! from the $g/4!$ in $\mathcal{H}$. We display this weight as a vertex with four lines meeting. Note that momentum is conserved at the vertex.

2) Each external line (incoming or outgoing particle) contributes a weight 1 to M or $1/((2\pi)^3\sqrt{2E(k)}$ to the T-matrix.

3) There is an overall combinitorial factor of 4! contracting field in $\mathcal{H}$ against those in the in - and out - states

The associated diagram is shown in Fig. 6.1. Feynman diagrams are not cartoons! Each line or vertex refers to some specific mathematical function.

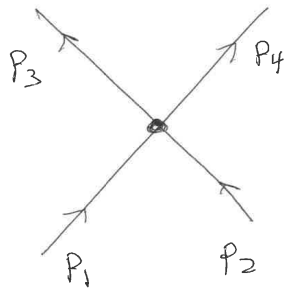

Figure 6.1: Lowest order scattering Feynman diagram for $\mathcal{H} = g\phi^4/4!$.

6.5 Example 2: Discovering the propagator

For our next example, imagine three species of scalar fields A, B, and C, none of which is its own antiparticle. Let

$$\mathcal{H}(x) = g[A(x)B(x)C(x) + h.c.] \tag{6.58}$$

Let's calculate $BC \to BC$ scattering. We need second order perturbation theory.

$$S(k_b + k_c \to k_b' + k_c') = \frac{(-i)^2}{2!}\int d^4x d^4y \, \langle \Phi_{k_b',k_c'}|T(\mathcal{H}(x)\mathcal{H}(y))|\Phi_{k_b,k_c}\rangle \,. \tag{6.59}$$

The two vertices are at space-time points x and y. At one of these points, b' and c' are created, and at the other point b and c are annihilated. Each choice gives the same weight to S, so pick one of them and cancel the 2!. As before, we have to evaluate

$$\langle 0|b(k_b')c(k_c')\{T((a_1^\dagger + a_{c1})(b_2^\dagger + b_{c2})(c_3^\dagger + c_{c3})(a_4 + a_{c4}^\dagger)(b_5 + b_{c5}^\dagger)(c_6 + c_{c6}^\dagger))\}b(k_b)^\dagger c(k_c)^\dagger|0\rangle \,. \tag{6.60}$$

b_2 annihilates $b(k_b)^\dagger$, c_3 annihilates $c(k_c)^\dagger$, and only

$$\langle 0|T((a_1^\dagger + a_{c1})(a_4 + a_{c4}^\dagger)|0\rangle \tag{6.61}$$

is left. In full gory detail we have

$$S = -g^2\int d^4x d^4y \frac{e^{-ik_b y}}{(2\pi)^{3/2}}\frac{e^{-ik_c y}}{(2\pi)^{3/2}}\frac{e^{ik_b' x}}{(2\pi)^{3/2}}\frac{e^{ik_c' x}}{(2\pi)^{3/2}}\langle 0|T(A(x)^\dagger A(y))|0\rangle \,. \tag{6.62}$$

The quantity $\langle 0|T(A(x)^\dagger A(y))|0\rangle$ is called the Propagator. Anticipating the calculation, it is a function of $(x-y)$. We can perform the x and y integrals by shifting variables to x and $x-y$. The x integral gives a delta function leaving

$$S = -g^2\frac{(2\pi)^4\delta^4(k_{b'} + k_{c'} - k_b - k_c)}{(2\pi)^6}[-i\Delta_F(k_b + k_c)] \tag{6.63}$$

where

$$-i\Delta_F(k) = \int d^4z e^{ikz}\langle 0|T(A(z)^\dagger, A(0))|0\rangle \,. \tag{6.64}$$

Of course,

$$M = (-ig)^2[-i\Delta_F(k_b + k_c)]. \tag{6.65}$$

Again we can draw a picture. See Fig. 6.2. And also note that x^0 is not necessarily later (or earlier) than y^0.

We have to calculate the propagator. We begin with the coordinate space version

$$-i\Delta_F(x-y) = \theta(x^0-y^0)\,\langle 0|A(x)^\dagger A(y)|0\rangle + \theta(y^0-x^0)\,\langle 0|A(y)A(x)^\dagger|0\rangle \tag{6.66}$$

The first term is

$$\frac{\theta(x^0-y^0)}{(2\pi)^3}\int \frac{d^3kd^3k'}{\sqrt{4E(k)E(k')}}\,\langle 0|[a(k)^\dagger e^{ikx}+a_c(k)e^{-ikx}][a(k')e^{-ik'y}+a_c(k')^\dagger e^{ik'y}]|0\rangle\,. \tag{6.67}$$

The $a_c a_c^\dagger$ term is the only non-vanishing one, and $k=k'$ is needed, so the first term collapses to

$$\frac{\theta(x^0-y^0)}{(2\pi)^3}\int \frac{d^3k}{(2\pi)^3 2E(k)}e^{-ik(x-y)}. \tag{6.68}$$

The second term is similar. The coordinate space propagator is

$$-i\Delta_F(x-y) = \int \frac{d^3k}{(2\pi)^3 2E(k)}[\theta(x^0-y^0)e^{-ik(x-y)}+\theta(y^0-x^0)e^{ik(x-y)}].. \tag{6.69}$$

Let's write this as a four dimensional Fourier transform. We use the integral expression with a convergence factor

$$\int_{-\infty}^{\infty} dt e^{i\omega t}\theta(t) = \int_0^\infty dt e^{i\omega t}e^{-\epsilon t} = \frac{i}{\omega+i\epsilon} \tag{6.70}$$

and so

$$\theta(t) = \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega \frac{e^{-i\omega t}}{\omega+i\epsilon} \tag{6.71}$$

Convert this to a contour integral and close the contour in the upper half plane if $t<0$, in the lower half plane if $t>0$, pick up the residue, and it checks.

Then (here note the $(t,\vec{x})$ and define $E(k)=\sqrt{k^2+m^2}$)

$$-i\Delta_F(x) = \frac{i}{(2\pi)^4}\int \frac{d^3kd\omega}{2E(k)(\omega+i\epsilon)}[e^{-iE(k)t+i\vec{k}\cdot\vec{x}-i\omega t}+e^{iE(k)t-i\vec{k}\cdot\vec{x}+i\omega t}] \tag{6.72}$$

The Fourier transform is

$$\begin{aligned} -i\Delta_F(q) &= \int d^4x(-i\Delta_F(x))e^{iqx} \\ &= i\int \frac{d^4x}{(2\pi)^4}\frac{d^3kd\omega}{2E(k)(\omega+i\epsilon)}[e^{-iE(k)t+i\vec{k}\cdot\vec{x}-i\omega t}+e^{iE(k)t-i\vec{k}\cdot\vec{x}+i\omega t}]e^{iqx} \end{aligned} \tag{6.73}$$

and so $\vec{q}=\pm\vec{k}$ from the $\vec{x}$ integral (the sign doesn't matter). The $\int dx^0$ gives a $2\pi\delta(q^0-\omega-E(k))$ for the first term and a $2\pi\delta(q^0+\omega+E(k))$ for the second one. The expression collapses to

$$-i\Delta_F(q) = i\int \frac{d\omega}{2E(q)(\omega+i\epsilon)}[\delta(\omega+E(q)-q^0)+\delta(\omega+E(q)+q^0)] \tag{6.74}$$

and so we can write the result in two equivalent ways: First, separate the terms:

$$-i\Delta_F(q) = \frac{i}{2E(q)}[\frac{1}{q^0 - E(q) + i\epsilon} + \frac{1}{-q^0 - E(q) + i\epsilon}] \tag{6.75}$$

and second, combine them,

$$-i\Delta_F(q) = \frac{i}{q^2 - m^2 + i\epsilon}. \tag{6.76}$$

This is a very important object, the propagator for a scalar field.

There is an associate Feynman diagram, Fig. 6.2.

We are not going to use Eq. 6.75 any more, but let's look at it before moving on. The first term is the antiparticle piece, where the time ordering is that first the antiparticle is produced, and later it annihilates to produce the outgoing BC pair. The second term is the opposite time order, emission preceded absorption. The energy denominators are the ones in the "usual" quantum mechanical second-order perturbation theory formula for the T-matrix

$$T = \sum_I \frac{\langle f|V|I\rangle \langle I|V|i\rangle}{E_i - E_I}. \tag{6.77}$$

This you can see if you write $\vec{q} = \vec{k}_b + \vec{k}_c = -p_a$, $q^0 = E_b + E_c$, and change notation so $E_a = E(q)$. Then

$$M \propto 0g^2[\frac{i}{2E_a}\{\frac{1}{E_b + E_c - E_a + i\epsilon} + \frac{1}{-E_a - E_b - E_c + i\epsilon}\}] \tag{6.78}$$

The second term is trickier until you realize that the initial energy is $E_b + E_c$ and the intermediate energy $E_I = E_b + E_c + E_{b'} + E_{c'} + E_a$ so the denominator is also the correct energy difference.

In nonrelativistic situations, the intermediate states from the different time orderings are just different, and you can't use all our tricks to get a beautiful answer, like Eq. 6.76. This is all people had, before Feynman (1948-ish). How would you like to do quantum field theory this way? It was horrible... Imagine higher order calculations! Lorentz invariance is hidden, too. There is an old book, Heitler, "The quantum theory of radiation," take a look at it, on a rainy day.

Now go back to the S-matrix

$$\begin{aligned} S(k_b + k_c \to k'_b + k'_c) &= \frac{(-i)^2}{2!}\int d^4x d^4y \frac{e^{-i(k_b+k_c)y} e^{-i(k_{b'}+k_{c'})x}}{(2\pi)^6}[-i\Delta_F(x-y)] \\ &= -g^2\frac{(2\pi)^4}{(2\pi)^6}\delta^4(k_{b'} + k_{c'} - k_b - k_c)\frac{i}{(k_{b'} + k_{c'})^2 - m_a^2 + i\epsilon}] \end{aligned} \tag{6.79}$$

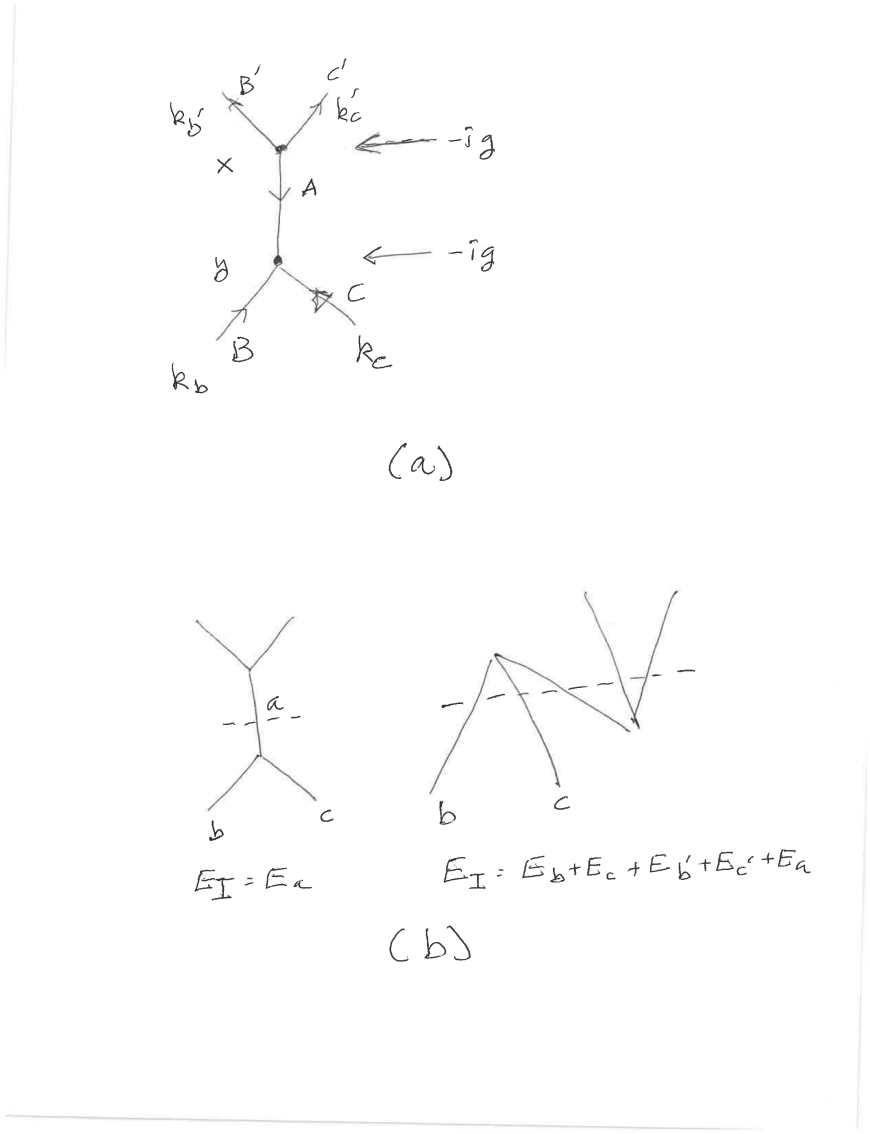

Figure 6.2: Lowest order scattering Feynman diagram for $BC \to BC$ with $\mathcal{H} = ABC + h.c.$

Here, $k_a = k_b + k_c = k_{b'} + k_{c'}$ so four-momentum is conserved at each vertex, but $k_a^2 = (k_{b'} + k_{c'})^2 \neq m_a^2$. We say "the intermediate particle is off mass shell."

Finally, what's the cross section? In a CM frame, $(k_b + k_c)^2 = E_{CM}^2$ and

$$\frac{d\sigma}{d\Omega} \sim |\frac{1}{E_{CM}^2 - m_a^2}|^2 \tag{6.80}$$

This diverges at $E_{CM} = m_a$. This is the mark of a resonance, when the CM energy is exactly right to produce the intermediate particle on shell, the cross section becomes very large. The divergence is not physical, it is an artifact of working to lowest order the perturbative expansion, where the resonance has zero width. The resonance is the physics!

6.6 Example 3: Another amplitude

For our next example, keep $\mathcal{H} = g(ABC + A^\dagger B^\dagger C^\dagger)$, but compute the reaction $B\bar{C} \to B\bar{C}$.

$$S = -g^2 \int\int d^4x d^4y e^{-(k_b - k_{c'})x} e^{-(k_{b'} - k_c)y} \langle 0|T(A(x)^\dagger A(y))|0\rangle \tag{6.81}$$

This gives

$$M = -g^2 \frac{i}{(k_b - k_{c'})^2 - m_a^2 + i\epsilon} \tag{6.82}$$

Let's write the S-matrix a little more suggestively,

$$\begin{aligned} S &= -g^2 \int\int d^4x d^4y e^{-(k_b - k_{c'})ix} e^{-i(k_{b'} - k_c)y} \int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m_a^2 + i\epsilon} e^{iq(x-y)} \\ &= -g^2 \frac{1}{(2\pi)^6} \int \frac{d^4q}{(2\pi)^4} [(2\pi)^4 \delta^4(k_b - k_{c'} + q)][(2\pi^4)\delta^4(k_{b'} - k_c + q)] \frac{i}{(k_b - k_{c'})^2 - m_a^2 + i\epsilon} \end{aligned} \tag{6.83}$$

Writing this expression tells us that we have conservation of four-momentum at each vertex. We also integrate over all internal momenta (in this case, just q). One of the delta-functions will give overall four momentum conservation when all the integrals are done.

The associated Feynman diagram is shown in Fig. 6.3.

6.7 Wick's theorem

Wick's theorem is used to reduce the nasty time ordered products of field operators into simple expressions, by recasting the time ordered product of many fields into a product

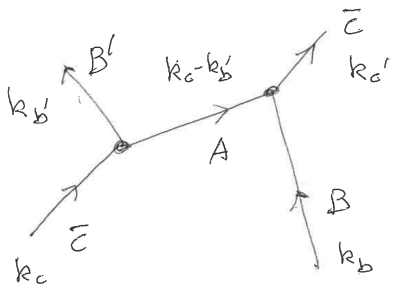

Figure 6.3: Lowest order scattering Feynman diagram for $B\bar{C} \to B\bar{C}$ with $\mathcal{H} = ABC + h.c.$

of time ordered products of two fields (which are just propagators). Let me describe it for bosons; fermions have some permutation factors (no surprise). Write each field in its creation and annihilation parts, $\phi = \phi^+ + \phi^-$ where $\langle 0| \phi^+ = \phi^- |0\rangle = 0$. Recall the normal ordering operator,

$$: \phi_i \phi_j := \phi_i^+ \phi_j^+ + \phi_i^+ \phi_j^- + \phi_j^+ \phi_i^- + \phi_i^- \phi_j^- . \tag{6.84}$$

We can write the time ordered product of two fields as the normal ordered product plus a constant C, since the rearrangements needed to go from one to the other just involve commutators of fields.

$$T(\phi_1 \phi_2) =: \phi_1 \phi_2 : + C. \tag{6.85}$$

Now take the vacuum expectation value of this expression. The vacuum expectation value of the normal ordered product vanishes, so

$$\langle 0|T(\phi_1 \phi_2)|0\rangle = 0 + C. \tag{6.86}$$

and so we have

$$T(\phi_1 \phi_2) =: \phi_1 \phi_2 : + \langle 0|T(\phi_1 \phi_2)|0\rangle . \tag{6.87}$$

This is for two fields. Wick's theorem generalizes this to n fields. For even n, it says

$$\begin{aligned} T(\phi_1 \phi_2 \dots \phi_n) \;=\; & : \phi_1 \phi_2 \dots \phi_n : \\ & + \langle 0|T(\phi_1 \phi_2)|0\rangle : \phi_3 \dots \phi_n : + \text{ order } n-2 \text{ permutations} \\ & + \sum \langle 0|T(\phi_i \phi_j)|0\rangle \, \langle 0|T(\phi_k \phi_l)|0\rangle : \phi \dots \phi : \\ & + \dots \\ & + \sum \langle 0|T(\phi_i \phi_j)|0\rangle \, \langle 0|T(\phi_k \phi_l)|0\rangle \dots \langle 0|T(\phi_m \phi_n)|0\rangle \end{aligned} \tag{6.88}$$

The proof is not so interesting. It is done by induction, and it is in Bjorken and Drell, "Relativistic quantum fields" (and probably other books). But the formula is quite useful. Return to $\mathcal{H} = ABC + h.c.$ where we need

$$\begin{aligned} \langle B' \bar{C}' |T(\mathcal{H}(x)\mathcal{H}(y)|B\bar{C}\rangle = \langle B'\bar{C}| \{ & \quad : A_i^\dagger B_1^\dagger C_1^\dagger A_2 B_2 C_2 : \\ & + \langle 0|T(A_1^\dagger A_2|0\rangle : B_1^\dagger C_1^\dagger B_2 C_2 : \\ & + \langle 0|T(B_1^\dagger B_2|0\rangle : A_1^\dagger C_1^\dagger A_2 C_2 : \\ & \quad + \dots \\ & \} \, |B\bar{C}\rangle \end{aligned} \tag{6.89}$$

Only the $\langle B'\bar{C}| : B_1^\dagger C_1^\dagger B_2 C_2 : |B\bar{C}\rangle$ term has the right combination of operators to annihilate $B\bar{C}$ and create $B'\bar{C}'$. Everything else gives zero.

6.8 A glance at higher orders

We return to

$$\mathcal{H} = \frac{g}{4!}\phi^4(x) \tag{6.90}$$

where ϕ is real, so $a = a_c$. Let us look at the order g^2 contributions to scatttering $\phi_1 + \phi_2 \to \phi_3 + \phi_4$.

$$S^{(2)} = \frac{1}{2!}\left(\frac{-ig}{4!}\right)^2 \int d^4x d^4y \, \langle \Phi_{34}|T(\phi(x)^4\phi(x)^4)|\Phi_{12}\rangle \tag{6.91}$$

When we use Wick's theorem to clean this up, we want to pull off terms with normal products of four fields.

$$T(\phi(x)^4\phi(x)^4) = \sum_{ijkl} \langle 0|T(\phi_j\phi_k)||0\rangle \, \langle 0|T(\phi_k\phi_l)|0\rangle : \phi_m\phi_n\phi_o\phi_p : \tag{6.92}$$

Four of $(ijklmnop)$ are at x, the other four are at y.

We have three generic kinds of terms (see Fig. 6.4):

- a) $i = k = x$, $j = l = y$ and permutations
- b) $i = x$, $j = y$, $k = l = x$ OR y
- c) In addition, we may have $i = j = x$, $k = l = y$. These are called "disconnected diagrams." They don't contribute to scattering (they won't be proportional to $\delta^4(k_1 + k_2 - k_3 - k_4)$. We will neglect them.

There are three "type (a)" graphs. They all have a common combinitorial factor C, which we will ignore for now. For eaxmple,

$$\begin{aligned} S_{a_1 a} &= C(-ig)^2 \int \frac{d^4q_1}{(2\pi)^4}\frac{d^4q_2}{(2\pi)^4}(2\pi)^4\delta^4(q_1 + q_2 - k_1 - k_3) \\ &\quad \times(2\pi)^4\delta^4(q_1 + q_2 - k_2 - k_4) \times \frac{-i}{q_1^2 - m^2 + i\epsilon} \times \frac{-i}{q_2^2 - m^2 + i\epsilon} \end{aligned} \tag{6.93}$$

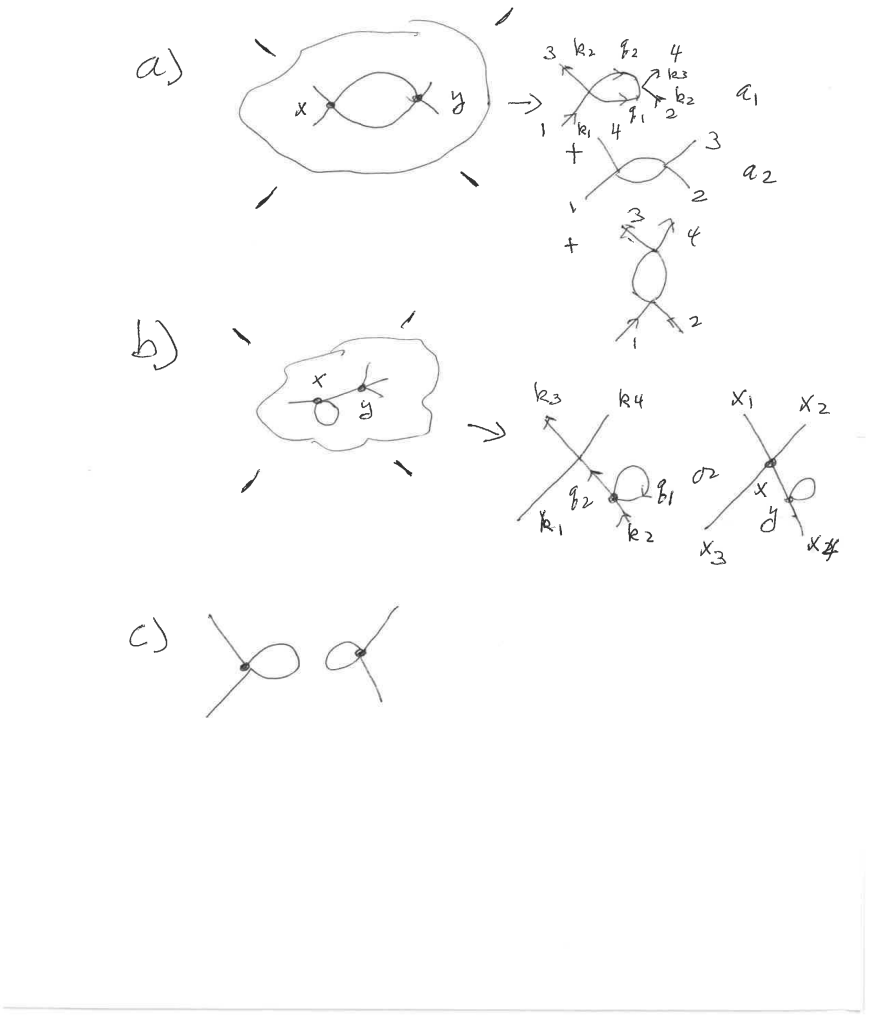

Figure 6.4: Higher order scattering processes with $\mathcal{H} = g\phi^4/4!$.

so that the invariant amplitude is

$$M_{a1} = -Cg^2 \int \frac{d^4q_1}{(2\pi)^4}[\frac{-i}{q_1^2 - m^2 + i\epsilon}][\frac{-i}{(k_1 - k_3 - q_1)^2 - m^2 + i\epsilon}]. \tag{6.94}$$

This is called a graph with a "loop integral." The momentum q_1 is unconstrained and must be integrated over.

There are four type (b) graphs.

$$\begin{aligned}(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)M_{b_1} \quad &\propto \quad (-ig)^2 \int \frac{d^4q_1}{(2\pi)^4}\frac{d^4q_2}{(2\pi)^4}(2\pi)^4\delta^4(q_2 - k_2)\delta^4(k_3 + k_4 - q_2 - k_1) \\ &\times \frac{-i}{q_1^2 - m^2 + i\epsilon} \times \frac{-i}{q_2^2 - m^2 + i\epsilon}\end{aligned} \tag{6.95}$$

This says that

$$M_{b_1} \propto (-ig)^2 \frac{-i}{k_2^2 - m^2 + i\epsilon} \int \frac{d^4q_1}{(2\pi)^4} \frac{-i}{q_1^2 - m^2 + i\epsilon}. \tag{6.96}$$

This is getting mechanical, but the result is worrisome: the integral is divergent!

$$\int_{|q|<\Lambda} \frac{d^4q_1}{(2\pi)^4} \frac{-i}{q_1^2 - m^2 + i\epsilon} \propto \int^{\Lambda} \frac{d^3q}{q^2} \propto \Lambda^2 \tag{6.97}$$

where I have put in a cutoff to make sense of the expression.

This is our first encounter with the "divergences" of quantum field theory. They arise from the short distance sector of the theory. This is most easily seen by working in coordinate space and keeping all the propagators,

$$\begin{aligned}M(x_1, x_2, x_3, x_4) \quad &\propto \quad (-ig)^2 \int d^4x d^4y [-i\Delta_F(x_1 - x)][-i\Delta_F(x_2 - x)][-i\Delta_F(x_3 - x)] \\ &\quad [-i\Delta_F(x - y)][-i\Delta_F(y - x_4)][-i\Delta_F(y - y)]\end{aligned} \tag{6.98}$$

The divergence comes from the last term: emission and absorption at the same point,

$$-i\Delta_F(x = 0) = \int \frac{d^4q_1}{(2\pi)^4} \frac{-i}{q_1^2 - m^2 + i\epsilon}. \tag{6.99}$$

This contribution is present even though we might only be interested in scattering (or other processes) taking place at very low energies, or at long distances. To proceed further at this point, we have to do several things:

1. Enumerate all possible divergent contributions: are there a finite number, or an infinite number? It turns out that in ϕ^4 theory, only two point functions and four point functions contain divergences.

2. Regulate the divergences: do something to make them well defined. In the example, I cut off the momentum integral, restricting $|q| < \Lambda$.

3. Do calculations in the regularized theory and make sense of them.

This is a project for later in the semester.

6.9 Feynman rules

You may have noticed a certain robotic similarity in all our examples. All our manipulations can be absorbed into a set of rules for constructing invariant amplitudes, in terms of Feynman diagrams.

Here are the Feynman rules for scalar field theory:

Given an interaction Hamiltonian density which is a product of fields (for example $\mathcal{H} = g\phi^4/4!$)

1) Draw all connected diagrams allowed by $\mathcal{H}$. For each vertex there is a factor $-i$, a factor of the coupling constant, and there is an overall $1/n!$ from the Dyson series. Each diagram will typically have some associated combinatorial factor.

2) Each internal boson carrying a momentum q contributes a $-i\Delta_F(x-y)$ in coordinate space, or

$$\int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\epsilon} \tag{6.100}$$

in momentum space.

3) Each vertex counts either (orient all the momenta in or out for this)

$$(2\pi)^4 \delta^4(\sum q) \tag{6.101}$$

or

$$\int d^4x e^{i\sum qx} \tag{6.102}$$

As an alternative to rules 2 and 3,

2'-3') Conserve momentum at every vertex, integrate

$$\int \frac{d^4q}{(2\pi)^4} \tag{6.103}$$

over all unconstrained momenta

And compute the invariant amplitude from the resulting expression by snipping off delta functions and 2π's associated with the incoming and outgoing particles

$$M\frac{(2\pi)^4}{(2\pi)^{3(n_i+n_f)/2}}\delta^4(\sum p_{out} - \sum p_{in}) \to M \tag{6.104}$$

The differential cross section is

$$d\sigma = \frac{1}{4E_1E_2v_{rel}}|M|^2(2\pi)^4\delta^4(\sum k_f - \sum k_i)\prod_{j=1}^{n_f}\frac{d^3k_j}{(2\pi)^3 2E(k_j)} \times \frac{1}{n_f!} \tag{6.105}$$

The new $1/n_f!$ is a counting factor for identical particles, basically Bose statistics. It is discussed in Bjorken and Drell, "Relativistic Quantum Mechanics," p. 136, or in Zee, a hint on p. 54.

6.10 Conclusion

This chapter has been a terse but fairly complete introduction to quantum field theory based on Hamiltonians. We only looked at scalar particles. The interesting (and more physical, at least before 2012) quantum field theories contain fermions and gauge fields. Perturbation theory for such systems is very similar to what we have done: there will be vertices and propagators, which only differ in detail (important details, true) from what we have found.

However, before we push on to these systems, I want to go back to the drawing board. You recall, that there are two ways of introducing quantum mechanics, either with Hamiltonians, or with the path integral. Path integrals are not so useful for practical problems in single particle quantum mechanics, but they really come into their own when they are applied to quantum field theories. We have to become comfortable with them. That is the subject of the next chapter.

Chapter 7

Path integrals in quantum mechanics and quantum field theories

7.1 Path integrals in quantum mechanics

This section is taken from my quantum mechanics notes.

What if H varies with time? How can we make sense of the time evolution operator? Let us begin with the formal expression for the evolution operator

$$U_{\alpha\beta}(t) = \langle\alpha| \exp(-\frac{i\hat{H}t}{\hbar})|\beta\rangle \,. \tag{7.1}$$

In this formula t is some finite quantity, so it might be hard to compute $U_{\alpha\beta}(t)$. We propose to evaluate the time evolution operator by slicing the time interval into a series of N steps of infinitesimal time interval Δt, arranging that $t = N\Delta t$. In each time interval Δt, the Hamiltonian will be regarded as constant, while of course it will be allowed to vary from time step to time step. That is,

$$U_{\alpha\beta}(t) = \lim_{N\to\infty;\Delta t\to 0} \langle\alpha| \prod_{i=1}^{N} \exp(-\frac{i\hat{H}(t_i)\Delta t}{\hbar})|\beta\rangle \,. \tag{7.2}$$

Inserting a complete set of states between each exponential factor, we have

$$\begin{aligned} U_{\alpha\beta}(t) \;=\; & \sum_{j_1}\sum_{j_2}\cdots\sum_{j_N} \\ & \langle\alpha|j_N\rangle \langle j_N| \exp(-\frac{i\hat{H}(t_N)}{\hbar}\Delta t)|j_{N-1}\rangle \\ & \langle j_{N-1}| \exp(-\frac{i\hat{H}(t_{N-1})}{\hbar}\Delta t)|j_{N-2}\rangle \cdots \\ & \langle j_1| \exp(-\frac{i\hat{H}(t_1)}{\hbar}\Delta t)|\beta\rangle \,. \end{aligned} \tag{7.3}$$

Rather than thinking of this expression time slice by time slice, let us link together a particular set of states $|j'_1\rangle$, $|j'_2\rangle$, $\ldots |j'_N\rangle$, and connect them together as a "path" in Hilbert space, as shown in Fig. 7.1. Each path contributes a complex number to $U_{\beta\alpha}$, the product of each particular matrix element $\langle j'_M| \exp(-iH(t_M)\Delta t/\hbar)|j'_{M-1}\rangle$. $U_{\beta\alpha}$ is the sum of contributions from all paths: hence the name "path integral" associated with this representation of the time evolution operator.

Let us suppose that our intermediate states are diagonal in coordinate space, so that we

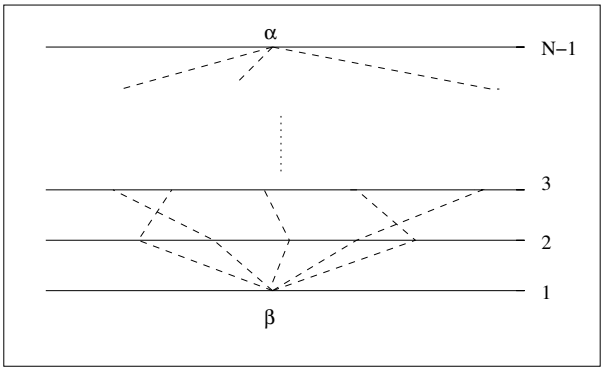

Figure 7.1: Paths in the time-sliced interval.

can interpret each path through Hilbert space as a "real" trajectory, $x(t)$. Then,

$$U_{\alpha\beta}(t) = \int dx_0 \int dx_N \cdots \int dx_1 \, \langle \alpha | x_N \rangle \, \langle x_N | \exp(-\frac{i\hat{H}(t_N)}{\hbar}\Delta t) | x_N \rangle \times \tag{7.4}$$

$$\langle x_N | \exp(-\frac{i\hat{H}(t_{N-1})}{\hbar}\Delta t) | x_{N-1} \rangle \cdots \langle x_1 | \exp(-\frac{i\hat{H}(t_0)}{\hbar}\Delta t | x_0 \rangle) \, \langle x_0 | \beta \rangle \, . \tag{7.5}$$

The time slicing helps us make sense of the evolution operator. Suppose next that the Hamiltonian is that of a particle in an external potential, $\hat{H} = \hat{H}_1 + \hat{H}_2$, where $\hat{H}_1 = \frac{\hat{p}^2}{2m}$ and $\hat{H}_2 = V(x)$. Then we can break the time evolution operator over a time Δt (which involves $\hat{H}$) into a product of two terms

$$\exp(-\frac{i\hat{H}\Delta t}{\hbar}) = \exp(-\frac{i\hat{H}_1\Delta t}{\hbar}) \exp(-\frac{i\hat{H}_2\Delta t}{\hbar}) + O(\Delta t^2). \tag{7.6}$$

This means that

$$\begin{aligned} \langle x_1 | \exp -\frac{i\hat{H}_1\Delta t}{\hbar} \exp -\frac{i\hat{H}_2\Delta t}{\hbar} | x_2 \rangle &= \int dx_3 \, \langle x_1 | \exp(-\frac{i\hat{p}^2\Delta t}{2m\hbar}) | x_3 \rangle \\ &\quad \times \langle x_3 | \exp(-\frac{iV(x)\Delta t}{\hbar}) | x_2 \rangle \\ &= \int dx_3 \, \langle x_1 | \exp(-\frac{i\hat{p}^2\Delta t}{2m\hbar}) | x_3 \rangle \\ &\quad \times \delta(x_3 - x_2) \exp(-\frac{iV(x_2)\Delta t}{\hbar}). \end{aligned} \tag{7.7}$$

Notice that the term first term is just the propagator for a free particle, and so

$$\begin{aligned}\langle x_1|\exp(-\frac{i\hat{H}\Delta t}{\hbar})|x_2\rangle &= \int dx_3 K(x_1,\Delta t;x_3,0)\delta(x_3-x_2)\exp(-\frac{iV(x_2)\Delta t}{\hbar}) \\ &= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{1}{2}} \exp[i\left(\frac{m(x_1-x_2)^2}{2(\Delta t)^2}-V(x_2)\right)\frac{\Delta t}{\hbar}].\end{aligned} \tag{7.8}$$

Thus the path integral will take the form

$$\begin{aligned}U_{\alpha\beta}(t) &= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{N}{2}}\int dx_N\cdots dx_1\,\langle\alpha|x_N\rangle\,\langle x_1|\beta\rangle \\ &\quad \exp\left(\frac{i\Delta t}{\hbar}\sum_{j=1}^{N}\left(\frac{1}{2}m\left(\frac{x_{j+1}-x_j}{\Delta t}\right)^2-V(x_j)\right)\right).\end{aligned} \tag{7.9}$$

(We have factorized the Δt dependence in the expression in a useful way.) In the limit that $\Delta t\to 0$ the exponential factor associated with a single path $x(t')$ becomes

$$\exp\left(\frac{i}{\hbar}\int_0^t dt'\left(\frac{1}{2}m\dot{x}^2(t')-V(x(t'))\right)\right). \tag{7.10}$$

The reader should recognize the integrand as the classical Lagrangian for a nonrelativistic particle in an external potential., The integral is the classical action associated with its integral along a path specified by $x(t')$. This is the expression for the contribution to the path integral from a single path (a single term in the sum over intermediate states as defined in Eq. 7.3.) Each contribution is a phase, whose value is the ratio of the classical action for that particular path divided by $\hbar$. The time evolution operator is a sum of contributions over all possible paths the particle can take from the initial time to the final time.

We can now understand how classical dynamics can arise from quantum mechanical motion. Suppose that the action for every possible path the particle could take is large, much greater than $\hbar$. Compare two paths (which could correspond to paths which are similar in the classical sense). An action S_0 is associated with one path while the other has a different action, which we will write as $S_0+\delta S$. These two paths combine to produce a contribution to the evolution operator of

$$\exp\frac{iS_0}{\hbar}+\exp\frac{i(S_0+\delta S)}{\hbar}=\exp\frac{iS_0}{\hbar}\left(1+\exp\frac{i\delta S}{\hbar}\right). \tag{7.11}$$

If $\delta S/\hbar$ is a large number, the phase difference between the paths, $\exp(i\delta S/\hbar)$, will not be small. There will be destructive interference between the two paths and their contribution

to the transition amplitude will be small. Obviously, the only sets of paths which contribute will be the ones for which their contributions will add coherently. These will be the ones for which $\delta S \simeq 0$, This is the statement of the principle of least action: the dominant paths are the ones which extremize the classical action – that is, the particle follows the path which is the solution to the classical equations of motion.

As a contrast, it could be the case that all paths have a classical action which is on the order of, or smaller than, $\hbar$. Then there are no sets of nearby paths which dominate the evolution operator. These systems are fully quantum mechanical; attempting to describe a particle's motion using the idea of a classical trajectory will not result in correct physics.

7.2 Path integrals in quantum field theory (for bosons)

Once again, quantum field theory is quantum mechanics with lots of indices and some relabelling. I specialize to the case of quantum field theories containing bosons (but without fermions – see below). Our variables are the $\phi(x,t)$'s and the $\pi(x,t)'s$, defined to obey canonical commutation relations

$$[\pi(x,t), \phi(x',t)] = -i\delta^3(x-x'). \tag{7.12}$$

The x_i's or $x(t)$'s of the quantum mechanical path integral are replaced by the $\phi(x,t)$'s. We replace the integration measure

$$\prod_i dx_i \to \prod_{x_j,t_i} \phi(x_j,t_i) \equiv [d\phi]. \tag{7.13}$$

If $\phi(x,t)$. has components, we must integrate over them, too. For example, if ϕ is a complex field we must integrate over its real part and its imaginary part.

Next we have to deal with the in and out states, $|\alpha\rangle$ and $|\beta\rangle$. There are many choices here! Most often, we take $|\alpha\rangle = |\beta\rangle = |0\rangle$, the vacuum. Why? The analogy is with a

correlation function for a quantum mechanical system in its ground state, which is

$$\begin{aligned}\langle 0|T(q(t_A)q(t_B))|0\rangle &= N\int dq(t_0)dq(t)\prod_{t_i}dq(t_i)\,\langle 0|q(t)\rangle\left\langle q(t)|e^{iH\Delta t}|q(t_n))\right\rangle \\ &\quad\times\ldots\left\langle q(t_A)|q(t_A)e^{iH(t_A)\Delta t)}|q(t_A-\Delta t)\right\rangle \\ &\quad\times\ldots\langle q(t_0)|0\rangle \\ &= N'\int\prod_i dq_i e^{iS}q(t_A)q(t_B).\end{aligned} \tag{7.14}$$

In quantum field theory the analogous object is $\langle 0|T(\phi(x_1)\phi(x_2)|0\rangle$, which will give us Green's functions. Just copying formulas we have already written down and changing labels, the matrix element of an arbitrary operator between $|0\rangle$. the vacuum at $t=-\infty$ and $\langle 0|$, the vacuum at $t=\infty$, is

$$\langle 0|\mathcal{O}(\phi(x_1)\phi(x_2),\ldots\phi(x_n))|0\rangle = N\int[d\phi]\exp(i\int\mathcal{L}d^4x)|\mathcal{O}(\phi(x_1)\phi(x_2),\ldots\phi(x_n)). \tag{7.15}$$

that is, just the average of the operator weighted by the measure $\exp(i\int\mathcal{L}d^4x)$.

What is N? We can find it with a trick – calculate $\langle 0|1|0\rangle = 1$. Solving for N in Eq. 7.15,

$$\langle 0|\mathcal{O}(\phi(x_1)\phi(x_2),\ldots\phi(x_n))|0\rangle = \frac{\int[d\phi]e^{iS(\phi)}\mathcal{O}(\phi(x_1)\phi(x_2),\ldots\phi(x_n))}{\int[d\phi]e^{iS(\phi)}} \tag{7.16}$$

This is a story you have seen before, in statistical mechanics. The density matrix $\rho(p,q)$ gives the thermal probability of occupancy,

$$\rho = Ne^{-\beta H}. \tag{7.17}$$

Thermal expectation values are

$$\langle\mathcal{O}\rangle = \mathrm{Tr}\,\rho O = N\mathrm{Tr}\,e^{-\beta H}\mathcal{O} \tag{7.18}$$

and of course $\langle 1\rangle = \mathrm{Tr}\,\rho = N\mathrm{Tr}\,e^{-\beta H}$ so

$$\langle\mathcal{O}\rangle = \frac{\mathrm{Tr}\,e^{-\beta H}\mathcal{O}}{e^{-\beta H}} \tag{7.19}$$

and $\rho = e^{-\beta H}/\mathrm{Tr}\,e^{-\beta H}$.

The analogy is even more precise with the so-called "Euclidean path integral." Consider a Lagrange density

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial_\mu\phi)^2 - V(\phi) \\ &= \frac{1}{2}(\frac{\partial\phi}{\partial t})^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi).\end{aligned} \tag{7.20}$$

Change variables from t to $x_0 = it$. This makes two changes in familiar formulas,

$$\begin{aligned} i\int dt &= \int dx_0 \\ \frac{1}{2}(\frac{\partial\phi}{\partial t})^2 &= -\frac{1}{2}(\frac{\partial\phi}{\partial x_0})^2\end{aligned} \tag{7.21}$$

Then

$$\langle 0|\mathcal{O}|0\rangle = \frac{\int[d\phi]\mathcal{O}\exp(-\int d^4x[\sum_j \frac{1}{2}(\frac{\partial\phi}{\partial x_j})^2 + V(\phi)])}{\int[d\phi]\exp(-\int d^4x[\sum_j \frac{1}{2}(\frac{\partial\phi}{\partial x_j})^2 + V(\phi)])}. \tag{7.22}$$

There is a complete analogy of bosonic quantum field theory in its path integral form with the partition function for a spin model in $d+1$ (spatial) directions:

- $\phi(x,t) \leftrightarrow$ a classical spin variable defined on every site in a $d+1$ dimensional space
- action $\leftrightarrow \beta H(\phi)$
- (or alternatively) Euclidean Lagrangian $\leftrightarrow \beta\times$ Hamiltonian density
- Green's function $\leftrightarrow$ an expectation value of some function of ϕ's (a correlation function).

Path integrals for fermionic theories are a bit more complicated, because fermionic operators obey anticommutation relations rather than commutation relations. The field variables in the path integral must encode this fact. The objects that do this are called "Grassman variables" or "anticommuting c-numbers." For us, living in a too-short one-semester course, this is a technical difficulty which I will skip over due to time constraints, and I will not discuss the fermion path integral any further. Look in any good quantum field theory text if you are interested.

7.3 The propagator and Feynman rules, again

(This section is copied from a section in the lecture "Secret Symmetry" in Sidney Coleman's "Aspects of Symmetry" and from Pierre Ramond's book.) We would like to calculate the following integral

$$\begin{aligned} \rho &= \int_{-\infty}^{\infty} \prod_{i=1}^{N} dx_i \exp(-\frac{1}{2} x_i M_{ij} x_j) \\ &\equiv \int \prod dx_i \rho(x) \end{aligned} \tag{7.23}$$

(Einstein convention summing repeated indices, of course).

Start with $N = 1$. This is

$$\begin{aligned} \rho &= \sqrt{\frac{2\pi}{M}} \\ \langle x^2 \rangle &\equiv \frac{\int x^2 \rho(x) dx}{\int dx \rho(x)} = M^{-1}. \end{aligned} \tag{7.24}$$

This formula is obviously true when M is real and it is also true for complex M whenever Re $M > 0$.

We can generalize this formula to N dimensions if we think x as a vector and M as a matrix in that space. Then if M is real, symmetric, and positive definite, we can make a rotation which diagonalizes M

$$\begin{aligned} x_i &= R_{ij} y_j \\ \prod dx_i &= \prod dy_i \\ x_i M_{ij} x_j &= y_l R^{-1}_{li} M_{ij} R_{jk} y_k = y_l M'_{ll} y_l \end{aligned} \tag{7.25}$$

so

$$\begin{aligned} \rho &= (2\pi)^{N/2} (\prod_l M'_{ll})^{-1/2} \\ &= (2\pi)^{N/2} (\det M)^{-1/2}. \end{aligned} \tag{7.26}$$

Again, if M is a complex symmetric matrix with positive-definite real part, the result is also this formula, by analytic continuation.

Now suppose we want to compute

$$\langle x_i x_j \rangle = \frac{\int \prod dx_k (x_i x_j) \rho(x)}{\int \prod dx_k \rho(x)}. \tag{7.27}$$

We can do this integral by a trick. To begin, define

$$(dx) = \prod_{k=1}^{N} dx_k (2\pi)^{-N/2} \tag{7.28}$$

so

$$\int (dx) \exp(-\frac{1}{2} x_i M_{ij} x_j) = (\det M)^{-1/2}. \tag{7.29}$$

Now consider the integral

$$\int (dx) \exp(-Q(x)) \tag{7.30}$$

where $Q(x) = \frac{1}{2}(x, Mx) + (b, x) + c$ and $(,)$ represents an abstract dot product, b is a vector, $(b, x) = \sum_j b_j x_j$, and c is a constant. We evaluate the integral as we do for any Gaussian, by completing the square:

$$\begin{aligned} \bar{x} &= -M^{-1} b \\ Q(x) &= Q(\bar{x}) + \frac{1}{2}(x - \bar{x}, M(x - \bar{x})) \\ Q(\bar{x}) &= \frac{1}{2}(b, M^{-1} b) + c \end{aligned} \tag{7.31}$$

so

$$\int (dx) \exp(-Q(x)) = \exp(-Q(\bar{x}))(\det M)^{-1/2}. \tag{7.32}$$

Now notice:

$$\langle x_i x_j \rangle = \frac{\int (dx) \exp(-\frac{1}{2}(x, Mx)) x_i x_j}{\int (dx) \exp(-\frac{1}{2}(x, Mx))} \tag{7.33}$$

and if we write

$$S(J) = \frac{1}{2}(x, Mx) + (J, x), \tag{7.34}$$

this is

$$\langle x_i x_j \rangle = \frac{\frac{\partial}{\partial J_i} \frac{\partial}{\partial J_j} \int (dx) \exp(-S(J))}{\int (dx) \exp(-S(J))} |_{J=0}, \tag{7.35}$$

which evaluates to

$$\frac{\frac{\partial}{\partial J_i}\frac{\partial}{\partial J_j}\exp(-\frac{1}{2}(J,M^{-1}J))(\det M)^{-1/2}}{\exp(-\frac{1}{2}(J,M^{-1}J))(\det M)^{-1/2}}|_{J=0} = \frac{1}{2}\times 2(M^{-1})_{ij}. \tag{7.36}$$

(The extra "2" is because i can be the index of the first J or of the second J in the expression – M is symmetric.)

Well, so what? These formulas are true regardless of the dimension of the space, and we propose to apply them to spaces of infinite dimensionality. The x variables are our field variables $\phi(x,t)$ defined everywhere in space and time. For continuous variables in D dimensions we define

$$\begin{aligned}(J,\phi) &= \int d^D x J(x)\phi(x)\\ (\phi,M\phi) &= \int d^D x \int d^D x' \phi(x)M(x,x')\phi(x').\end{aligned} \tag{7.37}$$

Generally, M will be a differential operator and M^{-1} will be its inverse. We use the sources $J(x,t)$ as crutches to construct Green's functions.

So let's look once again at scalar field theory with an action

$$S = \int d^4x[\frac{1}{2}(\partial_\mu\phi)^2 - V(\phi)] \tag{7.38}$$

with $V(\phi) = \frac{1}{2}m^2\phi^2 +$ some higher order $\bar{V}(\phi)$. Green's functions are defined as (the middle column is the Hamiltonian version, the right column is the path integral version)

$$\begin{aligned}G^{(2)}(x_1,x_2) &= \langle 0|T(\phi(x_1)\phi(x_2))|0\rangle = \langle\phi(x_1)\phi(x_2)\rangle\\ G^{(4)}(x_1,x_2,x_3,x_4) &= \langle 0|T(\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4))|0\rangle = \langle\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle\end{aligned} \tag{7.39}$$

and we compute them through an object called the "generating functional,"

$$W(J) = \frac{\int[d\phi]\exp(i\int d^4x[\mathcal{L}+J(x)\phi(x)])}{\int[d\phi]\exp(i\int d^4x\mathcal{L})} \tag{7.40}$$

via

$$G^{(N)} = (-i)^N\frac{\partial}{\partial J_1}\frac{\partial}{\partial J_2}\cdots\frac{\partial}{\partial J_N}W(J)|_{J=0}. \tag{7.41}$$

As a check, doing the differentiation before evaluating $W(J)$ gives

$$G^{(2)}(x_1, x_2) = \frac{\int [d\phi \exp(i \int d^4x \mathcal{L}) \phi(x_1)\phi(x_2)}{\int [d\phi] \exp(i \int d^4x \mathcal{L})}, \tag{7.42}$$

as expected.

If you look closely at $W(J)$, you see that the integral is highly oscillatory. There are two ways to deal with this:

- Include a convergence factor in the weight of $\exp(-\frac{1}{2}\epsilon \int d^4x \phi(x)^2)$
- Define W in Euclidean space, $x_0 = -i\bar{x}_4$, $d^4x = -id^4\bar{x}$, $\partial_\mu\phi\partial^\mu\phi \to -\sum_{j=1}^4 (\partial_j\phi)^2$ so

$$W_E(J) \propto \int [d\phi] \exp(-\int d^4x [\frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\phi^2 + \bar{V}(\phi)] \tag{7.43}$$

  Now the Boltzmann factor really is a suppression factor.

Most formal quantum field theory is done in Euclidean space since only there is the path integral well defined. Essentially all nonperturbative work (numerical simulations, for example) is done in Euclidean space because computers can't handle violent oscillations. The price to be paid is that you have to undo the "Wick rotation" ($x_0 = -ix_4$) at the end of the day. Away from perturbative calculations, that is usually not so straightforward to carry out. Since I am pointed at Feynman rules, I will (mostly) stay in Minkowski space in what follows.

Let's see how path integrals reproduce the propagator. We set $\bar{V} = 0$ and write

$$W_0(J) = N \int [d\phi] \exp(i \int d^4x [\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 + i\epsilon\phi^2 + J\phi]) \tag{7.44}$$

(note the convergence factor!). We can define Fourier transforms ($(1/\sqrt{2\pi})^4 = (2\pi)^2$)

$$\begin{aligned} F(p) &= \int \frac{d^4x}{(2\pi)^2} e^{-ipx} F(x) \\ F(x) &= \int \frac{d^4p}{(2\pi)^2} e^{ipx} F(p) \end{aligned} \tag{7.45}$$

so that the object in the exponential of Eq. 7.44 is

$$\begin{aligned} & i\int d^4x\frac{d^4pd^4p'}{(2\pi)^4}[(-pp'-m^2+i\epsilon)\phi(p)\phi(p')e^{i(p+p')x}+J(p)\phi(p)e^{i(p+p')x}] \\ = \;& i\int d^4p(\phi(-p)[p^2-m^2+i\epsilon]\phi(p)+J(-p)\phi(p)) \end{aligned} \tag{7.46}$$

This is (again) a shifted Gaussian so we can write

$$\phi'(p)=\phi(p)+\frac{1}{p^2-m^2+i\epsilon}J(p) \tag{7.47}$$

and complete the square.

$$W_0(J)=\exp(\frac{i}{2}\int d^4p\frac{J(p)J(-p)}{p^2-m^2+i\epsilon} \tag{7.48}$$

and, once again Fourier transforming,

$$W_0(J)=W_0(0)\exp(-\frac{i}{2}\left\langle J_1\Delta_F(1,2)J_2\right\rangle \tag{7.49}$$

where I've introduced the shorthand

$$\left\langle J_1\Delta_F(1,2)J_2\right\rangle=\int d^4x_1d^4x_2J(x_1)\Delta_F(x_1-x_2)J(x_2) \tag{7.50}$$

and

$$\Delta_F(x)=\int\frac{d^4p}{(2\pi)^4}e^{-ipx}\frac{1}{p^2-m_2+i\epsilon}, \tag{7.51}$$

the usual Feynman propagator, once again.

Let's evaluate some Green's functions:

$$\begin{aligned} G^{(2)}(x,y) &= (-i)^2\frac{\partial}{\partial J_i}\frac{\partial}{\partial J_j}\exp(-\frac{i}{2}\sum_{kl}J_k\Delta_F^{kl}J_l)|_{J=0} \\ &= i\Delta_F(x,y)=\int\frac{d^4p}{(2\pi)^4}e^{-ipx}\frac{i}{p^2-m_2+i\epsilon}. \end{aligned} \tag{7.52}$$

and

$$\begin{aligned} G^{(4)}(x_1,x_2,x_3,x_4) &= \frac{\partial}{\partial J_1}\frac{\partial}{\partial J_2}\frac{\partial}{\partial J_3}\frac{\partial}{\partial J_4}W_0(J)|_{J=0} \\ &= [(i\Delta_F(x_1-x_2))(i\Delta_F(x_3-x_4)+(1-3)(2-4)+(1-4)(2-3). \end{aligned} \tag{7.53}$$

The last expression is Wick's theorem in path integral form.

Now for general formulas – this is very dull, but you need to have some procedure which gives all the Feynman diagrams with their (correct) combinatorial weights, to some order in perturbation theory. Here I am (mostly) following Ramond. We begin with

$$W(J) = N \int [d\phi] \exp(i \int d^4x [\frac{1}{2}(\partial_\mu \phi^2) - \frac{1}{2}(m^2 - i\epsilon)\phi^2 - \bar{V}(\phi) + J(x)\phi(x)]). \tag{7.54}$$

In what follows I will assume that $\bar{V}(\phi)$ is a polynomial in ϕ with a small coefficient, and we want to find a perturbative expansion for $W(J)$. For example, we could consider $\bar{V} = \lambda\phi^4/4!$ and imagine that we want to find $W(J) = W_)(J) + \lambda W_1 + \lambda^2 W_2 + \ldots$. In that case

$$\frac{\partial}{\partial J_1} \ldots \frac{\partial}{\partial J_N} W = G_n^{(0)} + \lambda G_n^{(1)} + \lambda^2 G_n^{(2)} + \ldots. \tag{7.55}$$

To begin this long exercise, write

$$W(J) = N \int [d\phi] \exp(i \int d^4x \bar{V}(\phi)) \exp(iS_0) \exp(i \int d^4x J(x)\phi(x)) \tag{7.56}$$

and I find that it is also convenient to replace all integrals by sums as I go along:

$$\begin{aligned} \exp(i \int d^4x \bar{V}(\phi)) &= \exp(i \sum_k \bar{V}(\phi_k)) \\ \exp(i \int d^4x J(x)\phi(x)) &= \exp(i \sum_k J_k \phi_k). \end{aligned} \tag{7.57}$$

Now the procedure is to de-expand the exponential in $\exp(i \int \bar{V})$. Each term is a polynomial in ϕ. Schematically, a term is

$$W(J) = N \int [d\phi] e^{iS_0} \phi_l^M e^{i \sum J_k \phi_k} = \left\langle \phi_l^M \right\rangle. \tag{7.58}$$

But

$$\frac{1}{i} \frac{\partial}{\partial J_l} \exp(i \sum_k J_k \phi_k)_{J=0} = \phi_l. \tag{7.59}$$

So we can replace the argument ϕ_k in $\bar{V}(\phi_k)$ with $\frac{1}{i}\frac{\partial}{\partial J_k}$,

$$\exp(-i \int d^4x (J(x)\phi(x) + \bar{V}(\phi(x)) = \exp(-i \int d^4x (J(x)\phi(x) + \bar{V}(\frac{\partial}{i\partial J(x)}))) \tag{7.60}$$

or

$$\exp(-i\sum_k (V(\phi_k) + J_k\phi_k) = \exp(-i\sum_k (V(\frac{\partial}{i\partial J_k}) + J_k\phi_k))..\tag{7.61}$$

This seems like an absurd thing to do, but now notice that

$$W(J) = N\int [d\phi]\{\exp(-i\int d^4x V(\frac{\partial}{i\partial J(x)})\} \times \exp(iS_0 - i\int d^4x J(x)\phi(x))..\tag{7.62}$$

The expression in curly brackets does not depend on ϕ – we can pull it outside the functional integral. And what is left is just $W_0(J)$. Thus we have found that

$$W(J) = \exp(-i\sum_k \bar{V}(\frac{\partial}{i\partial J_k}))W_0(J)\tag{7.63}$$

and again I have written the shorthand expression

$$W_0(J) = \exp(-\frac{i}{2}J_m\Delta_F^{ml}J_l) = \exp(-\frac{i}{2}\int d^4x d^4y J(x)\Delta_F(x-y)J(y)).\tag{7.64}$$

To find Green's functions in the interacting system, just de-expand the curly-bracket exponential and take $J-$derivatives, term by term.

As an example, suppose $\bar{V}(\phi) = \lambda\phi^4/4!$.

$$W(J) = [1 - i\sum_k \frac{\lambda}{4!}(\frac{1}{i})^4\frac{\partial^4}{\partial J_k^4} - \frac{1}{2}\sum_{kl}(\frac{\lambda}{4!})^2(\frac{1}{i})^8\frac{\partial^4}{\partial J_k^4}\frac{\partial^4}{\partial J_l^4} + \ldots]\exp(-\frac{i}{2}J_m\Delta_F^{mn}J_n)\tag{7.65}$$

Of course,

$$G_{12} = \frac{\partial^2 W(J)}{\partial J_1\partial J_2}|_{J=0}; \qquad G_{1234} = \frac{\partial^4 W(J)}{(\partial J_1\partial J_2\partial J_3\partial J_4}|_{J=0}\tag{7.66}$$

and so on.

Differentiating the "1" term gives the lowest order $G_{12} = \Delta_{12}$. Let's look at the $O(\lambda)$ term, which comes from

$$\frac{-i\lambda}{(-i)^4 4!}\frac{\partial^4}{\partial J_k^4}\exp(-\frac{i}{2}J_m\Delta_F^{mn}J_n).\tag{7.67}$$

Several lines of uninteresting algebra give an intermediate answer; the expression (before taking $J = 0$) is

$$-i\frac{\lambda}{4}[-3\Delta_{kk}^2 - 6i\Delta_{kp}\Delta_{kk}\Delta_{kq}J_pJ_q + \Delta_{kp}\Delta_{kq}\Delta_{kr}\Delta_{ks}J_pJ_qJ_rJ_s]\exp(-\frac{i}{2}J_m\Delta_F^{mn}J_n).\tag{7.68}$$

Now we can differentiate with respect to the J's and then set them to zero. The order λ correction to $G^{(2)}$ consists of two terms:

- (a) $-i\frac{\lambda}{4!}(-6i)\Delta_{k1}\Delta_{k2}\Delta_{kk}\times 2$
- (b) $-3\frac{\lambda}{4!}\Delta_{kk}^2\Delta_{12}$

The interesting order λ contribution to $G^{(4)}$ is $-i(\frac{\lambda}{4!})4!\Delta_{ka}\Delta_{k2}\Delta_{k3}\Delta_{k4}$. The corresponding graphs are shown in Fig. 7.2.

A couple of things are to be noticed: First of all, some of the diagrams are "connected" and some are "disconnected." When you evaluate the connected diagrams, they will contain delta functions which insure overall momentum conservation, and the delta functions do not factorize. These graphs correspond to real scattering amplitudes ($1+2\to 3+4$, for example, with a $\delta^4(p_1+p_2-p_3-p_4)$). The disconnected diagrams do not correspond to a full scattering process. Physically, they correspond to a product of individual smaller scattering processes. Momentum is separately conserved in each subprocess. For example, graph (b) only has a factor $\delta^4(p_1-p_2)$.

Second, the formalism I have described generates what are called "unamputated graphs" meaning that the amplitude includes propagators for the incoming and outgoing states. T-matrix elements want something different, the amplitude without propagators for the external particles. For example, the $O(\lambda)$ scattering amplitude for $1+2\to 3+4$ is just $-i\lambda$. Passing from unamputated graphs to amputated ones can be done at increasingly higher levels of sophistication. A good place to read about this is in the text by Srednicki; look up "LSZ formalism" (for Lehmann, Symanzik, Zimmerman). Here I am just clipping off the incoming and outgoing propagators "by eye."

A good place to read about generating functionals for connected graphs is in Ramond's book. Briefly, one writes

$$W(J)=\exp(iZ(J)) \tag{7.69}$$

and discovers that the connected $n-$point function is

$$G^{(n)}_{conn}=\frac{\partial^n Z}{\partial J_1\partial J_2\ldots}|_{J=0} \tag{7.70}$$

This is easy to check for free field theory:

$$Z_0(J)=\frac{i}{2}J_k\Delta_{kl}J_l \tag{7.71}$$

and immediately we see

$$G^{(2)}_{conn}=\frac{i}{2}\times 2\Delta_{ij};\qquad G^{(n>2)}_{conn}=0. \tag{7.72}$$

He has useful tricks for carrying out this program for interacting systems.

I admit, the last few pages are pretty dense. But if you are going to do a perturbative calculation, you have to be sure that you have written down all the Feynman diagrams you need, and you have to get the combinatorial weight of each graph correctly. For simple graphs, that can be done by inspection, but intuition may not be good enough in all cases.

And let's sum up: we are pretty much done with formalism for Feynman diagrams for a while. You have seen how perturbation theory is developed in the language of Hamiltonians and in the language of path integrals. Both languages have their good points (and not-so-good points). Hamiltonians are closest to what you have already seen in your quantum mechanics classes and may be more familiar. However, there are many problems for which Hamiltonian methods are much more complicated than path integral methods. An example is the quantization of non-Abelian gauge theories. Most of my career has been spent working with path integrals, studying QCD and related systems. Most of the calculations I have done would basically be impossible using Hamiltonians. (And there are interesting problems which are basically impossible with path integrals, and Hamiltonians are starting to come back into my research field.) Of course, a good physicist speaks both languages!

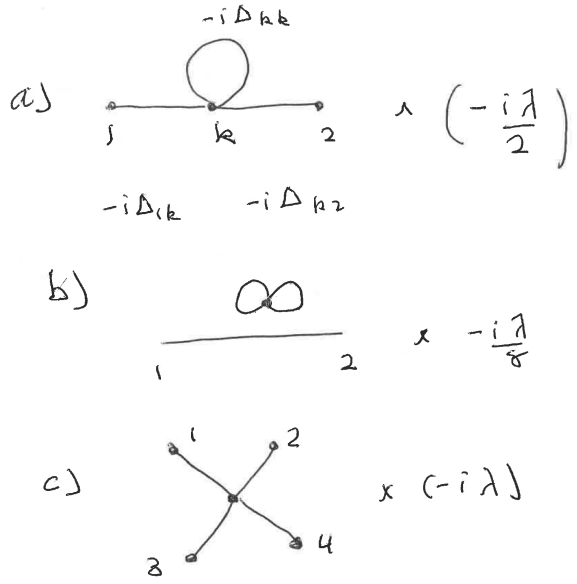

Figure 7.2: Lowest order scattering Feynman diagram for $\mathcal{H} = g\phi^4/4!$. (a) and (b) are the contributions to $G^{(2)}$ and (c) is the contribution to $G^{(4)}$.

Chapter 8

Spin-$\frac{1}{2}$: foundations, generalities, practicalities

8.1 The Dirac equation

Recall that the wave functions of nonrelativistic electrons have two components corresponding to the two possible eigenvalues of one direction of their angular momentum. They are spinors; under rotations they transform as spin-$\frac{1}{2}$ objects.

We now face the issue of embedding spin into special relativity. In fact, it is more interesting to look the other way around: to start with special relativity, or more precisely with the invariance of the squared length of four vectors, and to see spin-$\frac{1}{2}$ appear as part of a general discussion of the properties of states under Lorentz transformations. This is very much like constructing states which are eigenstates of J^2 and J_z by considering irreducible representations of the rotation group, rather than writing down solutions to partial differential equations and studying their properties.

Along the way, we will encounter the Dirac equation. The approach I am taking has an advantage that we can see the extent to which the Dirac equation is "unique" as a description of spin-$\frac{1}{2}$ particles – as well as other ways that spinors appear in the context of special relativity.

And we cannot stop there. Relativistic quantum mechanics is internally inconsistent, and so the Dirac equation, as well as other formulations of spinors in special relativity, are just ingredients in relativistic quantum field theories. We have to complete that connection, too.

8.2 The Lorentz group and relativistic quantum field theories

Reference: Ramond, pp 6-32.

We all know that the squared length of a four-vector

$$s^2 = x^\mu x^\nu g_{\mu\nu} = t^2 - r^2 \tag{8.1}$$

is unchanged under Lorentz transformations. We want to find the most general set of linear transformations

$$x'^\mu = \Lambda^\mu_\nu x^\nu \tag{8.2}$$

which preserves s^2. We require that

$$g_{\mu\nu}x'^{\mu}x'^{\nu} = g_{\mu\nu}\Lambda^{\mu}_{\rho}\Lambda^{\nu}_{\sigma}x^{\rho}x^{\sigma} = g_{\rho\sigma}x^{\rho}x^{\sigma} \tag{8.3}$$

or

$$g_{\mu\nu}\Lambda^{\mu}_{\rho}\Lambda^{\nu}_{\sigma} = g_{\rho\sigma}. \tag{8.4}$$

We can classify the transformations as proper or improper depending on whether Λ has an infinitesimal limit. To do this, consider x as a four-component column vector and L the matrix generalization of Λ,

$$x' = Lx. \tag{8.5}$$

Space inversions such as

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{8.6}$$

are improper. Boosts, such as

$$L = \begin{pmatrix} \gamma & 0 & 0 & \beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \beta\gamma & 0 & 0 & \gamma \end{pmatrix}, \tag{8.7}$$

are proper. We will be concerned almost exclusively with proper transformations in the next few pages – boosts and rotations.

Consider an infinitesimal Lorentz transformation

$$\Lambda^{\mu}_{\nu} = \delta^{\mu}_{\nu} + \epsilon^{\mu}_{\nu}. \tag{8.8}$$

δ^{μ}_{ν} is a Kroneker delta, $\delta^{\mu}_{\mu} = 1$ for all μ. The parameter ϵ characterizes the Lorentz transformation. It is constrained by Eq. 8.4,

$$\begin{aligned} g_{\rho\sigma} &= g_{\mu\nu}\Lambda^{\mu}_{\rho}\Lambda^{\nu}_{\sigma} \\ &= g_{\mu\nu}[\delta^{\mu}_{\rho} + \epsilon^{\mu}_{\rho}][\delta^{\nu}_{\sigma} + \epsilon^{\nu}_{\sigma}] \\ &= g_{\rho\sigma} + g_{\mu\nu}(\epsilon^{\mu}_{\rho}\delta^{\nu}_{\sigma} + \delta^{\mu}_{\rho}\epsilon^{\nu}_{\sigma}) \\ &= g_{\rho\sigma} + \epsilon_{\rho_\sigma} + \epsilon_{\sigma\rho}. \end{aligned} \tag{8.9}$$

This means that $\epsilon_{\sigma\rho} + \epsilon_{\rho_\sigma} = 0$ or that $\epsilon_{\mu\nu}$ is an antisymmetric tensor, This in turn means that it has six nonzero components. These will correspond to the magnitudes of the boosts and the rotations in each of the three dimensions.

We can introduce a sort of "generalized angular momentum" as a generator of Lorentz transformations:

$$L_{\mu\nu} \equiv i[x_\mu \partial_\nu - x_\nu \partial_\mu] \tag{8.10}$$

(recall $\partial_\mu = (\partial/\partial t, \vec{\nabla})$) in terms of which the change in x can be written as

$$\delta x^\mu = \frac{1}{2}\epsilon^{\rho\sigma} L_{\rho\sigma} x^\mu \tag{8.11}$$

(this is $\delta x^\mu = \epsilon^\mu_\nu x^\nu$ after taking the derivatives, of course.) Again, because of antisymmetry, $L_{\mu\nu}$ has six nonzero components; the number of generators is equal to the number of ϵ's. It's easy but tedious to work out the commutator,

$$[L_{\mu\nu}, L_{\rho\sigma} = ig_{\nu\rho}L_{\mu\sigma} - ig_{\mu\rho}L_{\nu\sigma} - ig_{\nu\sigma}L_{\mu\rho} + ig_{\mu\sigma}L_{\nu\rho}. \tag{8.12}$$

This says that the L's are the generators of a Lie algebra, This should not be surprising, since for the space-like components, we can define

$$L_i \equiv \frac{1}{2}\epsilon_{ijk}L_{jk} \tag{8.13}$$

and Eq. 8.12 says

$$[L_1, L_2] = [L_{23}, L_{31}] = ig_{33}L_{21} = -iL_{21} = iL_{12} = iL_3 \tag{8.14}$$

($\mu = 2$, $\nu = 3$, $\rho = 3$, $\sigma = 1$) which is one of the usual commutation relations for the ordinary angular momentum operators.

In an analogy with the relation between spin and ordinary orbital angular momentum, we can imagine more generators $S_{\mu\nu}$ which act on the internal structure of our states, with the same commutation relations among themselves, and with $[L_{\mu\nu}, S_{\rho\sigma}] = 0$, We can write the most general representation of the generators of Lorentz transformations as

$$M_{\mu\nu} = L_{\mu\nu} + S_{\mu\nu}. \tag{8.15}$$

Let's pull the commutator Eq. 8.12 apart so we can make sense of it. Define

$$\begin{aligned} J_i &= \frac{1}{2}\epsilon_{ijk}M_{jk} \\ K_i &= M_{0i} \end{aligned} \tag{8.16}$$

so that the six nonzero M's are recast as three J's and three K's, and then take linear combinations

$$\begin{aligned} A_i &= \frac{1}{2}(J_i + iK_i) \\ B_i &= \frac{1}{2}(J_i - iK_i). \end{aligned} \tag{8.17}$$

Then we find

$$[A_i, B_j] = 0 \tag{8.18}$$

$$[A_i, A_j] = i\epsilon_{ijk}A_k \tag{8.19}$$

$$[B_i, B_j] = i\epsilon_{ijk}B_k. \tag{8.20}$$

This says that the A and B's obey the algebra of $SU(2)$!

In the case of ordinary angular momentum, states which transform irreducably under rotations may be characterized by a single set of angular momentum quantum numbers, J^2 and J_z, and for example $J^2 |\psi_j\rangle = j(j+1) |\psi_j\rangle$. States transforming as irreducable representations under the Lorentz group may be classified by two angular momentum quantum numbers, a and b:

$$\begin{aligned} A^2 |\psi_{ab}\rangle &= a(a+1) |\psi_{ab}\rangle \\ B^2 |\psi_{ab}\rangle &= b(b+1) |\psi_{ab}\rangle \end{aligned} \tag{8.21}$$

Note also that the two $SU(2)$'s are not independent. Under a parity transformation $J_i \to J_i$ (or $M_{jk} \to M_{jk}$) – the J's form an axial vector, while $K_i \to -K_i$, the K's form a vector. This means that a parity reflection is equivalent to exchanging A with B and vice versa. In general, representations of the Lorentz group are neither parity nor Hermitian conjugation eigenstates ($B_i = A_i^\dagger$).

And since $J_i = A_i + B_i$ we can identify the ordinary spin of the representation as $J = a+b$.

Rotation matrices transform states: Working in momentum space for a while, and imagining that we have some state which depends on p, the transformation law is

$$\psi(\Lambda p) = D(\Lambda)\psi(p). \tag{8.22}$$

The usual quantum mechanical rotation matrix is

$$\begin{aligned} D(\phi) &= \exp(-i\vec{\phi}\cdot\vec{J}) \\ &\simeq 1 - i\vec{\phi}\cdot\vec{J} \\ &= 1 - i\phi_i(\frac{1}{2}\epsilon_{ijk}M_{jk}) \\ &= 1 - \frac{i}{2}(\epsilon_{ijk}\phi_i)M_{jk} \end{aligned} \tag{8.23}$$

regrouping indices, to make the formula involve two-index generators. If ϕ is a rotation about the $z-$axis,

$$\epsilon_{ijk}\phi_i = \begin{pmatrix} 0 & \phi & 0 \\ -\phi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{8.24}$$

and we see that $\epsilon_{ijk}\phi_i = \epsilon_{jk} = -\epsilon_{kj}$ is the parameter characterizing the rotation. The natural generalization of a rotation to the Lorentz group is then

$$D(\Lambda) = 1 - \frac{i}{2}\epsilon^{\mu\nu}M_{\mu\nu}. \tag{8.25}$$

Let's work this out for a boost along the z direction,

$$B(\theta) = \begin{pmatrix} \gamma & 0 & 0 & \beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \beta\gamma & 0 & 0 & \gamma \end{pmatrix}, \tag{8.26}$$

where we introduce the rapidity angle θ to write $\gamma = \cosh\theta$, $\beta\gamma = \sinh\theta$. We can write this as a product if infinitesimal boosts $[B(\theta/n)]^n$ where

$$B\left(\frac{\theta}{n}\right) = \begin{pmatrix} 1 & 0 & 0 & \frac{\theta}{n} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{\theta}{n} & 0 & 0 & 1 \end{pmatrix}. \tag{8.27}$$

The infinitesimal rotation is

$$D\left(\frac{\theta}{n}\right) = 1 - \frac{i}{2}[\omega^{30}M_{30} + \omega^{03}M_{03}] \tag{8.28}$$

and $\omega^{30} = \theta/n$, $M_{30} = -K_3$, $K_i = A_i - B_i$, so the rotation matrix is

$$D\left(\frac{\theta}{n}\right) = 1 + \frac{1}{2}\frac{\theta}{n} 2(A_3 - B_3) \tag{8.29}$$

combining $\mu\nu$ and $\nu\mu$ and the i's. This exponentiates to

$$\begin{aligned} D(B(\theta)) &= \lim_{n\to\infty}[1 + \frac{\theta}{n}(A_3 - B_3)]^n \\ &= \exp(\theta(A_3 - B_3)) \\ &\to \exp[\theta(\vec{A} - \vec{B}) \cdot \hat{n}] \end{aligned} \tag{8.30}$$

generalizing to a pure boost along direction $\hat{n}$.

What does this mean? (A, B) representations of the Lorentz group can be written as

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} \equiv \psi_{ab} \tag{8.31}$$

where, in turn, if $A = J_A$ then ψ_A is a column vector with $2A + 1$ entries, and similarly for ψ_B. The transformation law is

$$\psi_{a'b'}(p', s') = [\exp(\theta\hat{n} \cdot \vec{A})]_{a'a}[\exp -\theta\hat{n} \cdot \vec{B})]_{b'b}\psi_{ab}(p.s). \tag{8.32}$$

For example, if the states are $A = B = \frac{1}{2}$, then the operators $A = B = \frac{1}{2}\sigma$ and

$$D(\Lambda) = \begin{pmatrix} \exp(\frac{1}{2}\theta\hat{n} \cdot \vec{\sigma}) & 0 \\ 0 & \exp(-\frac{1}{2}\theta\hat{n} \cdot \vec{\sigma}) \end{pmatrix}. \tag{8.33}$$

Now we can classify states in terms of their transformation properties under boosts and rotations.

The simplest states are scalars, with $A = B = 0$. They have a single component. The Lagrange density should be a scalar under Lorentz transformations and should be built of scalar functions of the field variables.These could be polynomials in ϕ. And we can combine operators which transform nontrivially under Lorentz transformations into scalars. This is the analog of combining objects with nonzero angular momenta into (rotational) scalars. For example, ∂_μ transforms as a four-vector operator; $\partial_\mu\phi$ is a four vector (if ϕ is a scalar), so its square

$$\mathcal{L}_K = (\partial_\mu\phi)(\partial^\mu\phi) \tag{8.34}$$

is a scalar. It could be part of a Lagrange density for scalar fields. Since I've already written quite a bit about quantum field theory for scalars, there isn't much more to say.

8.3 Technology for spinors

Maybe it is time to talk about spinors. The easiest ones to describe have two components. Possible states can be the $(A, B) = (\frac{1}{2}, 0)$ or $(0, \frac{1}{2})$ combinations. These objects describe a left handed or a right handed fermion, respectively. These states are *not* parity eigenstates, since under parity, $P(A, B) = (B, A)$. To get a parity eigenstate we need to combine two of these objects, $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ into a four component "Dirac spinor." The transformation rules for the two component spinors are

$$\begin{aligned} (\frac{1}{2}, 0) \to \psi_L(x); \qquad & \psi'_L(x') = \Lambda_L \psi_L(x) \\ (0, \frac{1}{2}) \to \psi_R(x); \qquad & \psi'_R(x') = \Lambda_R \psi_R(x) \end{aligned} \tag{8.35}$$

where

$$\Lambda_{L,R} = \exp(\frac{i}{2}\vec{\sigma} \cdot (\vec{\omega} \pm i\vec{\nu})) \tag{8.36}$$

for a rotation about direction $\vec{\omega}$ and a boost along direction $\vec{\nu}$.

Let's record some useful properties of the rotation matrices. Note that $\Lambda_L \sim 1 + \frac{i}{2}\vec{\sigma} \cdot \vec{\omega} + \frac{1}{2}\vec{\sigma} \cdot \vec{\nu}$ and $\Lambda_L^{-1} = 1 - \frac{i}{2}\vec{\sigma} \cdot \vec{\omega} - \frac{1}{2}\vec{\sigma} \cdot \vec{\nu} = \Lambda_R^\dagger$.

Property 1: $\Lambda_L^{-1} = \Lambda_R^\dagger$.

Since $\sigma_z \sigma_i \sigma_z = -\sigma_i^*$, we have Property 2: $\sigma_2 \Lambda_L \sigma_2 = \Lambda_R^*$ and $\sigma_2 \Lambda_L^* \sigma_2 = \Lambda_R$.

More Pauli matrix identities! $(\sigma_2 \sigma_i \sigma_2)^T = -(\sigma_i^*)^T$, so with $\Lambda_L^T = 1 - \frac{i}{2}\vec{\sigma}^T \cdot \vec{\omega} + \frac{1}{2}\vec{\sigma}^T \cdot \vec{\nu}$, we have Property 3 is $\sigma_2 \Lambda_L^{-1} \sigma_2 = \Lambda_L^T$ or $\sigma_2 \Lambda_L^T \sigma_2 \Lambda_L = 1$.

Finally, property 4: $\Lambda_L^T \sigma_2 \Lambda_L = \sigma_2$ and $\Lambda_R^T \sigma_2 \Lambda_R = \sigma_2$.

These odd identities have their uses. First, $\sigma_2 \Lambda_L^* \psi_L^* = \sigma_2 \Lambda_L^* \sigma_2 \sigma_2 \psi_L^* = \Lambda_R(\sigma_2 \psi_L^*)$. This says that, given a ψ which transforms as $\psi'_L = \Lambda_L \psi_L$, (alternatively , it is an $(\frac{1}{2}, 0)$ object) call it $\begin{pmatrix} a \\ b \end{pmatrix}$, one can construct a $\sigma_2 \psi_L^* = \begin{pmatrix} -ib^* \\ ia^* \end{pmatrix}$ which transforms as $(0, \frac{1}{2})$. This expands our set of states: ψ_L and $\sigma_2 \psi_R^*$ are $(\frac{1}{2}, 0)$, ψ_R and $\sigma_2 \psi_L^*$ are $(0, \frac{1}{2})$.

This relation has a potential use. If we have a set of left handed and right handed spinors, we can always trade in the right hand ones for left hand ones, and vice versa. In words, the physics is that the antiparticle of a left handed fermion is a right handed antifermion, plus the additional statement that the whole notion of particle versus antiparticle in some

absolute sense is ambiguous. In the Standard Model, left handed fermions and right handed fermions have different weak charges. W bosons don't couple to right handed fermions, only to left handed ones. If we want, (and there are people who work this way), we never have to talk about right handed fermions. We can work with left handed anti fermions, instead. In this way of doing things, the electron (a four component Dirac particle) is built of two left handed fields, an $e(x)$ and an $\bar{e}(x)$.

Another use of these relations is to construct combinations of spinors which transform as scalars and vectors.

Suppose we have two left handed spinors, χ_L and ψ_L. The quantity $S = \chi_L^T \sigma_2 \psi_L$ transforms into $\chi_L^T \Lambda_L^T \sigma_2 \Lambda_L \psi_L$ which (from property 4) is $S = \chi_L^T \sigma_2 \psi_L$ – that is, S is unchanged under a Lorentz transformation. This means that S is a scalar! To write it in another way, let $\chi_L = \sigma_2 \psi_R^*$ and then $S = -\psi_R^\dagger \psi_L$. This is not necessary real, but $\psi_R^\dagger \psi_L + \psi_L^\dagger \psi_r$ is always real. Moral: the product of a left handed fermion and a right handed fermion is a scalar.

Now we want vectors. $\psi_L^\dagger \psi_L$ is invariant under rotations (the transformation is unitary) but not under boosts. Under a boost

$$\psi_L^\dagger \psi_L \to \psi_l^\dagger \Lambda_L^\dagger \Lambda_L \psi_L = \psi_L^\dagger \exp(\vec{\sigma} \cdot \vec{\nu}) \psi_L. \tag{8.37}$$

If we Taylor expand this expression, it is $\psi_L^\dagger \psi_L + \vec{\nu} \cdot \psi_L^\dagger \vec{\sigma} \psi_L + \ldots$. Similarly $\psi_L^\dagger \sigma_i \psi_L \to \psi_l^\dagger e^{\sigma \cdot \nu/2} \sigma_i e^{\sigma \cdot \nu/2} \psi_l$ which is $\psi_L \sigma_i \psi_l + \nu_i \psi_L^\dagger \psi_L + \ldots$. That is, the changes are

$$\begin{aligned} \delta \psi_L^\dagger \psi_L &= \nu_i \psi_L^\dagger \sigma_i \psi_L \\ \delta \psi_L^\dagger \sigma_i \psi_L &= \nu_i \psi_L^\dagger \psi_L. \end{aligned} \tag{8.38}$$

This is $\delta V^\mu = \epsilon^\mu_\nu V^\nu$ with $\epsilon^{0i} = -\nu_i$ as the boost parameters. Thus the following objects are four vectors

$$\begin{aligned} \psi_L^\dagger \sigma^\mu \psi_L &= (\psi_L^\dagger \psi_L, \psi_L^\dagger \vec{\sigma} \psi_L) \\ \psi_R^\dagger \bar{\sigma}^\mu \psi_R &= (\psi_R^\dagger \psi_R, -\psi_R^\dagger \vec{\sigma} \psi_R) \end{aligned} \tag{8.39}$$

(And note for future reference the definitions of $\sigma^\mu = (1, \vec{\sigma})$ and $\bar{\sigma}^\mu = (1, -\vec{\sigma})$.)

Now we are in business. Suppose we want to write down a Lagrangian for left handed fermions. There are two terms which are quadratic in the fields. One of them is a kinetic term

$$\mathcal{L}_K = \psi_L^\dagger \sigma^\mu \partial_\mu \psi_L \tag{8.40}$$

(the right hand fermion analog is $\psi_R^\dagger \bar{\sigma}^\mu \partial_\mu \psi_R$). The other term is a Majorana mass term – maybe!

$$\mathcal{L}_{MM} = \frac{m}{2}(\psi_L^\dagger \sigma_2 \psi_L^* + \psi_L^T \sigma_2 \psi_L) \tag{8.41}$$

Why "maybe?" If ψ carries a charge Q, then $\psi^\dagger$ must have charge $-Q$ in order that $\mathcal{L}_K$ is charge neutral. For Majorana fermions, the particle *is* the antiparticle (we'll see that, below) so we had better have $Q = 0$ because ψ^T carries the same charge as ψ and $\mathcal{L}_{MM}$ had better also be neutral. More on this, later, but I can't resist remarking that neutrinos could be Majorana fermions, but electrons can't.

Finally we come to four component (Dirac) spinors, which are what we need when our interactions conserve parity - electrodynamics by itself, quantum chromodynamics. Recall that parity takes L into R, $P\psi_L = \psi_R$ and vice versa, so if we want an action which respects parity, we have to have both ψ_L and ψ_R bundled together into four component spinor. We can write

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \tag{8.42}$$

Notice

$$P\psi = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \tag{8.43}$$

To simplify notation, we define (careful, each entry is 2×2)

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{8.44}$$

and say $P\psi = \gamma^0 \psi$. Left and right hand projectors are $\frac{1}{2}(1 \pm \gamma_5)$ where

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.45}$$

Recall that that a scalar is $\psi_R^\dagger \psi_L + \psi_L^\dagger \psi_R = \psi^\dagger \gamma^0 \psi \equiv \bar{\psi}\psi$. (Note the definition $\bar{\psi} = \psi^\dagger \gamma^0$.) Vector operators come from the two component vectors:

$$V^\mu = \psi^\dagger(\gamma^0 \gamma^\mu)\psi \equiv \bar{\psi}\gamma^\mu \psi \tag{8.46}$$

where

$$\gamma^i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix}. \tag{8.47}$$

The four by four objects $\gamma^\mu = (\gamma^0, \gamma_i)$ are referred to as the "Dirac matrices" (with further adjectives, see below.) The Dirac matrices obey the anticommutation algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{8.48}$$

which is a property which descends from the Pauli matrices.

Recall that we could also bundle left handed and right handed fields as

$$\psi^c = \begin{pmatrix} \sigma_2\psi_R^* \\ -\sigma_2\psi_L^* \end{pmatrix} \propto \gamma^2\psi^*. \tag{8.49}$$

This is called the "charge conjugate spinor." Under Lorentz transformations it transforms exactly like ψ. Do it twice: $(\psi^c)^c = \psi$. The transformation itself is called "charge conjugation." We will see what it is used for, shortly.

The Majorana spinor

$$\psi^M = \begin{pmatrix} \psi_L \\ -\sigma_2\psi_L \end{pmatrix} \tag{8.50}$$

is equal to its conjugate, $\psi^M = (\psi^M)^c$ – check it! We will shortly show that the charge conjugation operator turns a fermion into its antiparticle, so we have just discovered that Majorana fermions are their own antiparticles. Also, like ψ_R and ψ_L, the "Weyl spinors," Majorana fermions only have two independent components. The Majorana mass term can also be written in terms of the Majorna spinor as

$$\mathcal{L}_{mm} = \frac{m}{2}\bar{\psi}^M\psi^M. \tag{8.51}$$

"The" Dirac Lagrangian is built from $\partial_\mu V^\mu$ and the scalar term,

$$\mathcal{L} = \bar{\psi}(x)[i\gamma^\mu\partial_\mu - m]\psi(x). \tag{8.52}$$

Incidentally, when $\mathcal{L}$ is embedded in an action, there are many choices for derivative operators, all related by integration by parts. The derivative could act only on the ψ, or only on the $\bar{\psi}$ or half on each. In Eq. 8.52 just varying $\delta\mathcal{L}/\delta\bar{\psi} = 0$ gives "the" Dirac equation

$$[i\gamma^\mu\partial_\mu - m]\psi(x) = 0. \tag{8.53}$$

For future use, there is an alternative "slash notation:" $\not{p} = \gamma_\mu p^\mu$ and so on.

8.4 Technical details about solutions of the Dirac equation

Let's develop some technology for dealing with solutions of the Dirac equation. We will need the Hamiltonian, which in turn will need a canonical field momentum; It is

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\bar{\psi}\gamma^0 = i\psi^\dagger. \tag{8.54}$$

The Hamiltonian density is

$$\mathcal{H} = \psi^\dagger[-i\vec{\alpha}\cdot\vec{\nabla} + \beta m]\psi \tag{8.55}$$

where (this is the first of several annoying changes in definition) $\gamma^0 = \beta$, $\vec{\alpha} = \beta\vec{\gamma}$. To continue being annoying, there are several competing version of the Dirac matrices. We began with what is called Weyl basis,

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \qquad \gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ -\sigma^\mu & 0 \end{pmatrix}. \tag{8.56}$$

This is a natural basis for talking about left handed and right handed fermions. In Bjorken-Drell convention, γ^0 is diagonal

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \qquad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}; \qquad \gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{8.57}$$

The anticommutation relation Eq. 8.48 is basis-independent, of course.

(Why are there competing versions? because the Dirac equation simplifies in various limits, $m/E \to 0$ or $m/E \to 1$, and the solutions are simpler in these limits in different bases.)

Let's look at solutions of the free Dirac equation. We assume

$$\psi(x) = \psi(p)\exp(-i(Et - \vec{p}\cdot\vec{x})) = \psi(p)e^{-ipx} \tag{8.58}$$

where $\psi(p)$ is a four component function of p. We work in Weyl basis, where $\psi(p) = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$. The Dirac equation becomes

$$m\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} 0 & E + p\cdot\sigma \\ E - p\cdot\sigma & 0 \end{pmatrix}\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{8.59}$$

We could, of course, keep going, but let's set $m = 0$. Then the equation factorizes into a pair of two-component equations

$$\begin{aligned}(E + \vec{p}\cdot\vec{\sigma})\psi_L &= 0 \\ (E - \vec{p}\cdot\vec{\sigma})\psi_R &= 0\end{aligned} \tag{8.60}$$

Also, $(E - \vec{p}\cdot\vec{\sigma})(E + \vec{p}\cdot\vec{\sigma})\psi_L = (E^2 - p^2)\psi_L = 0$ tells us that we have a massless particle, $E = |\vec{p}|$. Calling $\hat{n} = \vec{p}/E$ we have

$$\hat{n}\cdot\sigma\psi_L = -\psi_L \tag{8.61}$$

which tells us that the axis of quantization of the spin lies along the direction of motion of the particle, antiparallel to it for ψ_L and parallel for ψ_R. Hence the L,R label we had earlier introduced – it is physics. The solutions of the Dirac equation are massless two-component states of definite helicity.

Next, let's consider the nonzero mass case and work in Bjorken-Drell basis for a change of pace. Begin with $p = 0$:

$$i\frac{\partial\psi}{\partial t} = \beta m\psi = \begin{pmatrix} m & & & \\ & m & & \\ & & -m & \\ & & & -m \end{pmatrix}\psi \tag{8.62}$$

There are obviously four solutions: two of them are

$$\psi_1 = C\begin{pmatrix}1\\0\\0\\0\end{pmatrix}e^{-imt}; \qquad \psi_2 = C\begin{pmatrix}0\\1\\0\\0\end{pmatrix}e^{-imt}; \tag{8.63}$$

Recall that is $i\partial\psi/\partial t = E\psi, \psi \sim \exp(-iEt)$, so these are positive energy solutions, $E = +mc^2$, restoring the c's for psychological purposes Looking at the two upper components, they apparently have different internal states, which we can associate with spin up and spin down.

But there are two negative energy solutions, with $E = -mc^2$,

$$\psi_3 = C\begin{pmatrix}0\\0\\1\\0\end{pmatrix}e^{imt}; \qquad \psi_4 = C\begin{pmatrix}0\\0\\0\\1\end{pmatrix}e^{imt}; \tag{8.64}$$

Who ordered these? Let's do some more technical manipulations, then deal with the interpretation of these states. You can wait. After all, it took three years from Dirac's invention of the original Dirac equation to the discovery of the positron.

What about arbitrary momentum? We can find these solutions by boosting; recall

$$\psi(p'/m, s') = D(\Lambda)\psi(p/m, s). \tag{8.65}$$

The only annoyance is that we worked out D in Weyl basis, while now we are in Bjorken - Drell basis. But no worries, $R\gamma^\mu_{AB}R^{-1} = \gamma^\mu_{BD}$, where

$$R = R^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{8.66}$$

and "1" is the 2×2 identity matrix. Then. we just have a little matrix multiplication

$$\psi(p/m, s) = R\begin{pmatrix} e^{\frac{1}{2}\theta\hat{p}\cdot\sigma} & 0 \\ 0 & e^{-\frac{1}{2}\theta\hat{p}\cdot\sigma} \end{pmatrix} R^{-1}\psi(0, s_0) \tag{8.67}$$

and after some algebra, we find that the rotation matrix in Bjorken - Drell basis is

$$D(\Lambda) = \begin{pmatrix} \cosh\frac{\theta}{2} & \hat{p}\cdot\sigma\sinh\frac{\theta}{2} \\ \hat{p}\cdot\sigma\sinh\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \tag{8.68}$$

where there are a plethora of definitions built on $\cosh\theta = E/m$ and $\sinh\theta = p/m$: in particular

$$\begin{aligned} \cosh\frac{\theta}{2} &= \sqrt{\frac{E+m}{2m}} \\ \sinh\frac{\theta}{2} &= = \frac{p}{E+m}\cosh\frac{\theta}{2} \end{aligned} \tag{8.69}$$

so that

$$\psi(p/m, s) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} 1 & \frac{p\cdot\sigma}{E+m} \\ \frac{p\cdot\sigma}{E+m} & 1 \end{pmatrix}\psi(0, s) \tag{8.70}$$

The four $\psi(0, s)$'s of Eqs. 8.63 and 8.64 are so simple that the boosted wave functions can be read off from the individual columns of Eq. 8.70. For example, starting with $\psi_1(p = 0)$, we have

$$\psi_1(x) = C\sqrt{\frac{E+m}{2m}}\begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x - ip_y}{E+m} \end{pmatrix}.\exp(-i(Et - \vec{p}\cdot\vec{x}) \tag{8.71}$$

Let's pull off the plane wave and write $\psi_i(x) = w_i(p)\exp(-i\epsilon_i(p_\mu x^\mu)$ where $\epsilon_1 = \epsilon_2 = 1$ and $\epsilon_3 = \epsilon_4 = -1$; w is the four-component Dirac spinor. To keep going, the w's are actually not the most convenient states. Let's define the two positive energy solutions as $u(p,s) = w(p,s)$. They obey

$$(\not{p} - m)u(p,s) = 0 \tag{8.72}$$

The adjoint of this equation is

$$((\not{p} - m)u(p,s))^\dagger = u^\dagger(\gamma_\mu^\dagger p^{\mu\dagger} - m) = 0 \tag{8.73}$$

and there is (once again) a magic identity, $\gamma_0\gamma_\mu^\dagger\gamma_0 = \gamma_\mu$. Along with $\gamma_0^2 = 1$ this gives

$$\begin{aligned} 0 &= u^\dagger\gamma_0(\gamma_0\gamma_\mu^\dagger\gamma_0 p^\mu - m)\gamma_0 \\ &= \bar{u}(\not{p} - m). \end{aligned} \tag{8.74}$$

For the negative energy solutions we define $v(p,s)$ to be a negative energy solution with spin $(-s)$ in the rest frame, that is

$$\begin{aligned} w_3(p) &= v(p,\downarrow) \equiv v(p,-s) \\ w_4(p) &= v(p,\uparrow) \equiv v(p,s). \end{aligned} \tag{8.75}$$

We thus label an arbitrary spinor by its momentum, the sign of its energy, and the spin direction in the rest frame s. This seemingly odd definition of the v's will be very convenient when we come to interpret the negative energy states in the field operators. The v's obey

$$(\not{p} + m)v(p,s) = 0 \tag{8.76}$$

and

$$\bar{v}(p,s)(\not{p} + m) = 0. \tag{8.77}$$

Now for some useful properties of the u's and v's. One can show explicitly that

$$\bar{u}_i u_j = -\bar{v}_i v_j = |C|^2\delta_{ij} \tag{8.78}$$

(so these objects are scalars). A completeness relation is

$$\sum_i (u_i\bar{u}_i - v_i\bar{v}_i)_{\alpha\beta} = \delta_{\alpha\beta}|C|^2 \tag{8.79}$$

where α and β are Dirac indices. The minus sign can be motivated from $1 = u^\dagger u + v^\dagger v$ plus the minus sign in the Bjorken-Drell γ_0.

Projection operators are useful. If we have some general state ψ, it can contain both positive energy and negative energy solutions. Perhaps we only want the positive energy ones? Write

$$\sum_i u_i \bar{u}_i - v_i \bar{v}_i = \Lambda_+ + \Lambda_- \tag{8.80}$$

where $\sum_i u_i \bar{u}_i = \Lambda_+$ and $\sum_i -v_i \bar{v}_i = \Lambda_-$. W can find Λ_+ and Λ_- "by inspection:"

$$\Lambda_+ u_j = \sum_i u_i \bar{u}_i u_j = \sum_i u_i (\bar{u}_i u_j) = \sum_i u_i |C|^2 \delta_{ij} = |C|^2 u_j. \tag{8.81}$$

And we know $(\not{p} + m)u_j = 2mu_j$ so Λ_+ must be proportional to $(\not{p} + m)$. The answer is

$$\Lambda_+ = \frac{\not{p} + m}{2m} |C|^2 \tag{8.82}$$

and then

$$\Lambda_- = \frac{-\not{p} + m}{2m} |C|^2 \tag{8.83}$$

The sign is so that $\Lambda_+ + \Lambda_- = |C|^2 \delta_{ij}$.

Finally, what about $|C|^2$ There are two choices in the literature. First, one can take $|C|^2 = 1$. Then $\sum_i \psi_i^\dagger \psi_i = 1$. This means that $|\psi|^2$ is a probability, just like in ordinary quantum mechanics. This choice is made in old books like Bjorken and Drell and, more generally, in calculations where the fermions are taken to be heavy.

The other choice is $|C|^2 = 2m$ so that $\Lambda_\pm = \pm\not{p} + m$. In this convention, $u^\dagger(p,s)u(p,s') = E\delta_{ss'}$. Remember that $u^\dagger u$ is the zeroth component of the current four vector. E is the zeroth component of the momentum four vector. The two components have to be proportional to each other if we are dealing with plane waves. This is a very useful normalization convention for states with $E \gg m$, since we can just set $m = 0$ in our calculation and move on. With the first choice, and a $2m$ in the denominator, we have to do this cautiously. My convention is this one, $|C|^2 = 2m$, $\Lambda_\pm = \pm\not{p} + m$. Note that this is the same choice as was taken for scalar fields as described in the Appendix.

We are getting close writing down the field variables we will actually use in calculations. Define a plane wave state containing positive and negative energy states as

$$\psi(x,t) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2E(k)}} \sum_s [b(k,s)u(k,s)e^{-ik\cdot x} + d^\dagger(k,s)v(k,s)e^{ik\cdot x}] \tag{8.84}$$

The conjugate momentum is $\pi = i\bar{\psi}\gamma_0$ or $\pi = \psi^\dagger$. $b(p,s)$ and $d^\dagger(p,s)$ start out as classical Fourier coefficients. We make the usual conversion of the field ψ into a field operator by treating the classical Fourier coefficients as quantum operators. We want a theory of fermions, so we need to impose anticommutaton relations on them. The nonzero ones are

$$[b(k,s), b^\dagger(k',s')]_+ = [d(k,s), d^\dagger(k',s')]_+ = \delta_{ss'}\delta^3(\vec{k}-\vec{k}') \tag{8.85}$$

and we take all other anticommutators to vanish. It is a long boring exercise to show that this choice reproduces the appropriate anticommutator for π and ψ; the calculation involves all the projector identities we just wrote down.

And now for the Hamiltonian. It is easy:

$$\begin{aligned} \mathcal{H} &= \pi\dot{\psi} - \mathcal{L} \\ &= \psi^\dagger[-i\vec{\alpha}\cdot\nabla + \beta m]\psi \\ &= \psi^\dagger i\frac{\partial\psi}{\partial t} \end{aligned} \tag{8.86}$$

where the last line uses the Dirac equation. Plugging in the appropriate field definitions and recalling that the Hamiltonian is related to the its density by $H = \int d^3x\mathcal{H}$, we eventually arrive at

$$H = \sum_s \int d^3p E(p)[b^\dagger(p,s)b(p,s) - d(p,s)d^\dagger(p,s)]. \tag{8.87}$$

This is straightforward to derive, but we have a problem due to the $d(p,s)d^\dagger(p,s)$ term. By creating negative energy particles we can arbitrarily lower the energy of any state. The spectrum of H is unbounded from below.

Couldn't we forbid transitions to negative energy? Dirac's 1930 resolution of the problem was to assume that in nature all the negative energy levels are filled,

$$|vacuum\rangle = \prod d^\dagger(p,s)\,|0\rangle \tag{8.88}$$

and the Pauli principle forbids further occupying them. But we could remove states from this "Dirac sea." States we can access are ones like $b^\dagger(k_1,s_1)d(k_2,s_2)\,|vacuum\rangle$. This is a state with one positive energy particle and with the absence of a negative energy particle – a "hole." But physically, what were the holes? (This was an issue before the discovery of the positron.) Were they protons? (This was before the discovery of charge conjugation.)

Nowadays we just work with a normal ordered Hamiltonian

$$: H := \sum_s \int d^3p E(p)[b^\dagger(p,s)b(p,s) + d^\dagger(p,s)d(p,s)]. \tag{8.89}$$

The vacuum is just $|0\rangle$, annihilated by both $b(k,s)$ and $d(k,s)$. It has zero energy; states like $b^\dagger(k_1,s_1)d^\dagger(k_2,s_2)\,|0\rangle$ have (positive) energy $E(k_1)+E(k_2)$.

We had a homework problem earlier this semester involving electrons in a one dimensional solid which connects the two descriptions. The physics is identical to what we face now. Just workng with the normal ordered Hamiltonian means that at the end of the day, we have a lowest state, the vacuum, containing no fermions and no antifermions, and we have excitations all of whose energies are positive with respect to it. The spinor factors in $v(p,s)$ are designed to encode this reasonable situation as seamlessly as possible.

To reinforce the point (and hopefully not to belabor it), consider the Dirac equation for a particle of negative charge, coupled to electromagnetism:

$$[i\not\partial - e\not A - m]\psi = 0 \tag{8.90}$$

For a positron, a particle which has positive charge, it would be

$$[i\not\partial + e\not A - m]\psi^c = 0. \tag{8.91}$$

We could have equally as well began the story of the Dirac particle with positrons and interpreted the positive energy electrons as the absence of negative energy positrons. In fact, there is a symmetry which encodes this fact.

Do you recall the charge conjugation operator, which transforms Eq. 8.90 into Eq. 8.91? Let's construct the operator it again, but this time with a more physical argument. We have to flip the relative sign between $\not\partial$ and $e\not A$ to toggle between the two equations. Use the fact (we'll get to this in a while) that the vector potential operator creates and annihilates photons,

$$A_\mu = \int [b(k,\epsilon) + b^\dagger(k,\epsilon)] = A^\dagger \tag{8.92}$$

(or classically, $A_\mu = A_\mu^*$). Complex conjugate Eq. 8.90 and multiply by -1:

$$[(i\partial_\mu + eA_\mu)\gamma^{\mu *} + m]\psi^* = 0. \tag{8.93}$$

If we can find a transformation $(C\gamma_0)$ such that $(C\gamma_0)\gamma^{\mu *}(C\gamma_0)^{-1} = -\gamma^\mu$, Eq. 8.93 becomes

$$[i\not\partial + e\not A - m](C\gamma_0\psi^*) = 0. \tag{8.94}$$

and we find $C\gamma_0\psi^* = \psi^c$. The result: $C = i\gamma_2\gamma_0$ does the job.

To see what is happening, look at a negative energy electron, spin down, at rest, Bjorken and Drell γ convention,

$$\psi = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{imt} \tag{8.95}$$

The explicit form of the operator is

$$C\gamma_0 = i\gamma_2 = i \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} \tag{8.96}$$

and, just doing the multiplication and complex conjugation, the transformed field is

$$\psi^c = \psi = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{-imt} \tag{8.97}$$

Charge conjugation takes the negative energy electron into a positive energy one. Note that the absence of a spin down state is the same as the presence of a spin up state, in the spin counting. This is the origin of the flipped spin in the relation between $v(p,s)$ and $w(p,s)$ as previously alluded to.

The complete charge conjugation transformation includes

- $A_\mu \to -A_\mu$ (effectively flipping the sign of e)
- $\psi \to C\gamma_0\psi^*$

And, why are the masses of the electron and positron the same? Charge conjugation is a symmetry of the Dirac Lagrangian!

Finally, note that diagrams involving an odd number of external photons vanish due to charge conjugation. This happens because A is c-odd and

$$\langle T(AAA)\rangle = \left\langle T(CAC^{-1}CAC^{-1}CAC^{-1})\right\rangle = -\langle T(AAA)\rangle = 0 \tag{8.98}$$

This is called "Furry's theorem." A process where a photon splits into two photons is not allowed.

8.5 The spin-statistics theorem ($j = 0$ and $j = 1/2$ only)

When we wrote down Eq. 8.84 we simply stated that the field creation operators anticommute. Now we can check that that must be the case, that spin $\frac{1}{2}$ fields must be fermions. The argument parallels our demonstration that scalar fields, which annihilated particles, also created antiparticles.

To begin: in order that the S-matrix be Lorentz invariant, we needed that $[H(x), H(y)] = 0$ if $(x-y)^2 < 0$. Since H (and any other observable) is usually a bilinear or higher power in the field variables, we need either

$$[\phi(x), \phi^\dagger(y)] = 0 \tag{8.99}$$

or

$$[\phi(x), \phi^\dagger(y)]_+ = 0 \tag{8.100}$$

for spacelike separation, in order to satisfy this relation. We know that Eq. 8.99 is true for spin-zero fields when $[a(k), a^\dagger(k')] = \delta^3(k-k')$, but it is not true if the a's anticommute.

Recall the formula for scalars:

$$\begin{aligned}
[\phi(x), \phi^\dagger(y)]_\pm &= \int \frac{d^3k d^3k'}{(2\pi)^3\sqrt{4EE'}} \left([a(k)e^{-ikx} + a^{c\dagger}(k)e^{ikx}][a^\dagger(k')e^{ik'y} + a(k')e^{-ik'y}] \pm \dots \right) \\
&= \int \frac{d^3k d^3k'}{(2\pi)^3\sqrt{4EE'}} \left(e^{-i(kx-k'y)}[a(k), a^\dagger(k')]_\pm + e^{i(kx-k'y)}[a^{c\dagger}(k), a^c(k')]_\pm \right)
\end{aligned} \tag{8.101}$$

If the a's commute, the object in the large parentheses becomes $e^{-ik(x-y)} - e^{ik(x-y)}$ and the integral will vanish because $\Delta(x) = \Delta(-x)$ at spacelike separations. However, if the a's anticommute, we get a sum of Δ's, which will not vanish for spacelike separations and our requirement that the Hamiltonian commutes will fail.

Now we repeat the argument for spinors: relabel$a \to b$, $a^c \to d$.

$$\begin{aligned}
[\psi(x), \bar\psi(y)]_\pm &= \int \frac{d^3k d^3k'}{(2\pi)^3\sqrt{4EE'}} \Big(\sum_{ss'} u(k,s)\bar u(k',s')[a(k,s), a^\dagger(k',s')]_\pm e^{-i(kx-k'y)} \\
&\quad + \sum_{ss'} v(k,s)\bar v(k',s')[a^{c\dagger}(k,s), a^x(k',s')]_\pm e^{i(kx-k'y)} \Big).
\end{aligned} \tag{8.102}$$

If we assume that $[a, a^\dagger]_\pm = \delta_{ss'}\delta^3(k-k')$, we get

$$[\psi(x), \bar\psi(y)]_\pm = \int \frac{d^3k'}{(2\pi)^3 2E}\left(\sum_s u(k,s)\bar u(k,s)e^{-ik(x-y)} \pm v(k,s)\bar v(k,s)e^{ik(x-y)}\right). \quad (8.103)$$

The $\pm$ goes with the reverse order of a and $a^\dagger$ in the second (anti)commutator. We now use

$$\begin{aligned}\sum_s u(k,s)\bar u(k,s) &= \not k + m \\ \sum_s v(k,s)\bar v(k,s) &= \not k - m\end{aligned} \quad (8.104)$$

to simplify

$$[\psi(x), \bar\psi(y)]_\pm = \int \frac{d^3k}{(2\pi)^3 2E}\left(\not k[e^{-ik(x-y)} \pm e^{ik(x-y)}] + m[e^{-ik(x-y)} \mp e^{ik(x-y)}]\right) \quad (8.105)$$

With a commutator (the lower sign) the m terms add, but recall

$$\int \frac{d^3k}{2E}e^{ik(x-y)} \quad (8.106)$$

is even under exchange of x and y, for spacelike $x-y$, and so the commutator is nonzero. With the anticommutator, it vanishes. The first $\not k$ term is

$$\int \frac{d^3k}{2E}e^{-ik(x-y)}\not k = i\not\partial \int \frac{d^3k}{2E}e^{-ik(x-y)} \quad (8.107)$$

and so with the anticommutator we get $\not\partial_x[\Delta(x-y)-\Delta(y-x)]$; the argument in the brackets is zero because $\Delta(x-y)$ is even, and we get zero again. With the comutator, this term is nonzero. Thus, for spin-$\frac{1}{2}$ particles, only the anticommutator of the fields vanishes outside the light cone.

This is a specific example of the spin statistics theorem: Systems with integer spin must be quantized with commutation relations and hence are bosons. Systems with half integer spin require anticommutation relations and hence are fermions.

The original proof of the theorem was by Pauli, Phys. Rev. **58** 716 (1940). The article is reprinted in Schwinger, "Selected Papers on Quantum Electrodynamics," Dover Publications, 1958, ISBN 978-0-486-60444-2, 978-0-486-60444-2 (which also contains most of the seminal papers on QED). Incidentally, long before this, in 1924, he had proposed the exclusion principle for fermions, as an empirical observation.

8.6 The fermion propagator

The fermion propagator is

$$iS_F(x-y) = \left\langle 0|T(\psi(x),\bar{\psi}(y))|0\right\rangle \tag{8.108}$$

Field operators anticommute. It is necessary for the time-ordered product to respect this fact, so it must be that

$$\begin{aligned} T(\psi(x),\bar{\psi}(y)) &= \psi(x)\bar{\psi}(y); \qquad t_x > t_y \\ &= -\bar{\psi}(y)\psi(x) \qquad t_y > t_x \end{aligned} \tag{8.109}$$

Then

$$iS_F(x-y) = \int \frac{d^3p}{(2\pi)^3 2E(p)}[\theta(t)\Lambda_+(p)e^{-ipx} + \theta(-t)\Lambda_(p)e^{ipx}] \tag{8.110}$$

where (again)

$$\begin{aligned} \Lambda_+ &= \sum_s u(p,sP\bar{u}(p,s) = \not{p} + m \\ \Lambda_- &= \sum_s v(p,sP\bar{v}(p,s) = -\not{p} + m \end{aligned} \tag{8.111}$$

and, repeating the derivation for the propagator of the scalar field, we find

$$iS_F(q) = i\frac{\not{q} + m}{q^2 - m^2 + i\epsilon}. \tag{8.112}$$

The denominator is exactly as for a scalar field. The numerator reflects spinor structure – $iS_F(q)$ is a 4×4 matrix in spinor space.

8.7 A sample problem – an electron in an external field

We imagine that we have an electron which is scattered by an external (non-dynamical) electromagnetic field. The differential cross section will presumably resemble the case of Rutherford scattering. Let's do the calculation, as a way of introducing some necessary technology as much as for an answer to a physical question. Let's just assume that the

electromagnetic current is equal to $-e$ times the probability current, so that we have an interaction Hamiltonian

$$\mathcal{H}(x) = J_\mu A^\mu = -e\bar{\psi}(x)\gamma_\mu\psi(x)A^\mu(x). \tag{8.113}$$

You can show that this definition of the current is not unique – there are many possible terms in the most general expression for a current – but this is the only one whose integral corresponds to a conserved charge

$$Q = \int d^3x J_0(x). \tag{8.114}$$

And it is the actual electromagnetic current for a point particle, as we will later see.

We write the field operator as

$$\psi(x,t) = \sum_s \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2E(p)}}[b(p,s)u(p,s)e^{-ipx} + d^\dagger(p,s)v(p,s)e^{ipx}] \tag{8.115}$$

and our states are

$$|\psi(p\sigma)\rangle = b^\dagger(p,\sigma)\,|0\rangle\,. \tag{8.116}$$

The first order $T-$matrix is

$$T(p',\sigma';p,\sigma) = -i\int d^4x A_\mu(x)\,\langle\psi(p'\sigma')|J^\mu(x)|\psi(p\sigma)\rangle \tag{8.117}$$

where

$$\langle\psi(p'\sigma')|J^\mu(x)|\psi(p\sigma)\rangle = e\frac{e^{i(p-p')x}}{(2\pi)^3\sqrt{2E(p)2E(p')}}\bar{u}(p',\sigma' r)|\gamma^\mu u(p,\sigma) \tag{8.118}$$

To proceed, we take A to be time independent and do the time integral. This is scattering in an external field, just like potential scattering in nonrelativistic quantum mechanics. The electron's energy is conserved, but its momentum is not. We get

$$T(p',\sigma';p,\sigma) = -ie\bar{u}(p',\sigma')\gamma^\mu u(p,\sigma)A_\mu(p-p')\frac{1}{2E(p)}[2\pi\delta(E(p)-E(p')] \tag{8.119}$$

where

$$A_\mu(q) = \int \frac{d^3x}{(2\pi)^3}e^{-iqx}A_\mu(x). \tag{8.120}$$

To find the differential cross section, drop the $2\pi\delta()$ and square:

$$d\sigma = e^2|u(p',\sigma')\gamma^\mu u(p,\sigma)A_\mu(p-p')|^2\frac{1}{4E^2}\frac{(2\pi)^4}{v_{rel}}d^3p' \tag{8.121}$$

(To get the right factor of powers of 2π, think of the scattering as $p + P_1 \to p' + P_2$ where $p_1 = P_2 = (M, 0, 0, 0)$ and M is large; include fields for the two heavy states. The alternative is to set up the calculation from the start along the lines of nonrelativistic potential scattering in Born approximation.) We can then do the phase space integral, $d^3p' = p'^2 dp' d\Omega = p'(E'dE')d\Omega$, so

$$\frac{\delta(E - E')p'dE'}{4E^2 v_{rel}} = \frac{p}{4Ev_{rel}} \tag{8.122}$$

With $v_{rel} = p/E$, the differential cross section is

$$\frac{d\sigma}{d\Omega} = (2\pi)^4 \frac{e^2}{4} |u(p', \sigma')\gamma^\mu u(p, \sigma) A_\mu(p - p')|^2. \tag{8.123}$$

We have now hit the technical issue we always encounter when dealing with Dirac fermions – all these Dirac matrices and spinors. How can we evaluate the expression as painlessly as possible? Dealing with this issue involves a set of manipulations which go under the heading called the "trace calculus." Let's use this problem as an introduction to these manipulations.

The situation is simplest when we sum over final spins and average over initial spins of the fermions. Then, noting that $[\bar{u}(p', \sigma')\gamma^\mu u(p, \sigma)]^\dagger = [\bar{u}(p, \sigma)\gamma^\nu u(p', \sigma')]$, we have

$$\frac{d\sigma}{d\Omega} = (2\pi)^4 \frac{e^2}{4} \sum_{\sigma, \sigma'} A_\mu A_\nu^* [\bar{u}(p', \sigma')\gamma^\mu u(p, \sigma)][\bar{u}(p, \sigma)\gamma^\nu u(p', \sigma')]. \tag{8.124}$$

Let's re-introduce the spinor indices on the fermionic terms: they are

$$F_{\mu\nu} = \sum_{\sigma'} \bar{u}(p', \sigma')_\alpha \gamma^\mu_{\alpha\beta} \sum_\sigma u(p, \sigma)_\beta \bar{u}(p, \sigma)_\gamma \gamma^\nu_{\gamma\delta} u(p', \sigma')_\delta. \tag{8.125}$$

Recall the projector definition: $\sum_\sigma u(p, \sigma)_\beta \bar{u}(p, \sigma)_\gamma = (\not{p} + m)_{\beta\gamma}$. Using this lets us write

$$F_{\mu\nu} = \gamma^\mu_{\alpha\beta} (\not{p} + m)_{\beta\gamma} \gamma^\nu_{\gamma\delta} (\not{p}' + m)_{\delta\alpha}. \tag{8.126}$$

Look at the indices – this expression is a trace:

$$\frac{d\sigma}{d\Omega} = (2\pi)^4 \frac{e^2}{4} \sum_{\sigma, \sigma'} A_\mu A_\nu^* \mathrm{Tr}\,[\gamma^\mu (\not{p} + m)\gamma^\nu (\not{p}' + m)]. \tag{8.127}$$

There are a set of tricks which are used again and again to reduce expressions like Eq. 8.127 – the "trace rules." They descend from the fundamental relation

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}. \tag{8.128}$$

Trace over this: because $\mathrm{Tr}\,1 = 4$ and because the trace is cyclic,

$$\mathrm{Tr}\,\gamma^\mu\gamma^\nu = \mathrm{Tr}\,\gamma^\nu\gamma^\mu = 4g^{\mu\nu}. \tag{8.129}$$

This allows us to write the useful result that

$$\begin{aligned} \mathrm{Tr}\,\not{a}\not{b} &= a_\mu b_\nu \mathrm{Tr}\,\gamma^\mu\gamma^\nu \\ &= a_\mu b_\nu 4g^{\mu\nu} \\ &= 4a\cdot b. \end{aligned} \tag{8.130}$$

Another useful rule is that the trace of an odd number of γ's is zero. Check this for three γ's:

$$\begin{aligned} \mathrm{Tr}\,\not{a}\not{b}\not{c} &= \mathrm{Tr}\,\not{a}\not{b}\not{c}\gamma_5\gamma_5 \\ &= \mathrm{Tr}\,\gamma_5\not{a}\not{b}\not{c}\gamma_5 \\ &= -\mathrm{Tr}\,\not{a}\not{b}\not{c}\gamma_5\gamma_5 \\ &= 0. \end{aligned} \tag{8.131}$$

In the first line, $\gamma_5^2 = 1$. Then in the second line, the trace is cyclic, and in the third line, anticommute γ_5 with all the other γ's to return it to where it started.

Not a trace rule, but another useful fact is $\overline{\not{a}\not{b}\not{c}\ldots} = \ldots\not{c}\not{b}\not{a}$. And finally (for now) can you show this one?

$$\mathrm{Tr}\,\not{a}\not{b}\not{c}\not{d} = 4[(a\cdot b)(c\cdot d) - (a\cdot c)(b\cdot d) + (a\cdot d)(b\cdot c)]. \tag{8.132}$$

Our expression Eq. 8.127 thus simplifies to

$$\mathrm{Tr}\,\gamma^\mu\not{p}\gamma^\nu\not{p}' + m^2\gamma^\mu\gamma^\nu = 4[g^{\mu\nu}m^2 + p^\nu p'^\nu + p^\nu p'^\mu - g^{\mu\nu}p\cdot p'] \tag{8.133}$$

so that the differential cross section is

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= (2\pi)^4\frac{e^2}{2}A_\mu A_\nu^*(p^\nu p'^\nu + p^\nu p'^\mu - g^{\mu\nu}(p\cdot p' - m^2)). \\ &= (2\pi)^4\frac{e^2}{2}[(A\cdot p)(A^*\cdot p') + (A\cdot p')(A^*\cdot p) - |A|^2(p\cdot p' - m^2)]. \end{aligned} \tag{8.134}$$

An interesting special case of this expression is Coulomb scattering, $\vec{A} = 0$, $A^0 = Ze/(4\pi r)$. (I am working in Lorentz - Heaviside units where $e^2/(4\pi) = \alpha = 1/137$.) The Fourier transform is

$$\begin{aligned} A^0(q) &= \frac{Ze}{4\pi}\int [\frac{e^{iqr\cos\theta}}{r}]\frac{2\pi r^2 dr d\cos\theta}{(2\pi)^3} \\ &= \frac{Ze}{4\pi}\frac{2\pi}{(2\pi)^3}\int_0^\infty r dr \left[\frac{e^{iqr}-e^{-iqr}}{iqr}\right] e^{-\epsilon r} \\ &= \frac{Ze}{4\pi}\frac{1}{(2\pi)^2}\frac{1}{iq}\left[\frac{1}{-iq+\epsilon}-\frac{1}{iq+\epsilon}\right] \\ &= \frac{Ze}{4\pi}\frac{2}{(2\pi)^2}\frac{1}{|\vec{q}|^2} \end{aligned} \tag{8.135}$$

This gives

$$\frac{d\sigma}{d\Omega} = (2\pi)^4\frac{e^2}{2}(\frac{Ze}{4\pi})^2\frac{4}{(2\pi)^4}\frac{1}{|\vec{q}|^4}[2E^2-(p\cdot p'-m^2)^2] \tag{8.136}$$

With $q^2 = -\vec{q}^2 = (p-p')^2 = 2m^2 - 2p\cdot p'$, $p\cdot p' - m^2 = q^2/2$, and the expression in square brackets is $2E^2 - \vec{q}^2/2$. We arrive at the Mott formula (1929),

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2\alpha^2}{|\vec{q}|^4}[E^2 - \frac{\vec{q}^2}{4}]. \tag{8.137}$$

Note that as $q \ll E$, we have $E \approx m$ and

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2\alpha^2 m^2}{|\vec{q}|^4}, \tag{8.138}$$

which is the Rutherford formula.

For completeness, let's look at Coulomb scattering for positrons. We begin with

$$\langle p', s'|J^\mu(x)|p, s\rangle = \left\langle 0|d(p', s')\bar{\psi}\gamma^\mu\psi d^\dagger(p, s)|0\right\rangle \tag{8.139}$$

ψ is still given by Eq. 8.115, but the contraction of creation and annihilation operators is

$$\begin{aligned} & -e\left\langle 0|d(p', s')d(k_2, s_2)d^\dagger(k_1, s_1)d^\dagger(p, s)|0\right\rangle \bar{v}(k_2, s_2)\gamma^\mu v(k_1, s_1) \\ = \; & +e\bar{v}(p, s)\gamma^\mu v(p', s')\exp(-i(p-p')x) \end{aligned} \tag{8.140}$$

There is a minus sign after anticommuting the creation and annihilation operators in the ψ's to contract against the states, and the momentum order is reversed from the case of

electrons because the ψ field annihilates fermions but creates antifermions. The differential cross scattering for Coulomb scattering becomes

$$\frac{d\sigma}{d\Omega} = \frac{2Z^2\alpha^2}{|\vec{q}|^4} \sum_{ss'} |\bar{v}(p,s)\gamma^0 v(p',s')|^2. \tag{8.141}$$

The trace is evaluated as before, but with projection operators $\sum_\sigma v(p,\sigma)_\alpha \bar{v}(p,\sigma)_\beta = (-\not{p} + m)_{\alpha\beta}$ so that we must evaluate

$$\mathrm{Tr}\,[\gamma^0(-\not{p}+m)\gamma^0(-\not{p}'+m)] = \mathrm{Tr}\,[\gamma^0(\not{p}-m)\gamma^0(\not{p}'-m)] \tag{8.142}$$

compared to the electron trace. This will give our previous answer since the result only depended on m^2, not m.

There is a conventional shorthand for handling ordering problems like this, which is to follow the fermion world line. See Fig. 8.1. The idea is that an electron is a fermion going forward in time. An amplitude with a forward-going fermion is

$$M(p \to p'') \propto -e\bar{u}(p')\gamma u(p) \equiv -e\bar{f}(p')\gamma f(p). \tag{8.143}$$

A positron is a fermion going backward in time, so

$$M(p \to p'') \propto e\bar{v}(p)\gamma v(p') \equiv e\bar{f}(-p)\gamma f(-p). \tag{8.144}$$

The polarization sum for a fermion is

$$\sum_s f(p)\bar{f}(p) = \not{p} + m \tag{8.145}$$

which takes care of the sign flip between the projection formula for u's and v's. With this convention, there is no problem getting the propagator correct: see the example in the figure.

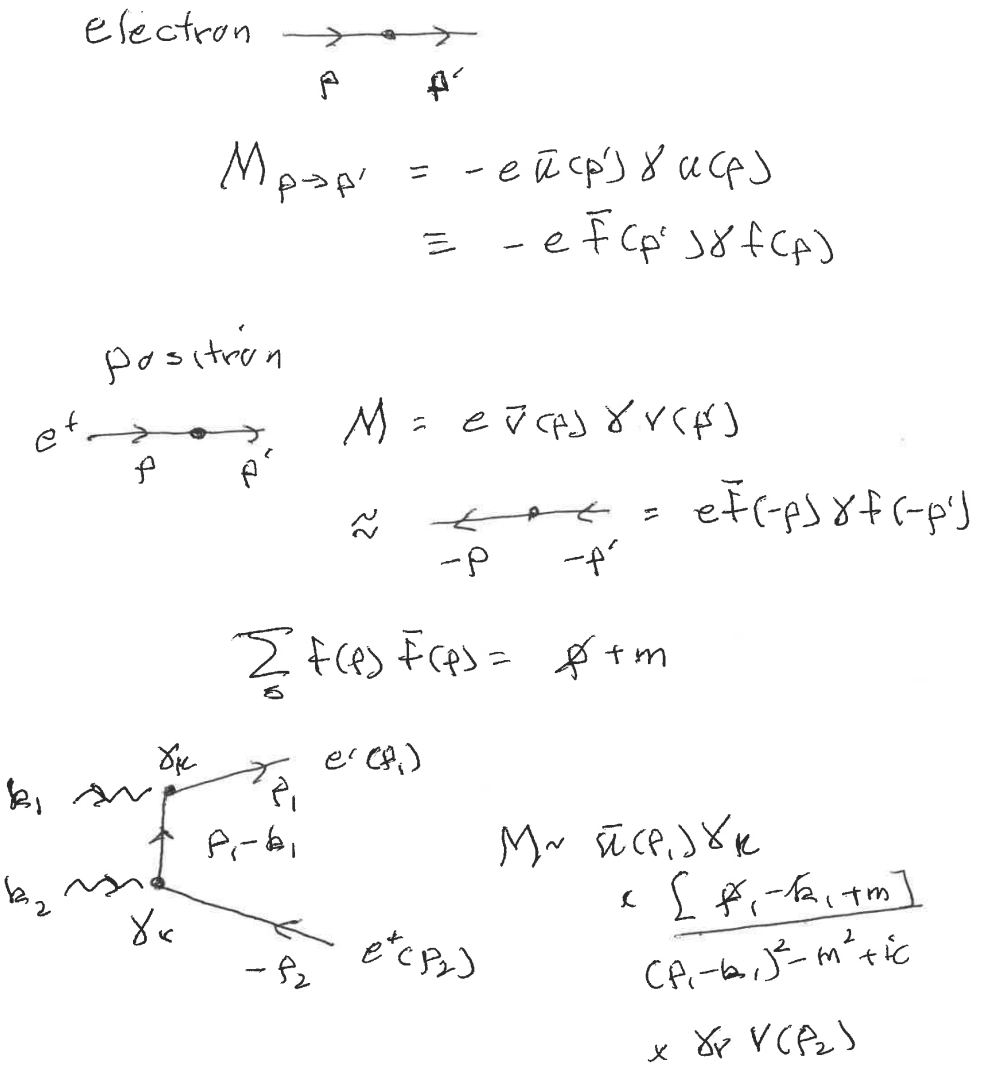

Figure 8.1: Following the fermion world line.

Chapter 9

Quantum electrodynamics, for itself and as a example of a gauge theory

9.1 Symmetries and conservation laws

It is often useful to identify symmetries of the classical Lagrangian. It turns out that for each continuous symmetry of $\mathcal{L}$ there is an associated conserved current and conserved charge. The relation is called "Noether's theorem." In quantum systems these conserved quantities often become constraints on physical states.

Let's work this out for fields ϕ and Lagrangfe density $\mathcal{L}(\phi_j, \partial_\mu \phi_j)$. Recall $\partial_\mu \phi = \partial\phi/\partial x^\mu$. The Lagrange equation of motion is

$$\frac{\partial \mathcal{L}}{\partial \phi_j} = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \right]. \tag{9.1}$$

Imagine a transformation $\phi_j \to \phi_j + \delta\phi_j$. The change in the Lagrangian is

$$\delta\mathcal{L} = \sum_j \frac{\partial \mathcal{L}}{\partial \phi_j} \delta\phi_j + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \delta(\partial_\mu \phi_j) \tag{9.2}$$

which becomes, invoking the equation of motion,

$$\delta\mathcal{L} = \sum_j \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \right] \delta\phi_j + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \delta(\partial_\mu \phi_j). \tag{9.3}$$

The two generic situations of symmetry transformations involve space-time transformations or internal ones. An example of a space-time symmetry is a translation $x^\mu \to x^\mu + \epsilon^\mu$. $\mathcal{L}$ does not depend directly on coordinates. It depends on field variables which themselves depend on coordinates. Varying x gives the variations

$$\begin{aligned} \delta\mathcal{L} &= \epsilon^\mu \partial_\mu \mathcal{L} \\ \delta\phi &= \epsilon^\mu \partial_\mu \phi \\ \delta(\partial_\mu \phi) &= \epsilon^\nu \partial_\nu (\partial_\mu \phi) = \epsilon^\nu \partial_\mu \partial_\nu \phi \end{aligned} \tag{9.4}$$

Then

$$\delta\mathcal{L} = \epsilon^\mu \partial_\mu \mathcal{L} = \sum_j \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \right] \epsilon^\nu \partial_\nu \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \epsilon^\nu \partial_\mu \partial_\nu \phi \tag{9.5}$$

or

$$\epsilon^\nu \partial_\mu [\delta^\mu_\nu \mathcal{L}] = \epsilon^\nu \sum_j \partial_\mu \left[\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \right] \partial_\nu \phi_j \right] \tag{9.6}$$

or

$$\epsilon^\nu \partial_\mu \left[\delta^\mu_\nu \mathcal{L} - \sum_j \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \partial_\nu \phi_j \right] = 0. \tag{9.7}$$

This is a conservation law for the quantity in the square brackets. Mixed incides are awkward, so raise the "ν" with a $g^{\nu\lambda}$ to write the conservation law as

$$\partial_\mu T^{\mu\nu} = 0 \tag{9.8}$$

wherea the conserved quantity

$$T^{\mu\nu} = -g^{\mu\nu}\mathcal{L} + \sum_j \partial^\nu \phi_j \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \right]] \tag{9.9}$$

is called the "energy-momentum" or "stress" tensor. Its conservation is encoded in Eq. 9.8.

Integrals of $T^{\mu\nu}$ are familiar (?) conserved quantities (see Jackson, for example),

$$P^\nu = \int d^3x T^{0\nu} \tag{9.10}$$

with $dP^\nu/dt = 0$. Another element of the tensor is

$$T^{00} = -\mathcal{L} + \sum_j \dot{\phi}_j \pi_j = \mathcal{H} \tag{9.11}$$

and

$$\frac{d}{dt} \int d^3x \mathcal{H} = 0 \tag{9.12}$$

is energy conservation.

Now for an example.

In the Dirac equation, ψ is the dynamical variable, the Lagrange density is $\mathcal{L} = \bar{\psi}[i\gamma^\mu \partial_\mu - m]\psi$, the canonical momentum is $\pi = i\bar{\psi}\gamma_0 = i\psi^\dagger$ and

$$T^{00} = i\psi^\dagger \frac{\partial \psi}{\partial x^\nu} \tag{9.13}$$

and the conserved total energy is

$$P^0 = \int d^3x T^{00} = -\int d^3x \frac{1}{i} \psi^\dagger \frac{\partial \psi}{\partial x^0} = \int d^3x \psi^\dagger i \frac{\partial \psi}{\partial t}. \tag{9.14}$$

Internal symmetries are often more interesting. Here we imagine a change of field variables

$$\delta\phi_i(x) = -i\epsilon\lambda_{ij}\phi_j(x) \tag{9.15}$$

The index i is internal. An example would be a situation with N identical scalar fields ϕ_i and the symmetry transformation would be a rotaton in basis among them. The i on the right hand side of Eq. 9.15 is conventional and the λ_{ij}'s are a set of coefficients which are constant in space (so we are dealing with what is called a "global" symmetry transformation.) If the transformation leaves $\mathcal{L}$ invariant, then

$$\delta\mathcal{L} = 0 = \sum_i \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta(\partial_\mu\phi_i) \tag{9.16}$$

and $\delta(\partial_\mu\phi_i) = \partial_\mu(\delta\phi_i) = -i\epsilon\lambda_{ij}\partial_\mu\phi_j$. Again, the equation of motion

$$\frac{\partial\mathcal{L}}{\partial\phi_i} = \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right] \tag{9.17}$$

can be used to rewrite the conservation law as

$$0 = -i\sum_j \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right]\epsilon_{ij}\phi_j + \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right]\epsilon_{ij}\partial_\mu\phi_j \tag{9.18}$$

which is a statement that there is a conserved current

$$\partial_\mu J^\mu = 0 \tag{9.19}$$

where

$$J^\mu = -i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\lambda_{ij}\phi_j. \tag{9.20}$$

Associated with the current there is a conserved charge

$$Q = \int J^0(x)d^3x. \tag{9.21}$$

For example, ordinary quantum mechanics is invariant under a global phase rotation $\psi(x) \to e^{i\theta}\psi(x)$. If θ is small, we can write this as $\psi(x) \to= (1 + i\theta)\psi(x)$, i. e. $\delta\psi = i\theta\psi$. The conserved current is the usual probability current.

In the Dirac Lagrangian $\mathcal{L} = \bar{\psi}[i\gamma^\mu\partial_\mu - m]\psi$, so the conserved current is

$$J^\mu = (-i)(i\bar{\psi}\gamma^\mu)\psi = \bar{\psi}\gamma^\mu\psi \tag{9.22}$$

(note $i\bar{\psi}\gamma^\mu = \partial\mathcal{L}/\partial(\partial_\mu\psi)$).

One could also think of a complex ϕ as two real fields, $\phi = \phi_1 + i\phi_2$. $\phi(x) \to e^{i\theta}\phi(x)$ is equivalent to

$$\begin{aligned}\delta\phi_1 &= -\theta\phi_2 \\ \delta\phi_2 &= \theta\phi_1\end{aligned} \tag{9.23}$$

or $-i\lambda_{12} = \theta$, $-i\lambda_{21} = \theta$ in Eq. 9.15. A potential term such as $V(\psi) = V(\psi_1^2 + \psi_2^2)$ obviously preserves the symmetry. Can you check that this is also true for a derivative term $(\partial_\mu \phi)^\dagger \partial_\mu \phi$?

9.2 Beginning to think about quantum electrodynamics

We can now begin to think about quantum electrodynamics, an interacting system of a charged matter fields (which we will take to a fermion) and a vector field (the photon). The outline of our plan of attack is as follows:

1. Motivate a $\mathcal{L}$ by thinking about gauge invariance
2. Pass from an interaction $\mathcal{L}$ to a $\mathcal{H}$ suitable for perturbative calculations
3. Work out the photon propagator
4. Perform some sample calculations

Let us return to our example of a global symmetry transformation, but this time with a slight change of notation. Here we have N kinds of fermions, and rotate each of them via $\delta\psi_n = i\epsilon q_n \psi_n$; that is, $\theta_n \leftrightarrow \epsilon q_n$. Obviously, we have a set of conserved currents $J_n^\mu \propto \bar{\psi}_n \gamma^\mu \psi_n$.

To proceed to describe electrodynamics, we promote ϵ from a global symmetry to a local one, by imagining it to be coordinate dependent. We imagine making a local transformation

$$\delta\psi_n(x) = i\epsilon(x) q_n \psi_n(x) \tag{9.24}$$

and ask whether this could be a symmetry. Following our noses, we compute

$$\begin{aligned}
\delta\mathcal{L} &= \sum_n \frac{\partial \mathcal{L}}{\partial \psi_n} i\epsilon(x) q_n \psi_n + \sum_n \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_n)} \partial_\mu (i\epsilon(x) q_n \psi_n) \\
&= i\partial_\lambda \left(\sum_n \frac{\partial \mathcal{L}}{\partial(\partial_\lambda \psi_n)} q_n \epsilon(x) \psi_n \right) \\
&= -\partial_\lambda [\epsilon(x) J^\lambda] \\
&= -(\partial_\lambda \epsilon) J^\lambda
\end{aligned} \tag{9.25}$$

where the last line is what we get because $\partial_\lambda J^\lambda = 0$. Eq. 9.24 does not appear to be a viable symmetry transformation on its own.

However, suppose we add another ingredient to $\mathcal{L}$, a vector field, whose simultaneous variation cancels the $\delta\mathcal{L}$ of the matter fields:

$$\delta A_\mu = \partial_\mu \epsilon(x). \tag{9.26}$$

Then, writing $\mathcal{L}_M$ as the matter Lagrangian which was invariant under global transformations, we have

$$\delta\mathcal{L} = \delta\mathcal{L}_M + \frac{\partial\mathcal{L}}{\partial A_\lambda}\delta_\lambda\epsilon + \frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\lambda}\partial_\nu\partial_\lambda\epsilon \tag{9.27}$$

We already found that $\delta\mathcal{L}_M = -(\partial_\lambda\epsilon)J^\lambda$, so $\delta\mathcal{L}' = 0$ as long as

$$\frac{\partial\mathcal{L}}{\partial A_\lambda} = J^\lambda \tag{9.28}$$

that is, the vector field couples to the conserved matter-field current which arose from the global symmetry transformation, and derivatives of A_μ are present in the Lagrangian with the constraint that

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\mu)}. \tag{9.29}$$

We solve the first constraint with an interaction term

$$\mathcal{L}_I = J_\mu A^\mu. \tag{9.30}$$

We solve the second constraint by specifying some Lagrangian only involving the A's.

The true symmetry we have encoded involves both the matter fields and the new vector (or gauge) fields,

$$\begin{aligned} A_\mu(x) &\to A_\mu + \partial_\mu\chi(x) \\ \psi_n(x) &= \exp(iq_n\chi(x))\psi_n(x) \end{aligned} \tag{9.31}$$

The symmetry is called a "local gauge transformation." We introduce quantum electrodynamics by imposing the symmetry from the start and constructing a dynamics which respects it. (A similar story, with a more conplicated local gauge symmetry, would also give us the entire Standard Model.

So to summarize: we start with

$$\mathcal{L} = \mathcal{L}_G + \mathcal{L}_M \tag{9.32}$$

(absorbing $\mathcal{L}_I$ into $\mathcal{L}_M$ for the moment). We solve the constraint Eq. 9.29 with

$$\mathcal{L}_G(A) = \mathcal{L}_G(F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu) \tag{9.33}$$

that is, the vector potential only appears in the action through the field strength tensor $F_{\mu\nu}$. The simplest quadratic scalar Lagrangian (which of course gives us Maxwell electrodynamics) is

$$\mathcal{L}_G = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{9.34}$$

The prefactor of $-1/4$ gives us electrodynamics in Lorentz-Heaviside units. To build $\mathcal{L}_M$, we can simply replace ∂_μ by the covariant derivative $D_\mu = \partial_\mu - iq_n A_\mu$. This compactly encodes gauge invariance: The expression $D_\mu \psi_n$ transforms covariantly (and identically to how ψ_n transforms),

$$\begin{aligned} D'_\mu \psi'_n &= [\partial_\mu - iq_n(A_\mu + \partial_\mu \chi)] e^{iq_n \chi} \psi_n \\ &= e^{iq_n \chi} [\partial_\mu - iq_n A_\mu - iq_n \partial_\mu \chi + iq_n \partial_\mu \chi] \psi_n \\ &= e^{iq_n \chi} D_\mu \psi_n. \end{aligned} \tag{9.35}$$

Then a gauge invariant $\mathcal{L}_M$ is

$$\begin{aligned} \mathcal{L}_M &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \\ &= i\bar{\psi}\gamma^\mu (\partial_\mu - ieA_\mu)\psi - m\bar{\psi}\psi. \end{aligned} \tag{9.36}$$

Looking ahead to perturbative calculations, we can write this as

$$\mathcal{L}_M = \mathcal{L}_0 + \mathcal{L}_I \tag{9.37}$$

where the free Lagrange density is

$$\mathcal{L}_0 = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi \tag{9.38}$$

and

$$\mathcal{L}_I = e\bar{\psi}\gamma^\mu \psi A_\mu \equiv J^\mu A_\mu \tag{9.39}$$

is the interactxaooin term. $\mathcal{L}_0$ gives us the fermion propagator, which we have already determined. $\mathcal{L}_I$ gives us a vertex. But we have one new ingredient, the vector field A_μ, and its new symmetry, the local gauge transformation. Let's look at these objects more closely.

9.3 Quantizing gauge theories

Let's look at the photon by itself. "Quantizing a gauge theory" means passing from a Lagrangian to a Hamiltonian (and then to a propagator, if we want to do perturbative calculations). This turns out to be complicated. The issue is local gauge invariance. Problems appear working either with a Hamiltonian or with the path integral. The source of the problem is that gauge theories have redundant degrees of freedom. The statement that

$$A_\mu \to A_\mu + \partial_\mu \chi \tag{9.40}$$

is a symmetry says that a gauge transformation of A does not change the physics. We have to preserve this symmetry as we proceed from Lagrangian to Hamiltonian to propagator.

Let's now restrict the discussion to consider only an Abelian gauge theory (QED), since the case of non-Abelian gauge theories is much more difficult to deal with.

Let's start with the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{9.41}$$

The field variable is $A_\mu(x)$. We need a canonical momentum for each A_μ,

$$\pi_\mu(x) = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu}. \tag{9.42}$$

For A_i this is actually the electric field $\vec{\pi} = \vec{E}$, or $\pi_i = F_{0i} = E_i$. However, there is no $\dot{A}_0$ in $\mathcal{L}$, so there is no π_0 conjugate to A_0. We have a mismatch between the number of fields (four A_μ's) and conjugate momenta (three E_i's). In fact, it is worse: Gauss' law tells us that $\vec{\nabla} \cdot \vec{E} = \rho$. This says that the three π_i's are not independent. We only have two independent field momenta.

One way to deal with this issue is to make a direct attack on the gauge symmetry problem and do a complete gauge fixing before we go any further. One such choice is Coulomb gauge,

$$\vec{\nabla} \cdot \vec{A} = 0. \tag{9.43}$$

We live with the fact that A_0 is constrained by directly solving

$$\nabla^2 A_0 = 4\pi\rho \tag{9.44}$$

and then invoking $\vec{E} = -\vec{\nabla} A_0 - \frac{\partial \vec{A}}{\partial t}$. Then

$$\vec{\nabla} \cdot \vec{E} = -\nabla^2 A_0 - \frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{A}) = -4\pi\rho. \tag{9.45}$$

so that A_0 is just a solution of the (static) equation

$$\nabla^2 A_0 = -4\pi\rho. \tag{9.46}$$

A_0 is not an independent dynamical variable in Coulomb gauge.

We can expand the time dependent and transverse vector potential in terms of Fourier modes:

$$\vec{A}(x,t) = \sum_\lambda \int \frac{d^3k}{(2\pi)^3\sqrt{2\omega_k}} \left[a(k,\lambda)\epsilon(k,\lambda)e^{-ikx} + a^*(k,\lambda)\epsilon(k,\lambda)e^{-kx}\right] \tag{9.47}$$

(summing over polarizations $\epsilon(k,\lambda)$ for the two choices labeled by λ). The gauge choice $\vec{\nabla}\cdot\vec{A} = 0$ means that $\vec{k}\cdot\vec{A} = 0$ or $\vec{k}\cdot\vec{\epsilon} = 0$, the usual transversality condition. Then there are two independent components of $\vec{A}$. They each have a conjugate $\vec{E}$, so the counting of field and momentum degrees of freedom is consistent. All of the $\vec{A}$'s and $\vec{E}$'s are transverse.

So much for classical physics. As usual, the classical Fourier coefficients $a(k,\lambda)$ and $a^*(k,\lambda)$'s become the operators $a(k,\lambda)$ and $a^\dagger(k,\lambda)$'s after quantization, with the usual commutation relations

$$[a(k,\lambda), a^\dagger(k',\lambda) = \delta_{\lambda,\lambda'}\delta^3(k-k'). \tag{9.48}$$

Now for the propagator. Temporarily pretending that there are four A's (A_μ), the photon propagator would be

$$\begin{aligned} iD^{tr}(x-x')_{\mu\nu} &= \langle 0|T(A_\mu(x)A_\nu(x')|0\rangle \\ &= \sum_\lambda \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left[\theta(t'-t)e^{-ik\cdot(x-x')}\epsilon_\mu\epsilon_\nu^* + \theta(t-t')e^{ik\cdot(x-x')}\epsilon_\mu^*\epsilon_\nu\right] \end{aligned} \tag{9.49}$$

where ϵ_μ has $\epsilon_0 = 0$ and $\vec{\epsilon}\cdot\vec{k} = 0$. There are two nonvanishing components of the polarization vector. Doing the integrals, we arrive at a propagator

$$iD^{tr}(x-x')_{\mu\nu} = \int \frac{d^4k}{(2\pi)^4} e^{ik(x-x')}\frac{i}{k^2+i\epsilon}\sum_\lambda \epsilon_\nu(k,\lambda)\epsilon_\mu^*(k,\lambda). \tag{9.50}$$

What is the polarization matrix? If we pick $\vec{k}$ to lie in the z direction, with $\vec{\epsilon}(k,1) = \hat{x}$

and $\epsilon(k,2) = \hat{y}$, we could define it as

$$\begin{aligned} \pi_{ij} &= \sum_\lambda \epsilon_\nu(k,\lambda)\epsilon^*_\mu(k,\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \delta_{ij} - \hat{n}_i\hat{n}_j \\ &= \delta_{ij} - \frac{k_i k_j}{|\vec{k}|^2} \end{aligned} \tag{9.51}$$

writing $\hat{n}_i$ as a unit vector along $\vec{k}_i$ and then generalizing our formula to an arbitrary direction for $\vec{k}$. This give a momentum space propagator

$$iD^{tr}(x-x')_{ij} = \frac{i}{k^2 + i\epsilon}\left[\delta_{ij} - \frac{k_i k_j}{|\vec{k}|^2}\right] \tag{9.52}$$

This is obviously noncovariant, just like the gauge choice was noncovariant.

The Coulomb propagator is static, and of course it is also noncovariant. Just Fourier transforming $1/|x-x'|$, we have

$$D^{coul}(x-x')_{\mu\nu} = \frac{1}{|\vec{k}|^2}\delta_{\mu 0}\delta_{\nu 0}. \tag{9.53}$$

It's a reasonably nontrivial (and to me, very contrived) exercise to show that the combination of two noncovariant propagators (the static one and the transverse one) combine to give a covariant propagator. Many quantum field theory books work this out explicitly. Of course, we expect things to work out because the original action was covariant, but still, it is an annoying situation.

Let's try writing a path integral expression for the propagator, instead. Now the problem is that gauge invariance induces flat directions in the path integral making it impossible to invert the expression for the action into an expression for the propagator. We'll start with

$$\begin{aligned} S &= -\frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu} \\ &= -\frac{1}{4}\int d^4x [\partial_\mu A_\nu - \partial_\nu A_\mu][\partial_\mu A^\nu - \partial^\nu A^\mu]. \end{aligned} \tag{9.54}$$

Go to momentum space, using

$$A_\mu(x) = \int d^4p (2\pi)^2 e^{ip\cdot x} A_\mu(p). \tag{9.55}$$

Then the action is

$$\begin{aligned} S &= \frac{1}{4}\int \frac{d^4p}{(2\pi)^4}[p_\mu A_\nu - p_\nu A_\mu][p^\mu A^\nu - p^\nu a^\mu] \\ &= \frac{1}{2}\int \frac{d^4p}{(2\pi)^4} A^*_\mu(p) A_\nu(p)[g^{\mu\nu}p^2 - p^\mu p^\nu]. \end{aligned} \tag{9.56}$$

The propagator is the inverse of the 4×4 matrix in the square brackets in this expression. This has zero eigenvalues for all values of p, so the matrix is non-invertable. As an example, set $p^\mu = (p, 0, 0, 0)$. In that case

$$[g^{\mu\nu}p^2 - p^\mu p^\nu] = \begin{pmatrix} p^2 - p^2 & 0 & 0 & 0 \\ 0 & -p^2 & 0 & 0 \\ 0 & 0 & -p^2 & 0 \\ 0 & 0 & 0 & -p^2 \end{pmatrix}. \tag{9.57}$$

which obviously has a zero determinant. We have to stop and think.

9.4 Gaussian integrals with zero eigenvalues

Let's consider a Gaussian integral

$$G(A) = \int_{-\infty}^{\infty} dx_1 \dots dx_N \exp(-x^T A x) = \frac{1}{\sqrt{\det A}} \tag{9.58}$$

(again neglecting factors of π). This integral is well defined as long as the determinant does not vanish. If the determinant does vanish, we have a problem – the integral blows up.

We can see the origin of this blowup if we change variables to a basis where A is diagonal. If A is a real, symmetric $N \times N$ matrix, then we can rotate x with an orthogonal matrix R,

$$y = Rx. \tag{9.59}$$

The integration measure is unchanged,

$$\prod_{i=1}^{N} dx_i = \prod_{i=1}^{N} dy_i \tag{9.60}$$

so our integral becomes

$$\begin{aligned} G(A) &= \int \prod dy \exp(-x^T R^T A R x) \\ &= \int \prod dy_i \exp(-y_i^2 d_i) \\ &= \sqrt{\frac{\pi}{\prod_i d_i}}. \end{aligned} \tag{9.61}$$

We have problems (we get an infinity) of one of the d_i's is zero. In that case, it comes from

$$\int_{-\infty}^{\infty} dy = \infty. \tag{9.62}$$

Can we get a sensible answer for the integral even when the determinant vanishes? If we can divide out the culprit infinite integral, the answer is Yes. How can we do this? Suppose we have n zero eigenvalues. Let's define the restricted Gaussian integral

$$G_R(A) = \int dy_1 \dots dy_{N-n} \exp(-x^T A x). \tag{9.63}$$

This is awkward, since we have to know the y's before hand. Instead, define new variables y_{N-n+1} to y_N and write

$$G_R(A) = \int dy_1 \dots dy_{N-n} dy_{N-n+1} \dots dy_N \delta(y_{N-n+1}) \dots \delta(y_N) \exp(-x^T A x). \tag{9.64}$$

We can now change variables back from y to x,

$$dy_1 \dots dy_N = dx_1 \dots dx_N \det|\frac{\partial y}{\partial x}| \tag{9.65}$$

so that

$$G_R(A) = \int \prod_{i=1}^{N} dx_i \{\det|\frac{\partial y}{\partial x}| \prod_{j=N-n+1}^{N} \delta(y_j)\} \exp(-x^T A x). \tag{9.66}$$

The extra factors, the terms inside the curly brackets, restrict the integration from its original $N-$dimensional space to an $N-n$ dimensional one. Of course, one must cleverly choose the y's to restrict the integration measure, or else the Jacobian $\det|\partial y/\partial x|$ will be singular.

9.5 More about gauge theories

I've just described the mechanical construction we will follow, to buiild the path integral for the photon. But we can do better. Let's try to derive some more general formulas for gauge theories, and then really look closely at gauge invariance. I will restrict all my discussions to the case of ordinary electrodynamics, but the language will be general enough to consider more complicated situations (like non-Abelian gauge symmetry).

Let's begin by looking at constant functions in a gauge theory. They actually don't exist! The closest one can come to a constant function ϕ is to define a function whose covariant derivative is zero,

$$D_\mu \phi = (\partial_\mu + iA_\mu)\phi = 0 \tag{9.67}$$

(Note: I have set $q = 1$ or redefined $A_\mu = qA_\mu$.) Now

$$\phi(x + dx) = \phi(x) + dx^\mu \partial_\mu \phi \tag{9.68}$$

which, if $D_\mu \phi = 0$, is

$$\phi(x + dx) = \phi(x) - i dx^\mu A_\mu \phi(x), \tag{9.69}$$

and this is also

$$\phi(x + dx) = \exp(-i dx^\mu A_\mu(x))\phi(x) + O(dx^2). \tag{9.70}$$

Imagine performing a local gauge transformation $\phi(x) \to \exp(-i\Lambda(x))\phi(x) \equiv V(x)\phi(x)$. The exponential in Eq. 9.70 becomes

$$\begin{aligned} \exp(idx^\mu A_\mu(x)) &\to \exp(idx^\mu(A_\mu(x) + \partial_\mu\Lambda(x))) \\ &= \exp(-i(\Lambda(x+dx) + A_\mu(x)dx^\mu - \Lambda(x))) \\ &= V(x+dx)\exp(-idx^\mu A_\mu(x))V(x)^\dagger \end{aligned} \tag{9.71}$$

with a slight overkill of notation, $V^\dagger = V^*$ here – it's just a phase rotation. (The more complicated notation allows an immediate generalization to non-Abelian gauge theories like QCD or the Standard Model.) Note that $\exp(-idx^\mu A_\mu(x))$ transforms by rotations V and $V^\dagger$ "at its ends."

Now iterate Eq. 9.70 along any curve P from initial point x to final point y:

$$\phi(y) = \exp(-i\int_x^y dx' \cdot A(x'))\phi(x) \equiv P\exp(-i\int dx' \cdot A(x'))\phi(x). \tag{9.72}$$

As $\phi(x)$ is carried ("parallel transported," is the jargon borrowed from General Relativity) from point x to point y along a path P, it picks up a phase, the line integral of A along the path. Calling $\phi(y) = U(y,x)\phi(x)$ where

$$U(y,x) = P\exp(-i\int_x^y dx' \cdot A(x')), \tag{9.73}$$

we see that under a gauge transformation $\phi(y)' = V(y)\phi(y)$, or

$$\phi(y)' = V(y)U(y,x)V^\dagger(x)V(x)\phi(x) = U(y,x)'\phi(x)'. \tag{9.74}$$

$U(x,y)$ is often called a "Wilson line," named after Ken Wilson. Under a gauge transformation, it rotates with a phase factor at each end. Notice that this expression also implies that $U(x,x)$, a "Wilson loop," is gauge invariant. (For a non-Abelian gauge theory, $\phi(x)$ would be a column vector and U would become a matrix; the gauge invariant object is Tr $U(x,x)$.)

Consider next an infinitesimal path of length a in the $\hat{\mu}$ direction, connecting the point $x - a\hat{\mu}$ to x:

$$\exp(-iaA_\mu(x)') = V(x)\exp(-iaA_\mu(x))V^\dagger(x - a\hat{\mu}). \tag{9.75}$$

Expanding this,

$$\begin{aligned} 1 - iaA_\mu(x)' &= V(x)[1 - iaA_\mu(x)][V(x)^\dagger - a\partial_\mu V^\dagger] \\ &= 1 - iaV\partial_\mu V^\dagger - aVAV^\dagger + \ldots \end{aligned} \tag{9.76}$$

gives us the formula for the gauge transformation of A_μ,

$$A(x)'_\mu = V(x)A_\mu(x)V(x)^\dagger - iV(x)\partial_\mu V(x)^\dagger. \tag{9.77}$$

This slightly complicated language comes into its own when considering non-Abelian gauge theories (where $\phi(x)$ becomes a column vector and U becomes a matrix).

Now let's focus on the gauge transformation itself

$$V(x) = \exp i\omega(x). \tag{9.78}$$

The V's are the elements of an Abelian group G parametrized by the ω's. There is an identity element, $\omega = 0$ or $V = 1$, an inverse, $VV^\dagger = 1$, and a closure relation

$$V(x)'' = V(x)'V(x) \in G. \tag{9.79}$$

Mathematicians call this G the group $U(1)$ – the group of unitary transformations in one dimension. The group is "compact:" the size of all the elements $||V|| = 1$. Notice that the $U(y, x)$'s are elements of G, too:

$$U(x, x') = U(x, x'')U(x'', x) \in G. \tag{9.80}$$

Now, remember all our issues with $\int_{-\infty}^{\infty} dA$? If we had compact variables, we would not encounter such integrals. All this discussion suggests that we should define gauge theories *not* directly in terms of the A_μ's, but in terms of the U's. We do this in the following (roundabout) way:

First, replace the four dimensional space time continuum by a four dimensional lattice of points, with a lattice spacing a. (This probably only makes sense for an Euclidean path integral.) Matter fields (fermions and bosons) are replaced by fields defined on the sites of the lattice, $\psi(x) \to \psi(x_i)$. Gauge transformations are also defined only on the sites (as the notation $V(x_i)$ indicates).

Second, gauge information is transmitted by the $U(x, y)$'s. The minimal size U will span one lattice link connecting two adjacent sites x and $x + a\hat{\mu}$

$$U(x + a\hat{\mu}, x) \equiv U_\mu(x) \tag{9.81}$$

and it transforms as

$$U(x + a\hat{\mu}, x)' = V(x + a\hat{\mu})U(x + a\hat{\mu}, x)V(x)^\dagger \tag{9.82}$$

Notice $U(x, x + a\hat{\mu}) \equiv U_\mu(x)^\dagger$. Thus, the U_μ's live on the links of the lattice.

Third, we define the integration measure for the path integral as an integration over all the allowed values of the U's. This is called the "Haar measure" (or "invariant integration over the group elements") and is denoted as dU. It is invariant in the sense that if $U' = VUV^\dagger$, then $dU' = dU$. If G is compact, the integration is over a finite range.

Two examples: in our case, where $G = U(1)$, we can parametrize $U = \exp(i\theta)$ and

$$\int dU = \int_{-\pi}^{\pi} d\theta. \tag{9.83}$$

Geometrically, Haar measure is a line integral around a unit radius circle. Another simple group is $SU(2)$. Any element of $SU(2)$ can be parametrized in terms of the Pauli matrices (plus the identity) and four real numbers,

$$U = U_0 \mathbf{1} + i\vec{U} \cdot \vec{\sigma} \tag{9.84}$$

in terms of which

$$\int dU = \int \prod_{i=0}^{3} dU \delta(1 - \sum_i U_i^2). \tag{9.85}$$

Then the gauge part of the (Euclidean space) partition function is defined as

$$Z = \int \prod_{i,\mu} dU_\mu(i) \exp(-S(U)) \tag{9.86}$$

The Minkowski space path integral would have an i in the exponent.

Now, what could $S(U)$ be? It must be gauge invariant, and so the most general S will be a sum of closed paths of U's, multiplied by a set of arbitrary coefficients,

$$S(U) = \sum_x \sum_{path\ j} U_{P_j}(x, x) c_{P_j} \tag{9.87}$$

where $U_{P_j} = U(x, x_1)U(x_1, x_2) \ldots U(x_n, x)$.

The theory ought to have an interesting continuum limit, that is, if we take the lattice spacing a to zero, the action ought to reduce to something proportional to

$$S = \int d^4x F_{\mu\nu} F^{\mu\nu}. \tag{9.88}$$

Actually, that is easy to achieve. Any closed path will reduce to this form. Consider the smallest one, a path around a 1×1 plaquette,

$$U = \exp(-a[A_\mu(x) + A_\nu(x + a\hat{\mu}) - A_\mu(x + a\hat{\nu}) - A_\nu(x)]) \tag{9.89}$$

Write this as $\exp(ia\Delta)$ and Taylor expand Δ for small a:

$$\begin{aligned}\Delta &= A_\mu(x) + (A_\nu(x) + a\partial_\mu A_\nu(x)) - (A_\mu(x) + a\partial_\nu A_\mu(x)) - A_\nu(x) \\ &= a(\partial_\mu A_\nu - \partial_\nu A_\mu) \\ &= aF_{\mu\nu}.\end{aligned} \tag{9.90}$$

Then

$$U = i + ia^2 F_{\mu\nu} - \frac{a^4}{2} F_{\mu\nu}^2 + \ldots. \tag{9.91}$$

Clearly

$$\sum_j (U(j) + U^\dagger(j) - 2) = -(a^4 \sum_j) F_{\mu\nu}^2 = -\int d^4x F_{\mu\nu}^2. \tag{9.92}$$

We could just take the action to be the left hand side of this expression, a sum over all the "plaquettes" of U's. This looks like stupendous overkill, but notice that the compact measure precludes infinities a la

$$\int_{-\infty}^{\infty} dx \exp(-0 \times x^2). \tag{9.93}$$

We could use this formalism to study the nonperturbative behavior of gauge theories. In fact, that is what people like me do for a living. Instead, let's use it to study gauge theories in perturbation theory. That's easy here. The Haar measure for a $U(1)$ gauge theory is, if $U = \exp(i\theta)$,

$$\int dU = \int_{-\pi}^{\pi} d\theta \tag{9.94}$$

so if we take $U = \exp(iaA)$,

$$dU = \int_{-\pi}^{\pi} d(aA) \to \lim_{a\to 0} \int_{-\pi/a}^{\pi/a} dA_\mu \to \int_{\infty}^{\infty} dA_\mu \tag{9.95}$$

and we are back to integration over all the values of a classical field (just like we had for scalar fields).

Unfortunately, this is the integral which had the flat directions, so we seem to be back to where we started.

9.6 A well behaved path integral

The resolution to our difficulties is to eliminate the flat directions from the functional integral. This is done precisely as was discussed in Sec. 9.4, specifically in Eq. 9.66. Gauge fixing involves a constraint $g(A) = 0$ for some function g. Recall the gauge transformation formula Eq. 9.77,

$$A_\mu^V = VA_\mu(x)V^\dagger - iV\partial_\mu V^\dagger \tag{9.96}$$

and consider the quantity

$$\Delta_g^{-1}(A_\mu) = \int DV\delta[g(A_\mu^V)] \tag{9.97}$$

that is, integrate the gauge fixing $\delta-$function over all the gauge transformations V at every site. Then invoke the identity

$$1 = \Delta_g(A_\mu)\int DV\delta[g(A_\mu^V)]. \tag{9.98}$$

Our partition function is

$$Z(J) = \int DA_\mu\{\Delta_g(A_\mu)\int DV\delta[g(A_\mu^V)]\}\exp(-iS(J)). \tag{9.99}$$

The terms in the curly brackets are our "1". Of course $S(J) = S + \int A_\mu J^\mu$ for use in generating Green's functions.

Now notice that $\Delta_g(A_\mu)$ is, in fact, gauge invariant:

$$\Delta_g^{-1}(A_\mu^{V'}) = \int DV\delta[g(A_\mu^{V'V})], \tag{9.100}$$

but $DV = D(V'V)$ (Haar measure at work) so

$$\Delta_g^{-1}(A_\mu^{V'}) = \int DV''\delta[g(A_\mu^{V''})] = \Delta_g^{-1}(A_\mu). \tag{9.101}$$

This means that we can perform a gauge transformation in $Z(J)$

$$Z(J) = \int dV\int DA_\mu\Delta_g(A_\mu)\delta[g(A)]\exp(-iS(J)). \tag{9.102}$$

The $\int dV$ just factors out of the expression. We can drop it. What does the rest of the expression mean? Integrate over all the A_μ's, subject to the gauge fixing constraint $\delta(g)$, and with the Jacobian $\Delta_g(A_\mu)$.

All we have to do is groom this expression to make it useful. We start (improbably) by changing variables from V to g itself,

$$DV = Dg\det(\frac{\delta V}{\delta g}) \tag{9.103}$$

so that

$$\Delta_g^{-1}(A) = \int Dg\det(\frac{\delta V}{\delta g})\delta(g) = \det(\frac{\delta V}{\delta g})_{g=0} \tag{9.104}$$

or

$$\Delta_g(A) = \det(\frac{\delta g}{\delta V})_{g=0}. \tag{9.105}$$

If we parametrize $V = \exp(i\omega)$, this is

$$\Delta_g(A) = \det(\frac{\delta g}{\delta \omega})_{g=0}. \tag{9.106}$$

(Be patient, an example is coming.)

We are nearly done. It would be very useful to get everything into the exponential. Let's pick a gauge choice, with a free parameter α, which will do that for us:

$$\begin{aligned} \tilde{g}(A) &= \int Dc(x)\delta(g-c)\exp(-\frac{i}{2\alpha}\int d^4x(c(x)^2) \\ &= \exp(-\frac{i}{2\alpha}\int d^4x(g(A)^2). \end{aligned} \tag{9.107}$$

For example, $g(A) = \partial^\mu A_\mu$ corresponds to

$$\tilde{g}(A) = \exp(-\frac{i}{2\alpha}\int d^4x(\partial_\mu A^\mu)(\partial_\nu A^\nu)). \tag{9.108}$$

Finally, *if* (this is a very big IF) $\det(\frac{\delta g}{\delta \omega})$ does not depend on A, it also factors out of the functional integral and we can drop it. Again, an example: $g(A) = \partial^\mu A_\mu$. Its gauge transformation is

$$g(A^V(x)) = \partial_\mu(A^\mu + \partial^\mu\omega) \tag{9.109}$$

so

$$\frac{\delta g}{\delta \omega} = \delta^4(x-y)\Box \tag{9.110}$$

($\Box$ is the d'Alembertian, $\partial^\mu\partial_\mu$) What is this? Pick some boundary conditions, solve $\Box\phi = \lambda\phi$, and then

$$\det\frac{\delta g}{\delta \omega} = \prod_j \lambda_j. \tag{9.111}$$

Even more explicitly,

$$\det\frac{\delta g}{\delta\omega} = \det\sum_{n,m}\int d^4x d^4y\ \phi_n^\dagger(x)\delta^4(x-y)\Box\phi(y) \tag{9.112}$$

which is $\det\delta_{nm}\lambda_n$. End of the example. The point is, $\det\frac{\delta g}{\delta\omega}$ has no A dependence, so we can factor it out and forget about it.

Coulomb gauge is similar. The gauge fixing term is $\delta(\vec{\nabla}\cdot\vec{A})\det\nabla^2$.

This was a mess. I dragged you through it, because there are cases where the factorization does not occur, and the determinant remains as part of the functional integral. Most gauge choices for non-Abelian gauge theories pick up such extra terms. Even for QED, it is possible to invent nonlinear gauge choices. In the literature, the extra terms involve new fields, called "ghosts." We will not discuss them any more.

We are nearly done. Let's focus on the action including the gauge fixing term Eq. 9.108:

$$\begin{aligned} S &= \int d^4x[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\alpha}(\partial_\mu A_\mu)(\partial^\nu A^\nu)] \\ &= -\int d^4x[\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\nu) + \frac{1}{2\alpha}(\partial_\mu A_\mu)(\partial^\nu A^\nu)] \\ &= -\int d^4x[\frac{1}{2}\partial_\mu A_\nu\partial^\mu A^\nu - \frac{1}{2}\partial_\mu A_\nu\partial^\nu A^\mu + \frac{1}{2\alpha}(\partial_\mu A_\mu)(\partial^\nu A^\nu)] \\ &= -\frac{1}{2}\int d^4x A_\rho[-\partial^\mu\partial_\mu g^{\rho\nu} + (1-\frac{1}{\alpha})\partial^\rho\partial^\nu]A_\nu, \end{aligned} \tag{9.113}$$

integrating the derivative terms by parts. Now we can go into momentum space,

$$S = \frac{1}{2}\int d^4p A_\mu(-p)A_\nu(p)[g^{\mu\nu}p^2 - (1-\frac{1}{\alpha}p^\mu p^\nu] \tag{9.114}$$

and we see we just have a quadratic form. The propagator will be

$$-iD_{\mu\nu}(p) = i[g_{\mu\nu}X(p) + p_\mu p_\nu Y(p)], \tag{9.115}$$

where X and Y are scalar functions of p. Going through the index-pushing exercise (a homework problem) of inverting the differential operator,

$$g_{\mu\nu} = [p^2 g_{\mu\rho} - (1-\frac{1}{\alpha})p_\mu p_\rho][g_{\rho\nu}X(p) + p_\rho p_\nu Y(p)], \tag{9.116}$$

gives the photon propagator

$$-iD_{\mu\nu}(p) = \frac{i}{p^2}[g_{\mu\nu} - (1-\alpha)\frac{p_\mu p_\nu}{p^2}]. \tag{9.117}$$

Our issue with "flat directions" has been removed by choosing a gauge, and our gauge choice was covariant, so the propagator is also covariant.

The parameter α is a gauge choice. Invariant amplitudes correspond to physical processes and they will be gauge invariant. Thus the parameter α will not be present in the final answer. Sometimes, one carries α through the calculation, as a check against mistakes: if it survives till the end, you have a mistake. More often, one picks a "convenient" choice of gauge; "convenience" meaning that with that choice the calculation simplifies. Two such choices are

- Feynman gauge: $\alpha = 1$
$$-iD_{\mu\nu}(p) = \frac{ig_{\mu\nu}}{p^2}. \tag{9.118}$$
- Landau gauge: $\alpha = 0$
$$-iD_{\mu\nu}(p) = \frac{i}{p^2}[g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}]. \tag{9.119}$$

Both are obviously covariant.

9.7 Feynman rules for QED

Now we can look at some tree-level precesses in quantum electrodynamics. The Lagrangian is

$$\mathcal{L} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{9.120}$$

plus gauge fixing terms, and it is more useful to rewrite the expression as

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi + A^\mu \mathcal{D}^{-1}_{\mu\nu} A^\nu + e\bar{\psi}\gamma_\mu \psi A^\mu \tag{9.121}$$

since we can read off the Feynman rules for QED from this expression.

The vertex (see Fig. 9.1) couples an incoming fermion, an outgoing fermion, and a gauge field. It has a factor $-ie\gamma_\mu$. The photon propagator (in Feynman gauge) is

$$\mathcal{D}_{\mu\nu} = \frac{ig_{\mu\nu}}{q^2 + i\epsilon} \tag{9.122}$$

The fermion propagator is

$$S(p) = \frac{i(\not{p} + m))}{p^2 - m^2 + i\epsilon} \tag{9.123}$$

To compute a process to any order in e, we have to draw all topologically distinct diagrams. (This is just Wick's theorem at work.)

Then there are factors asociated with external particles, first, forfermions

- for a particle entering an initial state – $u(p)$
- for a particle leaving a final state – $\bar{u}(p)$
- for a antiparticle entering an initial state – $\bar{v}(p)$
- for a particle leaving a final state – $v(p)$

Incoming and outgoing photons carry a factor $\epsilon_{\mu\lambda}$ or $\epsilon^*_{\mu\lambda}$. This comes from

$$A_\mu = \int d^3p \ldots [\epsilon_{\mu\lambda} a_{p\lambda} e^{ipx} + \epsilon^*_{\mu\lambda} a^\dagger_{p\lambda} e^{-ipx}] \tag{9.124}$$

Incidentally, there are no extra "external field" factors for scalar particles.

See Fig. 9.1 for cartoon versions of these rules.

Fermion spin averages or sums go into the trace rules with

$$\begin{aligned}\sum_s u(p,s)_\alpha \bar{u}(p,s)_\beta &= (\not{p}+m)_{\alpha\beta} \\ \sum_s v(p,s)_\alpha \bar{v}(p,s)_\beta &= (\not{p}-m)_{\alpha\beta}.\end{aligned} \tag{9.125}$$

There are several possibilities for dealing with the photon polarization terms.

- For explicit polatization information, use physical polarizations: $\epsilon \cdot k = 0$ with a four-vector $\epsilon^\mu = (0, \vec{\epsilon})$
- OR if you want to sum over them, use the Feynman gauge trick

$$\sum_\lambda \epsilon_\mu(k,\lambda)\epsilon^*_\nu(k,\lambda) = -g_{\mu\nu} \tag{9.126}$$

  This is a nice choice because identities can reduce the products of γ's. For example

$$\begin{aligned}\not{\epsilon}\not{p}\not{q}\not{\epsilon} &= -\gamma_\mu \not{p}\not{q}\gamma_\mu = -4p\cdot q \\ \not{\epsilon}\not{p}\not{\epsilon} &= -\gamma_\mu \not{p}\gamma_\mu = 2\not{p}\end{aligned} \tag{9.127}$$

  You can't use $\epsilon \cdot k = 0$ at the same time!

Loop factors are identical to what we found with scalars: either include a factor

$$\int \frac{d^4q}{(2\pi)^4} \tag{9.128}$$

over all "free" momenta (conserving momenta at all vertices) or integrate

$$\int \frac{d^4q}{(2\pi)^4} \tag{9.129}$$

over all internal momenta and include a factor of $\delta^4(\sum p_{in} - \sum p_{out})$ at all vertices.

And there are two more rules: There is an extra factor of -1 for every closed fermion loop and a factor of -1 between diagrams differing by identical fermion exchange. These are both antisymmetrization factors. The figure shows graphs for electron-electron scattering illustrating this point. The initial and final wave functions are

$$\psi_{in} = \frac{12-21}{\sqrt{2}}; \qquad \psi_{out} = \frac{34-43}{\sqrt{2}} \tag{9.130}$$

so

$$\langle out|T|in\rangle = \frac{1}{2}[\langle 34|21\rangle + \langle 43|21\rangle - \langle 34|21\rangle - \langle 43|12\rangle] = \langle 34-43|12\rangle \tag{9.131}$$

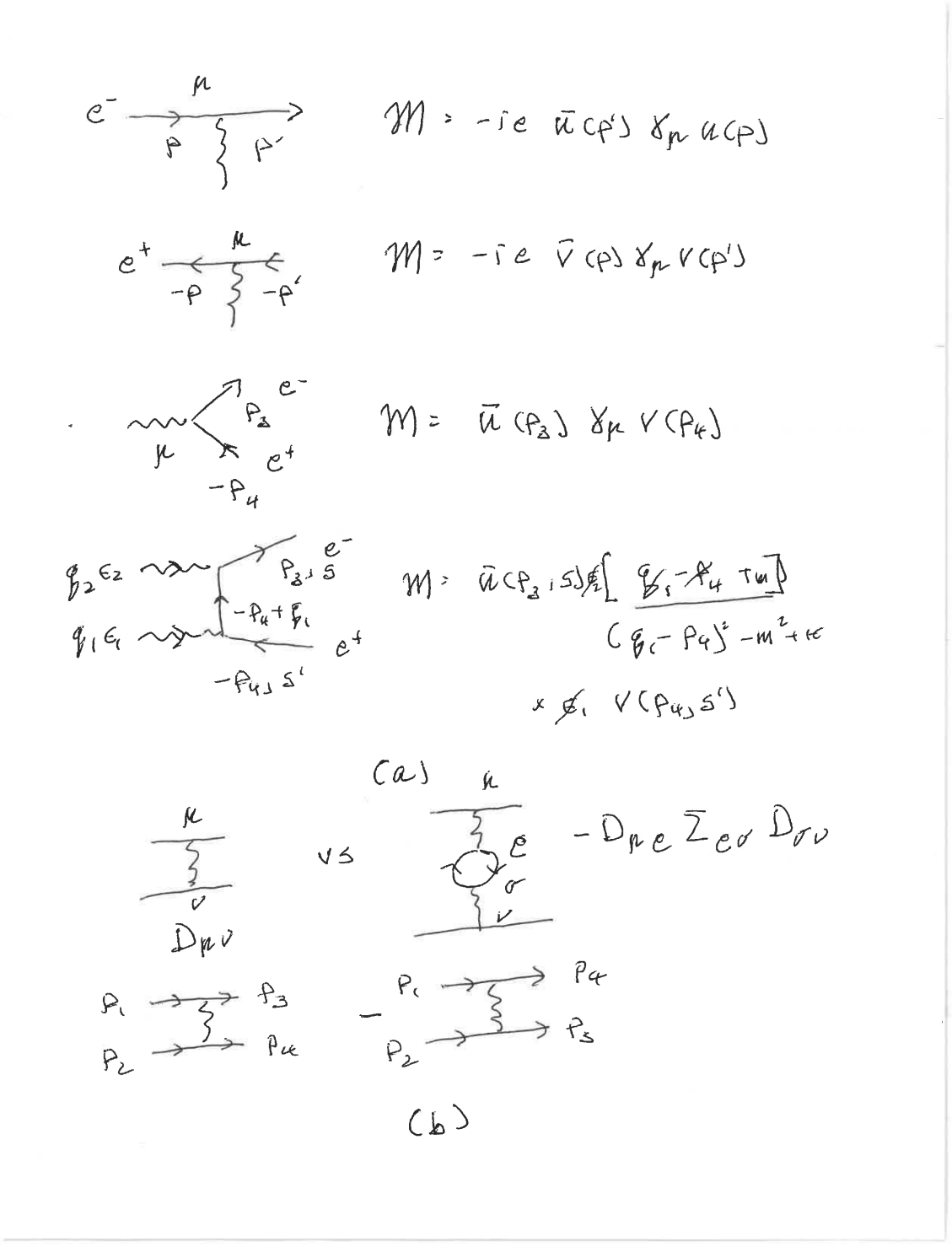

Figure 9.1: Figures illustrating Feynman rules (a) Sample amplitudes (b) Minus signs due to antisymmetry

9.8 Scattering processes in lowest order – e^+e^- annihilation

Our first process is electron-positron annihilation to a fermion-antifermon pair – $e^+e^- \to \mu^+\mu^-$ or $\tau^+\tau^-$ – or to quark-antiquark pairs. The Feynman graph is shown in Fig. 9.2a. I'll work in the center of mass and define the total energy as Q. I'll assume that $Q \gg m_e$, so that I can treat the electrons as massless. The amplitude is

$$\begin{aligned} M &= \bar{v}(p_2)(-ie\gamma_\mu)u(p_1)\left[\frac{ig^{\mu\nu}}{Q^2+i\epsilon}\right]\bar{u}(q_1)(-ie\gamma_\nu)v(q_2) \\ &= \frac{ie^2}{Q^2}[v(p_2)\gamma_\mu u(p_1)][\bar{u}(q_1)\gamma_\nu v(q_2)]. \end{aligned} \tag{9.132}$$

so that, averaging over initial spins and summing over final ones, the invariant amplitude is

$$\frac{1}{4}\sum_{ss'}|M|^2 = \frac{1}{4}\frac{e^4}{Q^4}\mathrm{Tr}\,\not{p}_2\gamma_\mu\not{p}_1\gamma_\nu\mathrm{Tr}\,(\not{q}_2 - m)\gamma_\mu(\not{q}_1 + m)\gamma_\nu. \tag{9.133}$$

There are two traces, one for each fermion's world line. The first trace, for the massless electrons, gives

$$4(p_{2\mu}p_{1\nu} + p_{2\nu}p_{1\mu} - g_{\mu\nu}p_1\cdot p_2) \tag{9.134}$$

The second trace is

$$4(q_{1\mu}q_{2\nu} + q_{2\mu}q_{1\nu} - q_{\mu\nu}q_1\cdot q_2) - 4m^2 g_{\mu\nu}. \tag{9.135}$$

With $e^2 = 4\pi\alpha$ the differential cross section is

$$\begin{aligned} \frac{d\sigma}{dt} &= \left(\frac{1}{16\pi Q^4}\right)(4\pi\alpha)^2\frac{16}{4}\times \\ &\quad [(p_2^\mu p_1^\nu + p_2^\nu p_1^\mu - g^{\mu\nu}p_1\cdot p_2)((q_{1\mu}q_{2\nu} + q_{2\mu}q_{1\nu} - q_{\mu\nu}q_1\cdot q_2) - m^2 g_{\mu\nu})]. \end{aligned} \tag{9.136}$$

Recalling that $g_{\mu\nu}g^{\mu\nu} = 4$, contracting the two expressions in the square brackets gives

$$2(p_1\cdot q_1)(p_2\cdot q_2) + 2(p_1\cdot q_2)(p_2\cdot q_1) + 2m^2 p_1\cdot p_2. \tag{9.137}$$

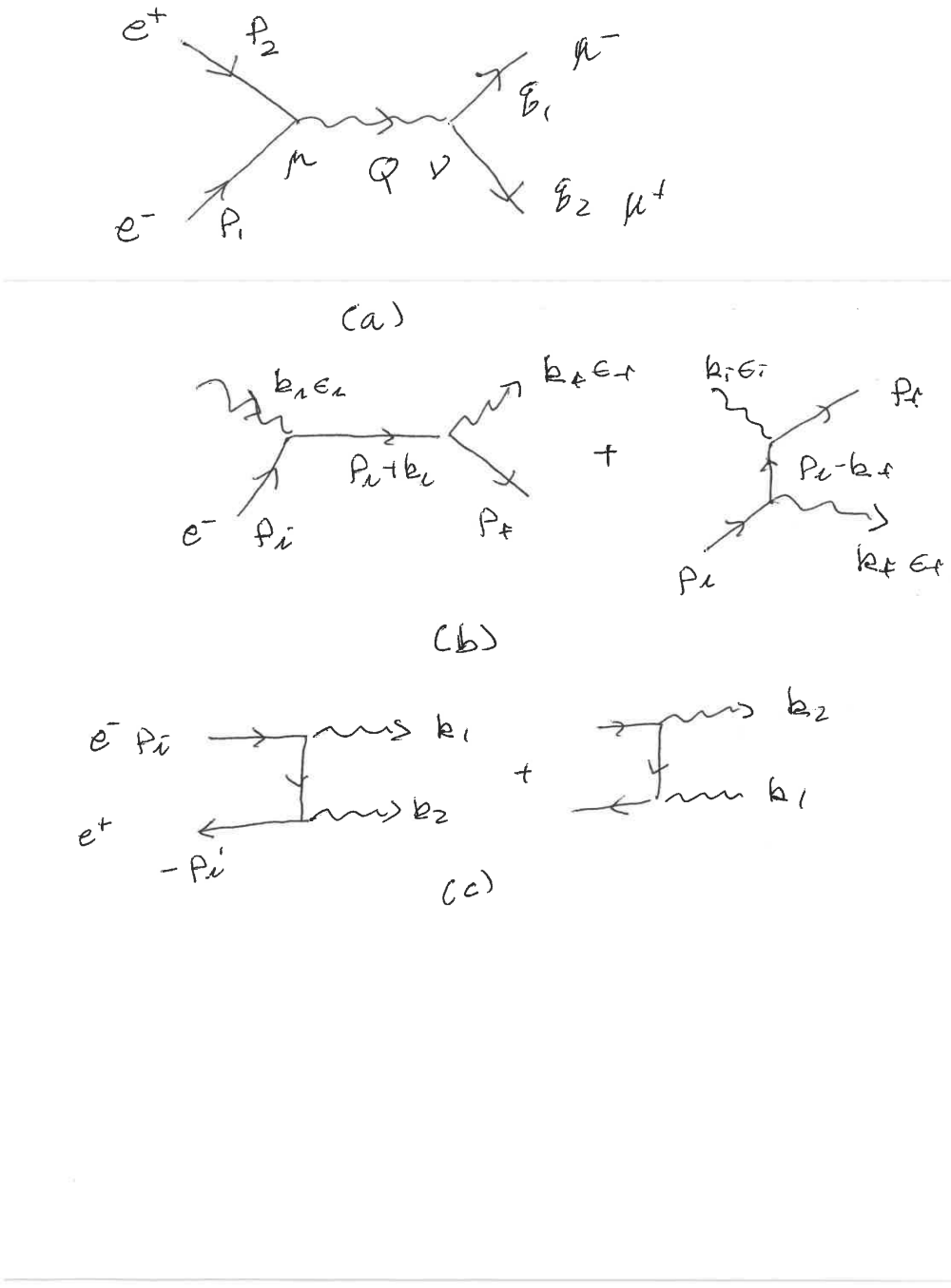

Figure 9.2: Tree-level scattering processes. (a) $e^+e^- \to \mu^+\mu^-$; (b) Compton scattering, $\gamma + e^- \to \gamma + e^-$; (c) Pair annihilation, $e^+e^- \to \gamma\gamma$;

Now we have to pick a coordinate system and write out the four vectors by components. In the center of mass frame (and with components (E, p_z, p_x, p_y)

$$\begin{aligned} p_1 &= \frac{Q}{2}(1,1,0,0) \\ p_2 &= \frac{Q}{2}(1,-1,0,0) \\ q_1 &= \frac{Q}{2}(1,v\cos\theta,v\sin\theta,0) \\ q_2 &= \frac{Q}{2}(1,-v\cos\theta,-v\sin\theta,0), \end{aligned} \tag{9.138}$$

so our kinematic expressions are

$$\begin{aligned} 2p_1\cdot p_2 &= Q^2 \\ p_1\cdot q_1 &= p_2\cdot q_2 = \frac{Q^2}{4}(1-v\cos\theta) \\ p_1\cdot q_2 &= p_2\cdot q_1 = \frac{Q^2}{4}(1+v\cos\theta) \\ t &= m^2 - 2p_1\cdot q_1 \\ dt &= \frac{Q^2}{2}v d\cos\theta \\ q_1^2 &= m^2 = \frac{Q^2}{4}(1-v^2) \\ v^2 &= 1-4\frac{m^2}{Q^2}. \end{aligned} \tag{9.139}$$

Using these relations, a line or two of algebra gives the (semi-)final result

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{2Q^2}v[1+v^2\cos^2\theta + 4\frac{m^2}{Q^2}]. \tag{9.140}$$

Maybe we can make the formula more memorable: consider the case $Q/m \gg 1$ where $v = 1$, and then

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{2Q^2}[1+\cos^2\theta] \tag{9.141}$$

or

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4Q^2}[1+\cos^2\theta]. \tag{9.142}$$

The total cross section is

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} = \frac{4\pi\alpha^2}{3Q^2}. \tag{9.143}$$

Note the units: α is dimensionless, a length is an inverse energy and at high energy the only relevant scale is the center of mass energy so $\sigma \propto 1/Q^2$.

How does the cross section behave at lower q? That is easy; omitting details, it is

$$\begin{aligned}\sigma &= \frac{\pi\alpha^2}{2Q^2} v[2(1+\frac{4m^2}{Q^2}) + \frac{2}{3}v^2] \\ &= \frac{4\pi\alpha^2}{3Q^2}\sqrt{1-\frac{4m^2}{Q^2}}(1+\frac{2m^2}{Q^2})\end{aligned} \tag{9.144}$$

and of course the threshold value of Q is $Q = 2m$ where the final state fermions are produced at rest.

Now for physics.

First, at threshold the cross section is proportional to v. This is a diagnostic for the production of a pair of spin-$\frac{1}{2}$ particles. An example of an application of this physics was the discovery of the tau lepton. It's useful to work in terms of the ratio

$$R_\tau = \frac{\sigma(\tau^+\tau^-)}{\sigma(\mu^+\mu^-)} = \sqrt{1-\frac{4m^2}{Q^2}}(1+\frac{2m^2}{Q^2}) \tag{9.145}$$

where $m = m_\tau$ and at $Q \sim 2m_\tau$, $m_\mu \simeq 0$. (The tau mass is $m_\tau = 1.77$ GeV, the muon mass is 105 MeV.) The shape of the turn-on of new physics was key to the discovery that the tau was a new spin-$\frac{1}{2}$ fermion. The situation was not so simple as this: the tau is unstable and decays weakly, into a tau neutrino plus an additional lepton pair (an electron and its antineutrino, or a muon and its antineutrino) or into hadrons. Neutrinos are effectively invisible in a small detector. Plus $2m_\tau$ is near the energies of bound states of a charm-anticharm quark pair. Everything was entangled and it was hard to tell what was going on. Take a look at the discovery paper: Phys. Rev. Lett. **35**, 1489-1492 (1975).

Next, the $(1+\cos^2\theta)$ angular distribution is diagnostic of the production of a spin-$\frac{1}{2}$ fermion-antifermion pair. This takes us to $e^+e^- \to$ hadrons. In QCD, the high energy expectation for the cross section is

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = 3\sum_i e_i^2 \tag{9.146}$$

where the "3" is a color factor (quarks come in three colors) and the sum on i is a sum over the quark flavors which are light enough to be produced: really light quarks are the

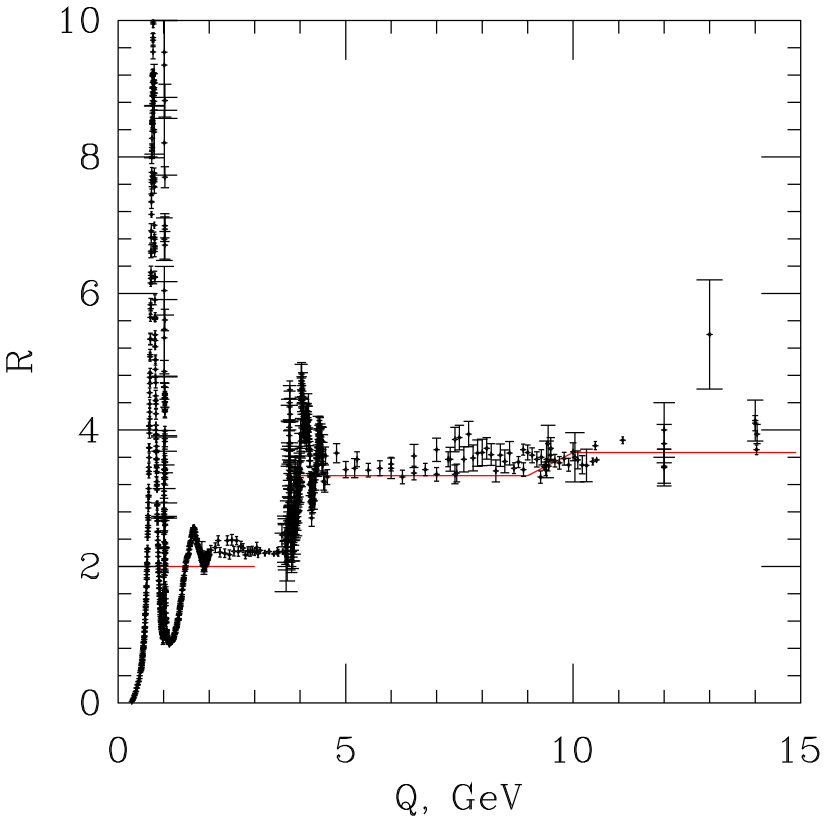

Figure 9.3: The R ratio of Eq. 9.146 with its expectation from charge counting overlaid in red.

up and down quarks with masses of about 4 and 7 MeV charge 2/3 and $-1/3$ (in units of the electron's charge). The strange quark at about 100 MeV and $q = -1/3$, and the heavy charm quark (at about 1200 MeV and charge 2/3), then the bottom quark at about 4.5 GeV ($e_b = -1/3$) and the top quark at about 175 GeV with charge 2/3.

Experimental data for Q up to 15 GeV is shown in Fig. 9.3. The red lines show the R ratio from Eq. 9.146 for up, down, and strange quarks from $0 < Q < 3$ GeV, then adding the charmed quark for $4 < Q < 9$ GeV, and finally with the bottom quark for $10 < Q < 15$ GeV. The bumps are at the masses of quark - antiquark resonances, where the simple calculation fails.

Hadronic jets were observed 50 years ago in e^+e^- annihilation into hadrons at what today would be considered an absurdly small center of mass energy (7 GeV) and the $(1 + \cos^2\theta)$ angular distributon was a diagnostic that pairs of fermions were being produced. See Phys.

Rev. Lett. **35**, 1609-1612 (1975).

9.9 Scattering processes in lowest order – Compton scattering and related

The process of Compton scatering is the reaction $\gamma + e^- \to \gamma + e^-$. The relevant graphs are shown in Fig. 9.2b. The amplitude has two terms:

$$\begin{aligned} M_{fi} &= (-ie)^2 \bar{u}(p_f) \not{\epsilon}_f \frac{i(\not{p}_i + \not{k}_i + m)}{(p_i + k_i)^2 + m^2} \not{\epsilon}_i u(p_i) \\ &\quad + (-ie)^2 \bar{u}(p_f) \not{\epsilon}_i \frac{i(\not{p}_i - \not{k}_f + m)}{(p_i - k_f)^2 + m^2} \not{\epsilon}_f u(p_i). \end{aligned} \tag{9.147}$$

One can take $\epsilon_i \cdot k_i = 0$ (this physical gauge choice would be useful if one wanted to calculate rates involving the initial or final photon polarizations). Alternatively, if one wanted to average or sum polarisations, the Feynman gauge trick

$$\sum_\lambda \epsilon_\mu(\lambda)\epsilon_\nu(\lambda) = -g_{\mu\nu} \tag{9.148}$$

is called for.

Pair annihilation into photons, $e^+e^- \to \gamma\gamma$ uses the same diagrams, only "twisted" as Fig. 9.2c shows. The translation dictionary

$$\begin{aligned} k_1 &\to k_f \\ k_1 &\to k_f \\ p_i &\to p_i \\ p_i' &\to -p_f \end{aligned} \tag{9.149}$$

converts the pair annihilation graphs to the Compton ones. This says that the amplitudes are related when written in terms of invariants:

$$\begin{aligned} &M_{Compton}(s = (p_i + k_i)^2, t = (k_i - k_f)^2, u = (k_i - p_f)^2) = \\ &\quad M_{pair}(t = (p_i - k1)^2, s = (k_1 + k_2)^2, u = (-k_2 - p_i')^2). \end{aligned} \tag{9.150}$$

Relations like these are called "crossing relations." It's a general result that the processes $AB \to CD$ and $A\bar{C} \to \bar{B}D$ (plus other permutations) are related.

9.10 Bremsstrahlung

Bremsstrahlung (Fig. 9.4) is the name for the emission of radiation when a charged particle is accelerated or decelerated (using classical language) by an external electromagnetic field. Let's study the process under the assumption that the emitted photon is "soft" –that it has low energy or (equivalently) long wavelength. The one can show that the general process factorizes as shown in Fig. 9.4. That is, almost all the radiation occurs off the initial state or final state charged particles and not in intemediate places in the graph. Physically, if the photon has momentum k and if the central blob has a spatial extent of $O(\lambda)$, then if $k\lambda \ll 1$ the photon cannot see it. In that case we can write

$$\begin{aligned} M(eA \to e\gamma B) &= ie\bar{u}(p_f)\not{\epsilon}\frac{\not{p}_f + \not{k} + m}{(p_f+k)^2 - m^2}\tilde{M}(p_i, p_f + k)u(p_i) \\ &\quad + ie\bar{u}(p_f)\tilde{M}(p_i - k, p_f)\frac{\not{p}_i - \not{k} + m}{(p_i-k)^2 - m^2}\not{\epsilon}u(p_i). \end{aligned} \tag{9.151}$$

The expression $\tilde{M}$ is the amplitude associated with the scattering without the extra photon emission. Now if $k \ll p_i$ and $k \ll p_f$ we can simplify the numerators and denominators,

$$M \sim ie\bar{u}(p_f)\left[\not{\epsilon}\frac{\not{p}_f + m}{2p_f \cdot k}\tilde{M}(p_i, p_f) + \tilde{M}(p_i, p_f)\frac{\not{p}_f + m}{-2p_i \cdot k}\not{\epsilon}\right]u(p_i). \tag{9.152}$$

We can further simplify this: first anticommute $\not{p}_i$ and $\not{\epsilon}$

$$(\not{p}_i + m)\not{\epsilon} = \not{\epsilon}(-\not{p}_i + m) + 2\epsilon \cdot p_i \tag{9.153}$$

and then use the Dirac equation,

$$(-\not{p}_i + m)u(p_i) = 0, \tag{9.154}$$

so the matrix element becomes

$$M \sim ie\left[\frac{\epsilon \cdot p_f}{k \cdot p_f} - \frac{\epsilon \cdot p_i}{k \cdot p_i}\right]\bar{u}(p_f)\tilde{M}(eA \to eB)u(p_i). \tag{9.155}$$

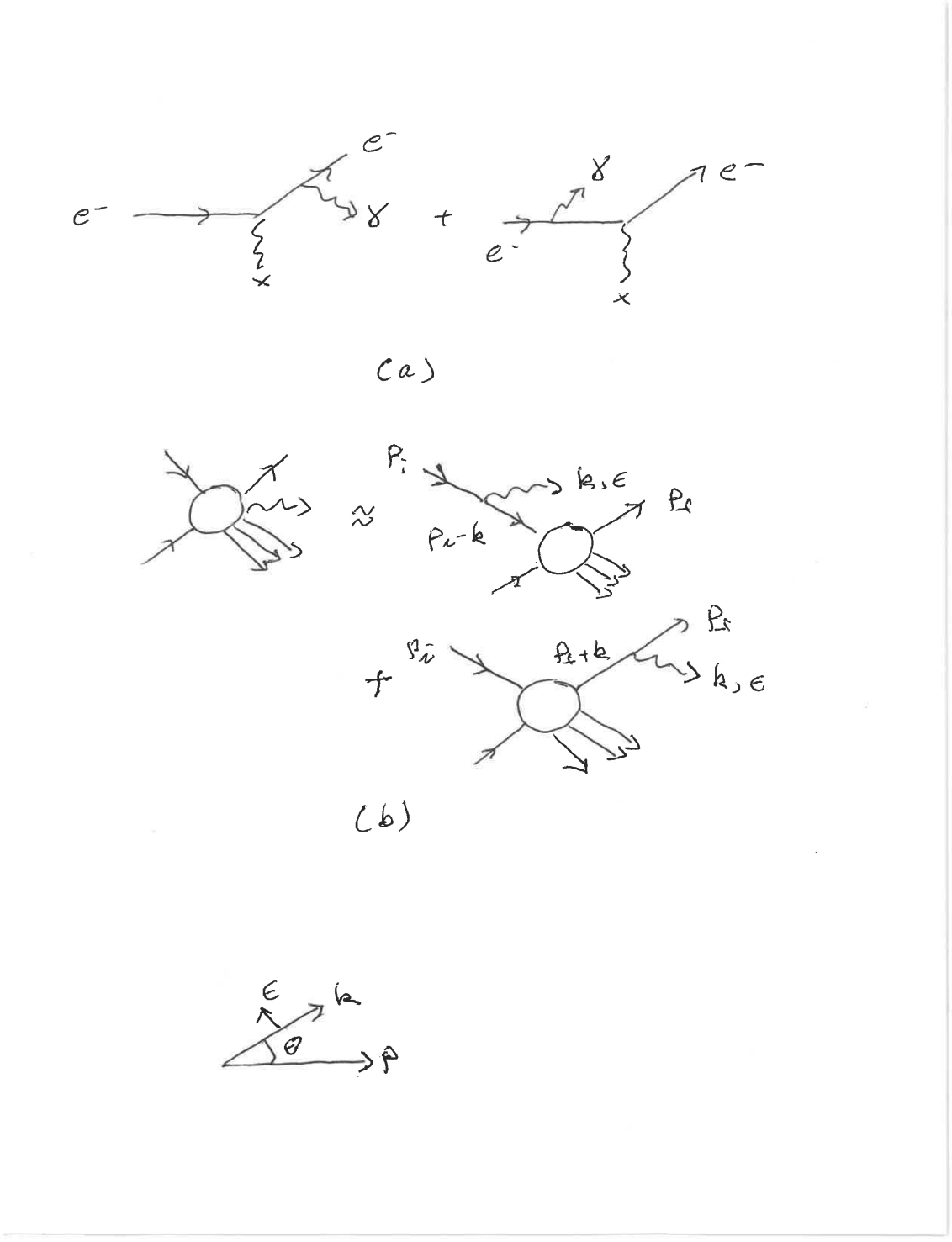

Figure 9.4: Figures illustrating Bremsstrahlung (a) Simple amplitudes (b) Kinematics for soft photon emission

We square M to find the cross section. Squaring $\tilde{M}$ gives the differential cross section for the reaction with no photon radiation, and all we have left is the photon's phase space, so

$$\begin{aligned}\frac{d\sigma(eA\to e\gamma B)}{d\Omega} &\times \frac{2\pi}{(2\pi)^3}\frac{k^2 dk d\cos\theta}{2k} \\ &= \frac{d\sigma(eA\to eB)}{d\Omega}\times e^2\{\left[\frac{\epsilon\cdot p_f}{k\cdot p_f}-\frac{\epsilon\cdot p_i}{k\cdot p_i}\right]^2\frac{2\pi}{(2\pi)^3}\frac{k^2 dk d\cos\theta}{2k}\}.\end{aligned} \tag{9.156}$$

The long expression in the curly brackets is the soft photon piece.

This is actually a classical formula (called the Weiszaker - Williams or equivalent photon approximation). When k is small and p is large, we can write

$$\frac{\epsilon\cdot p}{k\cdot p}\sim\frac{p\sin\theta}{kp(1-\cos\theta)}\sim\frac{\theta}{k\theta^2/2}\sim\frac{1}{k\theta} \tag{9.157}$$

so the cross section is

$$\begin{aligned}\sigma(\theta>\theta_0,k>k_0) &\sim \frac{\alpha}{2\pi}\sigma(\text{no rad})\int_{\mathrm{k_0}}^{\mathrm{E}}\frac{\mathrm{kdk}}{\mathrm{k^2}}\int_{\theta_0}\mathrm{d}\cos\theta\frac{\sin^2\theta}{(1-\cos\theta)^2} \\ &\sim \frac{\alpha}{2\pi}\sigma(\text{no rad})\int_{\mathrm{k_0}}^{\mathrm{E}}\frac{\mathrm{dk}}{\mathrm{k}}\int_{\theta_0}\frac{\theta^3\mathrm{d}\theta}{\theta^4} \\ &\sim \frac{\alpha}{2\pi}\sigma(\text{no rad})\left(\ln\frac{\mathrm{E}}{\mathrm{k_0}}\right)\left(\ln\frac{1}{\theta_0}\right).\end{aligned} \tag{9.158}$$

There is a lot of physics here!

- The ratio of the rate for photon emission is equal to the rate for no emission times α times (potentially large) logarithms (not just times α)
- The probability of radiating a soft photon is proportional to dk/k, that is, it is large, growing inversely with k, but
- The total energy carried off by these photons is finite

$$\langle k\sigma\rangle=\alpha\int_{k_0}^{E}[kdk\frac{1}{k^2}]k\sim\alpha E \tag{9.159}$$

There are an infinite number of soft photons being radiated, but the energy they carry off is finite and $O(\alpha)$. This is an example of what is called an "infrared divergence." It looks serious at first but it is not a problem as long as you ask the right question.

And I cannot resist – please re-read Jackson Chapter 15. For first year graduate students the chapter is usually a very cryptic one. The calculation we have just done was probably in the back of Jackson's mind when he wrote it.

The modern version of this topic is the production of jets in QCD: a produced quark radiates soft gluons (which then radiate softer quark - antiquark pairs) and so on.

Chapter 10

Renormalization and the renormalization group

The goal of the last part of the course is to give a somewhat informal description of renormalization. Our paradigm will be scalar field theory in D-dimensional Euclidean space, where $p^2 = \sum_{i=1}^{D} p_i^2$ and $G(p) = 1/(p^2 + m^2)$.

You recall the problem, from the study of the ϕ^4 field theory – higher order contributions to S-matrix elements, which we hoped would be small, had a strong dependence on processes involving large internal momenta. For example, the "tadpole graph" (Fig. 10.3) corresponds to an amplitude

$$\Sigma(p) = -\frac{\lambda_0}{2} \int \frac{d^4q}{(2\pi)^4} \frac{1}{q^2 + m_0^2}. \tag{10.1}$$

The integral receives a large contribution from values of q which are very large. In fact, as it stands, the integral does not make sense; it is divergent.

10.1 Regularization

The resolution of the problem involves several steps:

1. Regularization: do something to the theory to render all integrals finite. This involves introducing something like a cutoff in momentum, restricting $|q| < \Lambda$, or alternatively, introducing a short distance scale a in the problem and restricting lengths to be greater than a.

2. Then the question is, how does the cutoff Λ enter into physical processes. The answer has two parts, first a general classification scheme, then a careful investigation of where the Λ - dependent terms appear. At the end of the day, we will discover that two generic things can happen, either all the Λ dependence is hidden, or it is not. The two kinds of theories are called, respectively, "renormalizable" or "nonrenormalizable." Either way, the resulting theory can be used to make predictions for experiment.

Let's list a few kinds of regularization schemes.

The first one is a so-called hard momentum cutoff:

$$\int d^D q \to \int_{|q|^2 < \Lambda^2} d^D q. \tag{10.2}$$

A good feature of this choice is that it is simple (at least for scalar field theory). A disadvantage (if you are interested in studying a gauge theory) is that it violates gauge invariance.

Pauli-Villars regularization makes the propagator less singular by making the replacement

$$\frac{1}{q^2+m^2} \to \frac{1}{q^2+m^2} - \frac{1}{q^2+M^2} = \frac{M^2-m^2}{(q^2+m^2)(q^2+M^2)} \tag{10.3}$$

The Pauli-Villars regulator mass M is typically taken to be very large.

Dimensional regularization makes the replacement

$$\int d^4q \to \int d^Dq \tag{10.4}$$

where D is not an integer. The idea is to analytically continue D so that integrals are finite and then to express the answer for arbitrary D as a Laurant series and analytically continue back to near $D = 4$. As an illustration, the integral

$$\int d^Dq \frac{1}{q^2+m^2} \tag{10.5}$$

is convergent for $D < 2$. This is the "industry standard" of regularization schemes in the literature, but the procedure is not very intuitive. You can learn about it later if you want.

Finally, there is the lattice: define fields on the sites (or links, for gauge fields) of a lattice of lattice spacing a. Finite integrals result because the Brillouin zone is finite,

$$\int d^4q \to \prod_{i=1}^{4} \int_{-\pi/a}^{\pi/a} dq_i \tag{10.6}$$

An advantage of this scheme is that it can be made gauge invariant; a disadvantage is the propagators and vertices are very messy functions.

Let's stick with the hard momentum cutoff and classify the cutoff dependence of various diagrams. Generally, a graph will show power-law scaling in Λ, as Λ^d. (The case $d = 0$ corresponds to $\log\Lambda$ dependence.) As an example, the D-dimension tadpole graph is

$$I = \frac{\lambda}{2}\int \frac{d^Dq}{(2\pi)^D}\frac{1}{q^2+m^2} \sim \int^{\Lambda} \frac{q^{D-1}dq}{q^2} \sim \Lambda^{D-2} \tag{10.7}$$

so here $d = D-2$. The quantity d is called the "superficial degree of divergence" of a graph. It happens that there is a general expression for d:

Suppose we have a graph with V vertices, E external lines, and I internal lines. The superficial degree of divergence d comes from the independent loop integrals. The number of loops is $L = I - (V-1)$, as there are V $\delta-$ functions, but one of them gives an overall

$\delta(\sum k)$ for all the momenta entering and leaving the graph. As an example, the graph in Fig. 10.1a has $V = 2$, $I = 3$, and $E = 2$, from whence $L = 3-1 = 2$. Each loop is associated with a factor $\int d^D p \sim \Lambda^D$. A propagator, $1/(p^2 + m^2)$, contributes a Λ^{-2}. This says that the superficial degree of divergence is

$$d = LD - 2I = D(I - V + 1) - 2I. \tag{10.8}$$

Finally, if vertices have N legs (equivalently, there is a term in $\mathcal{L} \sim \phi^N$), we have

$$NV = E + 2I, \tag{10.9}$$

the "2" because each internal line connects two vertices. Again, another example in Fig. 10.1: $N = 4$, $V = 2$, $E = 4$, $I = 2$, $4 \times 2 = 4 + 2 \times 2$. Our final result uses

$$I = \frac{NV - E}{2} \tag{10.10}$$

so our desired general expression is

$$d = D + E(1 - \frac{D}{2}) + V\left((\frac{D}{2} - 1)N - D\right). \tag{10.11}$$

Some examples are in order. First $N = 4$, or $\mathcal{L} \sim \phi^4$, we find $d = D + E(1 - D/2) + V(D - 4)$. If $D = 4$, $d = 4 - E$ for any number of vertices. Only two processes show superficial cutoff dependence (see Fig. 10.1b), the two-point function, with $E = 2$ and $d = 2$, and the four point function, with $E = 4$ and $d = 0$.

Let me make a cryptic remark for future reference: the Lagrangian for this system has quadratic and quartic terms already in it,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}]\phi^4. \tag{10.12}$$

A two-point function, with $E = 2$ and $d = 2$, is generated by either of the two ϕ^2 terms in $\mathcal{L}$, while an $E = 4$ term comes from the ϕ^4 term in $\mathcal{L}$. This is an example of a situation with a finite number of $\Lambda-$ dependent processes, which correspond to terms already in $\mathcal{L}$. Such theories are called "renormalizable."

Note a caveat: higher n-point functions can be $\Lambda-$ dependent. An example is the ϕ^4 graph shown in Fig. 10.1c. It has $D = 2$. However, the Λ dependence is localized on the line, as it is for the two point function just described.

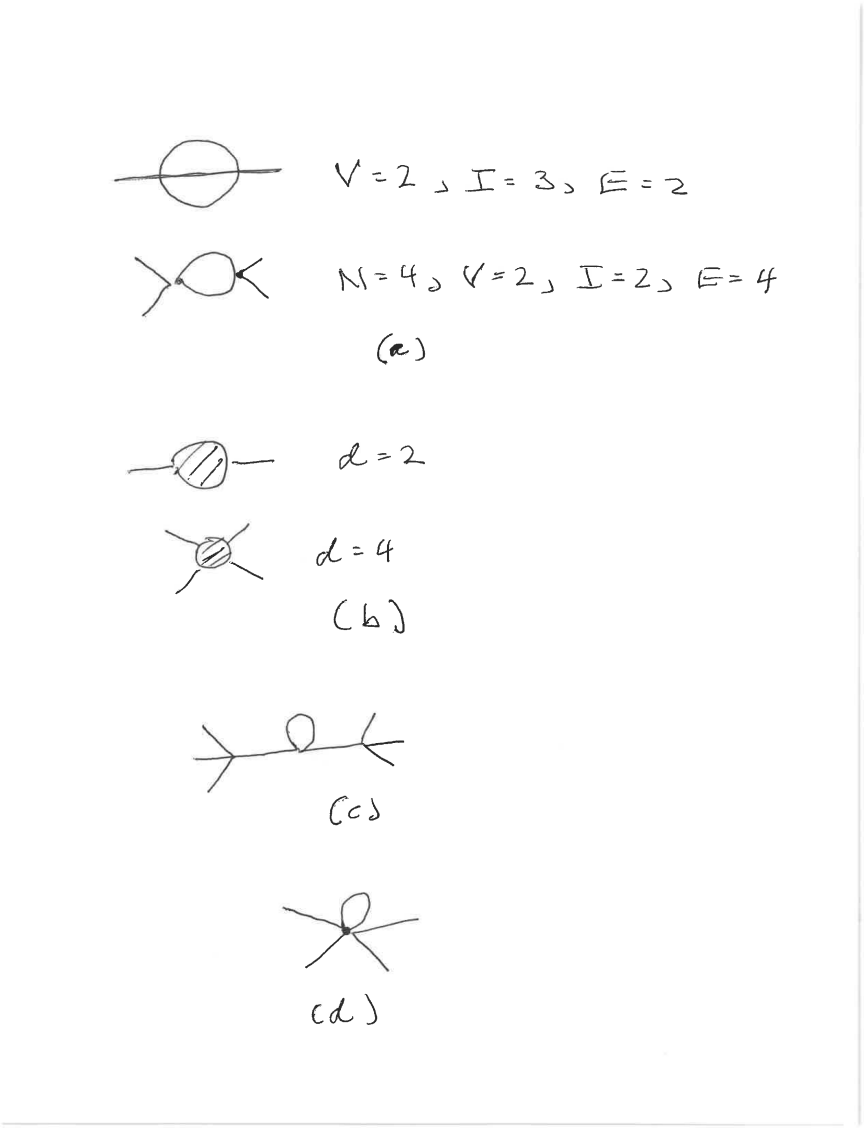

Figure 10.1: Graphs counting the degree of divergence. (a) examples. (b) The two professes in $D = 4$ ϕ^4 with $d \geq 0$. (c) A six-point functino with $d = 2$. (d) A divergent graph in $D = 2$.

A second example is $D = 4$, $N > 4$ (for example, from $\mathcal{L} \sim \phi^6$). Now

$$d = 4 - E + V(N - 4) \tag{10.13}$$

for V vertices. d grows with V at fixed E. This says that the Λ dependence increases order by order in the perturbative expansion. This is an example of what is called a nonrenormalizable interaction.

Finally, suppose $D = 2$. Then $d = 2 - 2V$, so only graphs with no vertices (?) or one vertex diverge. A $V = 1$ example is shown in Fig. 10.1d. It has an amplitude

$$\left(\int \frac{d^2p}{p^2}\right)^P \sim (\log \lambda)^P \tag{10.14}$$

for P closed loops. Two dimensions is obviously special.

Let's dig a little deeper and state two facts (which we have to show):

1. All coupling constants carry an engineering dimension of energy raised to a characteristic power $\sim \Lambda^p$)
2. Because of this fact, d is related to V.

10.2 Dimensional analysis and engineering dimensions

We'll begin by demonstrating the first item. Let's do a little dimensional analysis, looking at the Euclidean action (separating off the mass term for now)

$$\frac{S}{\hbar} = \int d^Dx[\frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\phi^2 + V(\phi)]. \tag{10.15}$$

I will use the same label as the momentum cutoff (Λ) as a generic marker for quantities with dimensions of energy. Equivalently, I will use the cutoff to set my energy scale. I will identify the dimensionality of an object by putting brackets around it.

The action S is dimensionless, so $\mathcal{L}$ scales like [length]$^{-D}$ or Λ^D. The kinetic and mass terms tell us that ϕ has an "engineering dimension" $[\phi] = L^{1-D/2}$ or $\Lambda^{D/2-1}$. We can find the dimensions of coupling constants, too: suppose we have a term $[\lambda_r\phi^r] = \Lambda^D$. Then

$$[\lambda_r] = \Lambda^{D-r(\frac{1}{2}D-1)} \equiv \Lambda^{\delta_r}. \tag{10.16}$$

For example,

$$[m^2] = \Lambda^2 \tag{10.17}$$

$$[\lambda_3] = \Lambda^{\frac{1}{2}(6-D)} \tag{10.18}$$

$$[\lambda_4] = \Lambda^{4-D} \tag{10.19}$$

$$[\lambda_6] = \Lambda^{6-2D} \tag{10.20}$$

and so on. This establishes fact number one, above. Note that λ_4 is dimensionless in $D = 4$.

Green's functions also have engineering dimensions:

$$[G^{(N)}(x_1, \ldots x_N)] = [\langle \phi(x_1)\phi(x_2)\ldots\rangle] \sim [\phi]^N = \Lambda^{N(D/2-1)}. \tag{10.21}$$

The Fourier transform has dimensions

$$[G^{(N)}(k_1, k_2 \ldots k_N)] = \Lambda^{N(D/2-1)}\Lambda^{-ND} = \Lambda^{-N(D/2+1)} \tag{10.22}$$

where the Λ^{-ND} comes from N factors of $d^D x$. However, recall that $G^{(N)}(k)$ has an overall momentum - conserving delta function. If we clip that off, $G = \bar{G} \times \delta^D(k)$, we have (recalling that $[\delta(k)] = \Lambda^{-D}$)

$$\begin{aligned}[\bar{G}^{(N)}(k_1, k_2 \ldots k_{N-1}; k_N = -\sum_i k_i)] &= \Lambda^{-N(D/2+1)}\Lambda^D \\ &= \Lambda^{D-N(D/2+1)}.\end{aligned} \tag{10.23}$$

Finally, there is a last piece of formalism. The vertex function $\Gamma^{(N)}$ is defined as the Green's function $G^{(N)}$, but with the external legs snipped off and with an extra conventional volume factor (equivalent to cutting off the $\delta^D(k)$ term) also removed:

$$\Gamma^{(N)}(x_1, \ldots x_N) = G^{(N)}(x_1 \ldots . x_N) V \prod_{i=1}^{N} G^{(2)}(x_i) \tag{10.24}$$

This is easiest to visualize in momentum space, as shown in Fig. 10.2, where

$$G^{(4)}(p_1 \ldots p_4) = (2\pi)^D \delta^D(\sum p_i) \prod_{i=1}^{4} \frac{1}{p_i^2 + m^2} \times \lambda \tag{10.25}$$

so

$$\bar{\Gamma}^{(4)}(p_1 \ldots p_4) = -\lambda + \ldots \tag{10.26}$$

(see the figure – the lowest order vertex function is just equal to $-\lambda$). And in this convention the two point function is

$$\Gamma^{(2)}(p, -p) = \frac{1}{G^{(2)}(p)} = p^2 + m^2 + \ldots \tag{10.27}$$

where the dots indicate potential higher order processes. After cutting off the $\delta^D(\sum p_i)$, we have

$$[\bar{\Gamma}^{(N)}] = \Lambda^{N+D-\frac{1}{2}ND}. \tag{10.28}$$

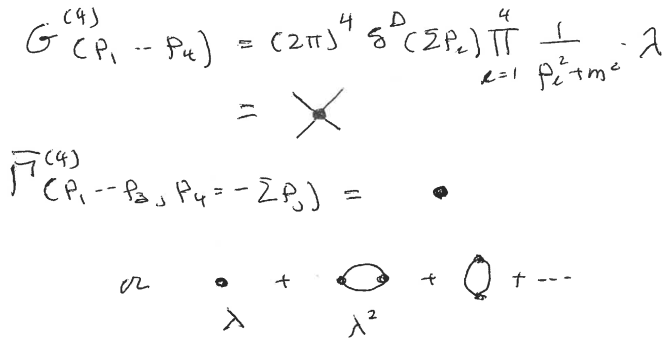

Figure 10.2: From $G^{(4)}$ to $\bar{\Gamma}^{(4)}$.

Note that $\Gamma^{(N)}$ has the same engineering dimension as λ_N, the N point coupling constant.

Now let's compute some one - loop amplitudes in the presence of a momentum cutoff Λ. Our goal is to identify the possible sensitivity of physical processes (scattering amplitudes) to physics at the cutoff scale.

10.3 Renormalization in scalar field theory

We have see that it is useful to consider how the bare parameters and cutoff combine to give renormalized parameters in a model renormalizable field theory. For a concrete example, consider (yet again) the Euclidean Lagrangian for a scalar field

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_0)^2 + \frac{1}{2} m_0 \phi_0^2 + \frac{\lambda_0}{4!}\phi_0^4 \tag{10.29}$$

and a momentum cutoff Λ. We assume perturbation theory is valid, and we want to consider various corrections to tree-level Green's functions. With one exception, we will restrict our calculation to the lowest nontrivial order in perturbation theory.

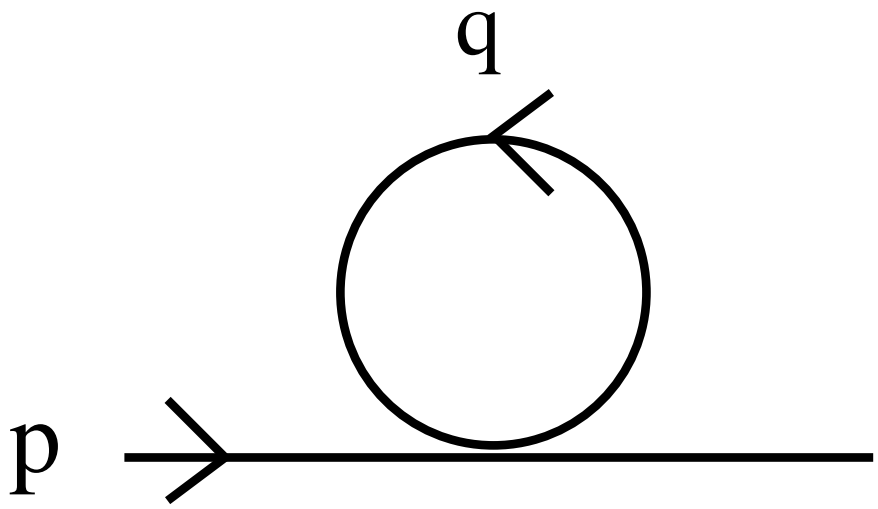

Figure 10.3: The self energy graph.

The simplest Green's function is the boson self energy, given by the Feynman graph of Fig. 10.3. It is

$$\Sigma(p) = -\frac{\lambda_0}{2}\int_0^\Lambda \frac{d^4q}{(2\pi)^4}\frac{1}{q^2+m_0^2}. \tag{10.30}$$

We write d^4q using four dimensinoal polar coordinates

$$q_i = q(\cos\theta_1, \sin\theta_1\cos\theta_2, \sin\theta_1\sin\theta_2\cos\phi, \sin\theta_1\sin\theta_2\sin\phi) \tag{10.31}$$

so that the volume integral is

$$\int d^4q = \int_0^\Lambda q^3 dq \int_0^\pi \sin^2\theta_1 d\theta_1 \int_0^\pi \sin\theta_2 d\theta_2 \int_0^{2\pi} d\phi. \tag{10.32}$$

Note also that

$$\int_0^\Lambda q^3 dq = \int_0^{\Lambda^2} \frac{1}{2} q^2 dq^2. \tag{10.33}$$

The integrand has no angular dependence, so (with $q^2 = x$)

$$\Sigma = -\frac{\lambda_0}{2} \frac{\pi^2}{(2\pi)^4} \int_0^{\Lambda^2} \frac{x}{x + m_0^2} dx, \tag{10.34}$$

which for $\Lambda^2 \gg m_0^2$ simplifies to

$$\Sigma \to -\frac{\lambda_0}{32\pi^2} \left[\Lambda^2 - m_0^2 \ln \frac{\Lambda^2}{m_0^2} \right]. \tag{10.35}$$

How does the self energy Σ contribute to the boson propagator? We can sum up its contribution to all orders in λ_0,

$$G = G_0 + G_0 \Sigma G_0 + G_0 \Sigma G_0 \Sigma G_0 + \cdots = G_0 \frac{1}{1 - \Sigma G_0}, \tag{10.36}$$

and if $G_0 = 1/(p^2 + m_0^2)$,

$$G = \frac{1}{p^2 + m_0^2 - \Sigma}, \tag{10.37}$$

i. e., we discover that the particle has an effective mass

$$m^2 = m_0^2 - \Sigma = m_0^2 + \frac{\lambda_0}{32\pi^2} \left[\Lambda^2 - m_0^2 \ln \frac{\Lambda^2}{m_0^2} \right]. \tag{10.38}$$

Next, let's look at the two-to-two scattering process shown in Fig. 10.4. We focus on the leftmost of the $M^{(2)}$ graphs, $M^{(2a)}$,

$$M^{(2a)} = \frac{\lambda_0^2}{2} \int_0^\Lambda \frac{d^4 l}{(2\pi)^4} \frac{1}{[l^2 + m_0^2][(p - l)^2 + m_0^2]}, \tag{10.39}$$

where $p = p_1 + p_2$. A useful way to do the integral is to use Feynman parameters. The trick is to write

$$\frac{1}{ab} = \int_0^1 dx \frac{1}{[ax + b(1 - x)]^2}. \tag{10.40}$$

Then

$$M^{(2a)} = \frac{\lambda_0^2}{2(2\pi)^4} \int d^4 l dx \frac{1}{[l^2 + m_0^2 - 2l \cdot px + p^2 x]^2}. \tag{10.41}$$

After a change of variables, $l' = l - px$, the denominator becomes $[l'^2 + m_0^2 + p^2 x(1 - x)]^2$, the angular dependence disappears from the integrand, the integral over l' can be done, and we have

$$M^{(2a)} = \frac{\lambda_0^2}{32\pi^2} \int_0^1 dx \left[\ln \left(1 + \frac{\Lambda^2}{m_0^2 + p^2 x(1 - x)} \right) - \frac{\Lambda^2}{\Lambda^2 + m_0^2 + p^2 x(1 - x)} \right]. \tag{10.42}$$

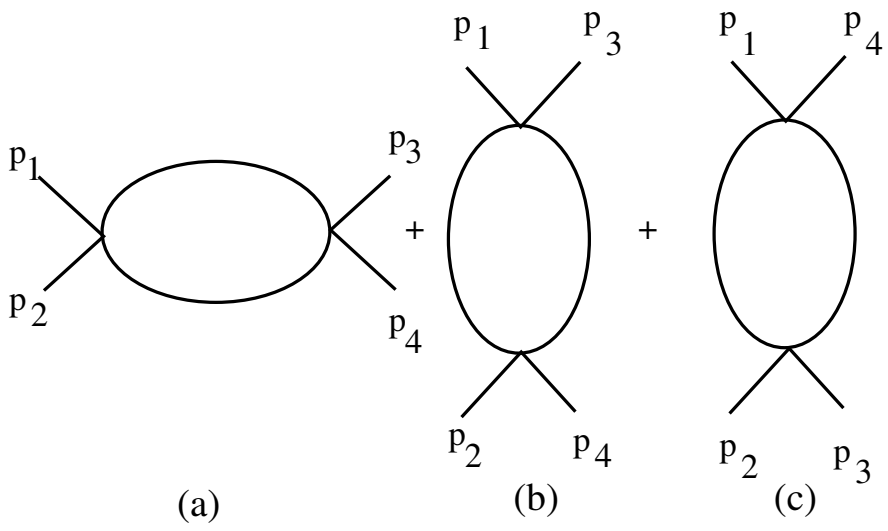

Figure 10.4: One-loop scattering amplitudes in ϕ^4 field theory.

If $\Lambda^2 \gg m_0^2$, we can drop the last term and the "1" in the first term. Let's also separate the p^2 dependence from the Λ dependence by introducing an arbitrary mass scale μ^2 (μ is called the "regularization point"). Then

$$\begin{aligned} M^{(2a)} &= \frac{\lambda_0^2}{32\pi^2}\int_0^1 dx\left[\ln\frac{\Lambda^2}{\mu^2} - \ln\frac{m_0^2 + p^2x(1-x)}{\mu^2}\right] \\ &= \frac{\lambda_0^2}{32\pi^2}\left[\ln\frac{\Lambda^2}{\mu^2} - \hat{I}(p^2)\right]. \end{aligned} \tag{10.43}$$

Note that $\hat{I}(p^2)$ is finite as $\Lambda \to \infty$, but it is also arbitrary because we must specify some value of μ. The other two diagrams are identical except for a relabeling of the external momenta. With the tree-level term included, the complete invariant amplitude up to second order in λ_0 is

$$M = -\lambda_0 + \frac{3\lambda_0^2}{32\pi^2}\ln\frac{\Lambda^2}{\mu^2} - \frac{\lambda_0^2}{32\pi^2}[\hat{I}(s) + \hat{I}(t) + \hat{I}(u)], \tag{10.44}$$

where $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$, and $u = (p_1 - p_4)^2$ are the Mandelstam invariants.

The momentum-independent terms in Eq. 10.44,

$$\lambda = \lambda_0 - \frac{3\lambda_0^2}{32\pi^2}\ln\frac{\Lambda^2}{\mu^2}, \tag{10.45}$$

always appear together in any set of diagrams. Indeed, if we computed to higher order, we would always find the same combination of bare coupling and cutoff. Note that the Λ

dependent terms only appear in combinations with λ_0 and m_0 to make up m and λ, and these terms appear in scattering amplitudes in a way that looks identical to terms in the original Lagrangian.

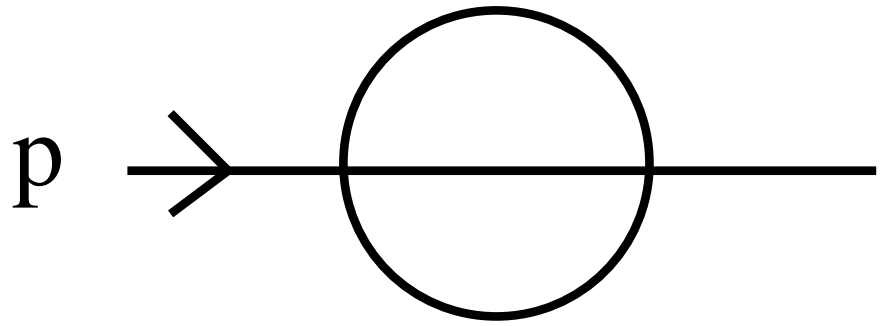

Figure 10.5: Lowest-order contribution to field renormalization.

Finally, there is one other kind of Λ dependence in Green's functions, a contribution to $\Sigma(p)$ that is proportional to p^2. It comes from the graph of Fig. 10.5, and it happens that it is equal to

$$\left[\frac{\lambda_0^2}{24(16\pi)^2}\ln\frac{\Lambda^2}{\mu^2}\right]p^2. \tag{10.46}$$

If we recalculate the ϕ_0 propagator while including it, we find

$$\langle\phi_0(-p)\phi_0(p)\rangle = \frac{1}{\left(1-\frac{\lambda_0^2}{24(16\pi)^2}\ln\frac{\Lambda^2}{\mu^2}\right)p^2 + m_0^2 + \frac{\lambda_0}{32\pi^2}(\Lambda^2 - m_0^2\ln\frac{\Lambda^2}{m_0^2})} \tag{10.47}$$

or

$$\langle\phi_0(-p)\phi_0(p)\rangle \equiv \frac{1}{Z_\phi}\frac{1}{p^2+m^2} \tag{10.48}$$

where $Z_\phi = 1 - \frac{\lambda_0^2}{24(16\pi)^2}\ln\Lambda^2/\mu^2$ is also cutoff-dependent. We can split Z_ϕ into two terms, one for each ϕ_0, and write

$$\phi = \frac{1}{\sqrt{Z_\phi}}\phi_0, \tag{10.49}$$

so

$$\langle\phi(-p)\phi(p)\rangle = \frac{1}{p^2+m^2}. \tag{10.50}$$

Thus, to this order in perturbation theory, all the cutoff dependence is confined to three places,

$$m^2 = m_0^2 - \Sigma = m_0^2 + \frac{\lambda_0}{32\pi^2}\left[\Lambda^2 - m_0^2\ln\frac{\Lambda^2}{m_0^2}\right], \tag{10.51}$$

$$\lambda = \lambda_0 - \frac{3\lambda_0^2}{32\pi^2} \ln \frac{\Lambda^2}{\mu^2}, \tag{10.52}$$

and

$$\phi = \left(\frac{1}{1 - \frac{1}{2} \frac{\lambda_0^2}{24(16\pi)^2} \ln \frac{\Lambda^2}{\mu^2}} \right) \phi_0. \tag{10.53}$$

In higher order, the cutoff dependence continues to appear only in m^2, λ, and ϕ – that is, physical scattering amplitudes only involve the quantities m^2, λ, and ϕ. Said differently, Green's functions parameterized in terms of m, λ and ϕ are finite as the cutoff is made very large. We could imagine determining m^2, λ, and ϕ from experiment (by fitting the parameters to experimental data at some value of the kinematic variables). The regularization point μ is fixed when we (arbitrarily) choose some prescription to define the scattering amplitude. For example, from Eq. (10.44) we could fix the value of λ knowing the value of M at some fiducial choice of the external momenta. If we follow this train of thought, we must regard the parameters m_0, λ_0, and ϕ_0 as implicit functions of the cutoff Λ; as Λ is varied, we must tune m_0, λ_0, and ϕ_0 so that m, λ and ϕ are unchanged.

Let's push on that point: but first, some vocabulary. One speaks of m_0, λ_0, and ϕ_0 (which are used to compute) as "bare quantities," and m^2, λ, and ϕ (which are measured in experiment) as "renormalized quantities." Then the first question is, how do the bare quantities depend on the cutoff Λ, assuming that the renormalized quantities are held fixed?

10.4 Towards the renormalization group

If we think of λ as being independent of Λ, then λ_0 must depend on Λ. How does it vary, to keep λ fixed? Start with

$$\lambda = \lambda_0 - \frac{3\lambda_0^2}{32\pi^2} \ln \frac{\Lambda^2}{\mu^2}, \tag{10.54}$$

and ask that

$$\Lambda \frac{\partial \lambda}{\partial \Lambda} = 0. \tag{10.55}$$

This is

$$\Lambda \frac{\partial \lambda}{\partial \Lambda} = \Lambda \frac{\partial \lambda_0}{\partial \Lambda} - \frac{6\lambda_0}{32\pi^2} \left(\ln \frac{\Lambda^2}{\mu^2} \right) \Lambda \frac{\partial \lambda_0}{\partial \Lambda} - \frac{6\lambda_0^2}{32\pi^2} = 0 \tag{10.56}$$

or

$$\Lambda \frac{\partial \lambda_0}{\partial \Lambda} [1 - \frac{6\lambda_0}{32\pi^2} \left(\ln \frac{\Lambda^2}{\mu^2} \right)] = \frac{6\lambda_0^2}{32\pi^2}. \tag{10.57}$$

If λ_0 is small, keep only the lowest order;

$$\Lambda \frac{\partial \lambda_0}{\partial \Lambda} = \frac{3\lambda_0}{16\pi^2}. \tag{10.58}$$

As Λ increases, so must the bare coupling λ_0, if λ is to remain fixed.

In fact, we can integrate this equation, to go beyond perturbation theory and find $\lambda_0(\Lambda)$:

$$\int_{\lambda_1}^{\lambda_2} \frac{d\lambda_0}{\lambda_0^2} = b \int_{\Lambda_1}^{\Lambda_2} \frac{d\Lambda}{\Lambda} \tag{10.59}$$

(where $b = 3/(16\pi^2)$) and doing the integral, we find

$$\frac{1}{\lambda_1} - \frac{1}{\lambda_2} = b \ln \frac{\Lambda_2}{\Lambda_1} \tag{10.60}$$

or

$$\lambda_2 = \frac{\lambda_1}{1 - \lambda_1 b \ln \frac{\Lambda_2}{\Lambda_1}}. \tag{10.61}$$

Again, this says that if $\Lambda_2 > \Lambda_1$ then $\lambda_2 > \lambda_1$.

So much for the theory in terms of bare quantities. What about the theory from the point of view of the renormalized ones?

Bare and renormalized Green's functions are proportional to each other because of the relation between ϕ_0 and ϕ:

$$\bar{\Gamma}_R^{(N)}(p_1 \dots p_N; \lambda, m, \mu) = Z_\phi^{N/2} \bar{\Gamma}_0^{(n)}(p_1 \dots p_N; \lambda_0, m_0, \Lambda). \tag{10.62}$$

The bare theory has cutoff dependence, so there is a Λ in the right hand expression. The renormalized theory depends on the renormalized coupling, the renormalized mass, and the renormalization point μ. We assume that $\bar{\Gamma}_R$ is finite as $\Lambda \to \infty$. Let's evaluate the ratio of $\bar{\Gamma}_R$'s at two different values of the renormalization point μ: we can do that because, from the point of view of the bare theory, μ is arbitrary. We just take a ratio,

$$\bar{\Gamma}_R^{(N)}(p_1 \dots p_N; \lambda_1, m_1, \mu_1) = \left[\frac{Z_\phi(\lambda_1, m_1, \mu_1)}{Z_\phi(\lambda_2, m_2, \mu_2)} \right]^{n/2} \bar{\Gamma}_R^{(N)}(p_1 \dots p_N; \lambda_2, m_2, \mu_2). \tag{10.63}$$

Because the $\bar{\Gamma}_R$'s are finite, the ratio $Z = Z_\phi(\lambda_1, m_1, \mu_1)/Z_\phi(\lambda_2, m_2, \mu_2)$ is finite as $\Lambda \to \infty$. (To check this point, it is

$$\frac{Z_\phi(\lambda_1, m_1, \mu_1)}{Z_\phi(\lambda_2, m_2, \mu_2)} = 1 - \frac{\lambda^2}{24(16\pi)^2} \ln \frac{\mu_1^2}{\mu_2^2} \tag{10.64}$$

to leading order in perturbation theory). Thus a change in the momentum scale μ at which the theory is normalized is equivalent to

1. A finite rescaling of the fields

2. A transformed coupling $\lambda_1 \to \lambda_2$

Let's see what that change is, by implementing the obvious fact that the bare theory does not know about the renormalization point: this means that

$$\mu \frac{d\bar{\Gamma}_0}{d\mu} = 0 \tag{10.65}$$

or

$$\left[\mu\frac{\partial}{\partial\mu} + \mu\frac{\partial\lambda}{\partial\mu}\frac{\partial}{\partial\lambda} + \frac{\mu}{2}\left(\frac{\partial}{\partial\mu}\ln m^2\right) m\frac{\partial}{\partial m} - \frac{N}{2}\frac{\partial \ln Z_\phi}{\partial\mu}\right]\bar{\Gamma}_R^{(n)} = 0. \tag{10.66}$$

It's conventional to define the quantities

$$\beta(\lambda) = \mu\frac{\partial\lambda}{\partial\mu} \tag{10.67}$$

$$\gamma_\phi = \frac{\mu}{2}\frac{\partial \ln Z_\phi}{\partial\mu} \tag{10.68}$$

$$\gamma_m(\lambda) = \frac{\mu}{2}\frac{\partial \ln m^2}{\partial\mu}. \tag{10.69}$$

Since β is dimensionless, it must depend only on the ratio m/μ.

This is not yet a useful equation. We don't care so much how $\bar{\Gamma}_R$ depends on μ as we do on how $\bar{\Gamma}_R$ depends on the momenta of the scattering particles at fixed μ. We can convert Eq. 10.66 into a useful one if we recall that $\bar{\Gamma}_R$ has an engineering dimension $[\text{mass}]^{d_n}$. Scale all the momenta $p \to sp$ and count powers of μ, s, and m – they must add up to d_n:

$$[\mu\frac{\partial}{\partial\mu} + s\frac{\partial}{\partial s} + m\frac{\partial}{\partial m} - d_n]\bar{\Gamma}_R^{(n)}(sp, m, \lambda, \mu) = 0. \tag{10.70}$$

Combining Eqs. 10.66 and 10.70 and solving for $\mu\partial\bar{\Gamma}_R/\partial\mu$ gives

$$\left[-s\frac{\partial}{\partial s} + \beta(\lambda)\frac{\partial}{\partial\lambda} + (\gamma_m - 1)m\frac{\partial}{\partial m} - n\gamma_\phi + d_n\right]\bar{\Gamma}_R^{(n)}(sp_1, sp_2 \ldots; m, \lambda, \mu) = 0. \tag{10.71}$$

This equation (called the Callan - Symanzik equation) gives the variation in Green's functions as we vary the external momenta. If we could solve it, we would know $\bar{\Gamma}(sp)$ in terns of $\bar{\Gamma}(p)$.

Let's write down the formal solution. We do this by first defining a "running" or "scale dependent" coupling constant and mass,

$$\frac{d\bar{\lambda}(s)}{ds} = \beta(\bar{\lambda}(s)); \qquad \bar{\lambda}(1) = \lambda \tag{10.72}$$

and

$$s\frac{d\bar{m}(s)}{ds} = [\gamma_m(\bar{\lambda}(s)) - 1]; \qquad \bar{m}(1) = m. \tag{10.73}$$

Then

$$\bar{\Gamma}^{(n)}(sp; m, \lambda, \mu) = s^{d_n}\Gamma^{(n)}(p; \bar{m}(s), \bar{\lambda}(s), \mu) \exp\left(-n\int_1^s \frac{ds'}{s'}\gamma_\phi(\bar{\lambda}(s'))\right). \tag{10.74}$$

To check the answer, just differentiate it. Here $d_n = n+D-\frac{1}{2}nD$ is the engineering dimension of $\bar{\Gamma}^{(n)}$. The formula says that a change in the momentum scale $p \to sp$ is equivalent of a change in the amplitude of

1. multiplication (scaling) by s^{d_n}. This would be the scaling law for $\bar{\Gamma}$ in the absence of interactions
2. An extra factor (the last integral) which gives rise to a so-called "anomalous dimension:" a change in the scaling of $\bar{\Gamma}$ as a result of interactions
3. A modified or "running" coupling constant and a "running" mass, whose values depend on the momentum scale.

The third item is the most important one, especially in describing the qualitative behavior of field theories. To introduce the physics, let's temporarily focus on a situation where we have one coupling $\lambda(s)$ with its own beta function $\beta(\lambda)$, and where the running mass $\bar{m}(s)$ is always zero. Let's also suppose that there is a special value of $\lambda(s)$, called λ^*, where the beta function is zero. The situation $\beta(\lambda^*) = 0$ is special; it is called a "fixed point." Let us defer a discussion of the theory at a fixed point, concentrating first on the physics of a running coupling. Let's imagine a simple form for β – a linear zero

$$s\frac{d\lambda}{ds} = \beta(\lambda) = (\lambda - \lambda^*)\beta'(\lambda^*) \tag{10.75}$$

If $\beta'(\lambda^*) < 0$ then for $\lambda > \lambda^*$, $sd\lambda/ds < 0$ and for $\lambda < \lambda^*$ $sd\lambda/ds > 0$. Under a change of scale λ flows into λ^* as s increases, and then it does not move ($\beta(\lambda^*) = 0$). This situation is called an "ultraviolet fixed point" or UV fixed point. Physics at large scale value s or large p is governed by λ^*. Conversely, if $ds < 0$ (we are going from short distance to long distance) we flow away from the fixed point. (See Fig. 10.6.) This situation can be called a UV attractive, UV stable, IR (infared) repulsive or IR unstable fixed point.

In particular, note that long distance physics will be very different for $\lambda > \lambda^*$ or $\lambda < \lambda^*$ (long distance compared to the distance scale where the coupling is equal to λ). In statistical

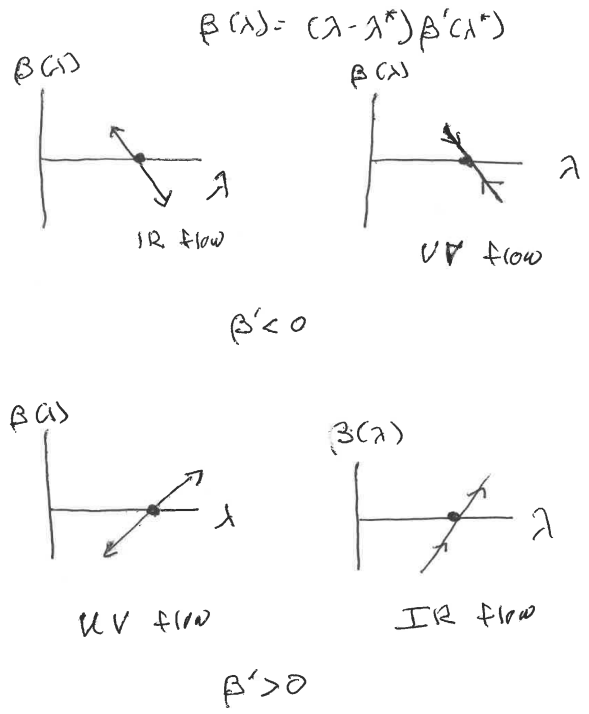

Figure 10.6: Examples of renormalization group flow around a linear fixed point $\beta(\lambda) = (\lambda - \lambda^*)\beta')\lambda^*)$. (a) $\beta')\lambda^*) < 0$; (b) $\beta')\lambda^*) > 0$.

mechanics, IR unstable fixed points are interesting because they characterize the locations of second (or higher) order phase transitions. Examples include order-disorder transitions in ferromagnets, where one long distance phase corresponds to the case where λ runs to a small value if it is less than λ^* or to large λ if $\lambda > \lambda^*$. Critical exponents are related to the anomalous dimensions γ_ϕ evaluated at λ^*.

If $\beta'(\lambda^*) > 0$ we have an IR attractive or IR stable or UV repulsive or UV unstable fixed

point. These fixed points characterize the long distance physics for our system.

We can check these statements by integrating Eq. 10.75. Suppose that at scale factor s_1 the coupling is $\lambda(s_1)$ and at s_2 we have $\lambda(s_2)$. Just doing the integral, we find

$$\frac{\lambda(s_1) - \lambda^*}{\lambda(s_2) - \lambda^*} = \left(\frac{s_1}{s_2}\right)^b, \tag{10.76}$$

which shows the evolution of the coupling toward or away from the fixed point according to the sign of b.

Now for $\lambda\phi^4$: what is the beta function? We can compute it starting with the bare charge.

$$\beta(\lambda) = \mu\frac{d\lambda}{d\mu} \tag{10.77}$$

and

$$\lambda = \lambda_0 - \frac{3\lambda_0^2}{32\pi^2}\ln\frac{\Lambda^2}{\mu^2}, \tag{10.78}$$

or

$$\lambda_0 = \lambda + -\frac{3\lambda^2}{32\pi^2}\ln\frac{\Lambda^2}{\mu^2} + \ldots \tag{10.79}$$

inverting Eq. 10.78. Then

$$\begin{aligned} \mu\frac{d\lambda_0}{d\mu} &= 0 \\ &= \mu\frac{d\lambda}{d\mu} - \frac{6\lambda}{32\pi^2}\left(\ln\frac{\Lambda^2}{\mu^2}\right)\mu\frac{d\lambda}{d\mu} - \frac{6\lambda^2}{32\pi^2} \end{aligned} \tag{10.80}$$

or

$$\mu\frac{d\lambda}{d\mu}[1 - \frac{3\lambda}{16\pi^2}\left(\ln\frac{\Lambda^2}{\mu^2}\right)] = \frac{3\lambda^2}{16\pi^2}. \tag{10.81}$$

To lowest order,

$$\beta(\lambda) = \frac{3\lambda^2}{16\pi^2} + O(\lambda^4) \tag{10.82}$$

or

$$s\frac{d\lambda}{ds} = \beta(\lambda) = \frac{3\lambda^2}{16\pi^2}. \tag{10.83}$$

This says that as the momentum scale grows, so does the effective or running coupling. Alternatively, as s shrinks, $\lambda(s)$ falls. This says that the effective coupling becomes smaller and smaller at longer and longer distance. This behavior is common to all field theories except for non-Abelian gauge theories (like QCD).

In QCD, $\beta(\lambda) < 0$; in fact, $\beta(\lambda) \sim -b\lambda^2$. The point $\lambda = 0$ is a UV attractive fixed point This means that in theories like QCD the constituents become increasingly weakly interacting at large momentum scales. We can integrate this beta function and see this result:

$$s\frac{d\lambda}{ds} = -b\lambda^2 \tag{10.84}$$

$$-\int_{\lambda_1}^{\lambda_2} \frac{d\lambda}{\lambda^2} = b \ln \frac{s_2}{s_1} \tag{10.85}$$

or

$$\frac{1}{\lambda_2} - \frac{1}{\lambda_1} = b \ln \frac{s_2}{s_1} \tag{10.86}$$

$$\frac{1}{\lambda_2} = \frac{1}{\lambda_1} + b \ln \frac{s_2}{s_1} \equiv b \ln \frac{s_2}{\bar{s}} \tag{10.87}$$

and finally

$$\lambda(s) = \frac{1}{b \ln \frac{s}{\bar{s}}} \tag{10.88}$$

absorbing $1/\lambda_1$ and s_1 into a common factor $\bar{s}$. The coupling decreases as s rises. This behavior is called "asymptotic freedom."

Samll λ at large s means that processes at large s – large momentum scales – can be calculated in perturbation theory. Indeed, in QCD predictions such as

$$R = \frac{\sigma(e^+e^- \to hadrons)}{\sigma(e^+e^- \to \mu^+\mu^-)} = 3\sum_i e_i^2, \tag{10.89}$$

computed at large center of mass energies, are very well satisfied.

Conversely, as $s \to 0$, λ rises and perturbation theory is no longer valid. Of course, we know that the interactions of quarks at very long distance is very complicated. Another active area of research is the calculation of predictions of QCD at low energy or long distance.

The behavior of ϕ^4 theory (and the same behavior also arises in QED) is called "triviality." At long distance the effective λ goes to zero and the theory exhibits nearly free field behavior. This has several interesting consequences. Here is one.

Suppose we imagine introducing a UV cutoff Λ, and we define λ_0 at the cutoff scale. The coupling $\lambda(Q)$ at scales $Q < \Lambda$ is smaller than λ_0. Now take Λ to be very large. If λ_0 is not also taken to be very large, then $\lambda(Q)$ at low Q will fall to zero. Conversely, suppose we measure a nonzero λ at some IR scale s. This implies that there is another (higher) scale s_{UV} where λ has grown to be order unity. At that scale our simple perturbative picture

breaks down and must be replaced. Some new physics must appear. We speak of the theory "defining its own region of validity."

Now let us return to look at physics close to a fixed point. To do this we have to generalize a bit what we mean by a "coupling" – in addition to coefficients of ϕ^n, I will include the mass (or squared mass) as another coupling. All these couplings flow with scale change. (I am now always talking about IR flow.) We can make a multidimensional plot with axes λ, m, and so on, and fill it with arrows showing the flow of the couplings in parameter space. Fig. 10.7 is a two-dimensional cartoon illustrating this situation. There can be fixed points, places where renormalization group flow vanishes. But the notion of stability or instability with respect to scale changes must be generalized: is is possible (and it is often the case) that the fixed point is repulsive in one or more directions and attractive in other directions. This means that in order for the system to flow into a fixed point the values of the couplings have to be tuned to take special values so that the flow is entirely attractive, into the fixed point. We call these couplings "relevant couplings" or "relevant directions" with respect to the fixed point. In a statistical mechanics context, the couplings in the field theory are functions of temperature or other external probes. The long distance behavior of the system depends on where the external probes are set. And in principle, the experimentalist can tune them. Then, once the relevant couplings are appropriately set, as the system is observed at ever longer distance, the other (irrelevant) couplings will flow into the fixed point and the long distance physics of the system will be governed by physics at the fixed point.

In Fig. 10.7 we see that starting anywhere on a line which flows into the fixed point, we will end up at the fixed point. Each of the points on that line have a different set of short distance couplings. The long distance physics for all of these systems will be governed by the same fixed point; it will not depend on what the short distance physics was. This phenomenon is called "universality." Universality generally depends on only a few quantities, such as the dimensionality of space and the local symmetries of the short distance degrees of freedom (real versus complex scalar fields, for example). A statistical mechanics example: all three-dimensional isotropic ferromagnets form a universality class, close to criticality they have identical behavior.

And what happens at a fixed point? There, since $\bar{\lambda}$ doesn't change, the last integral in

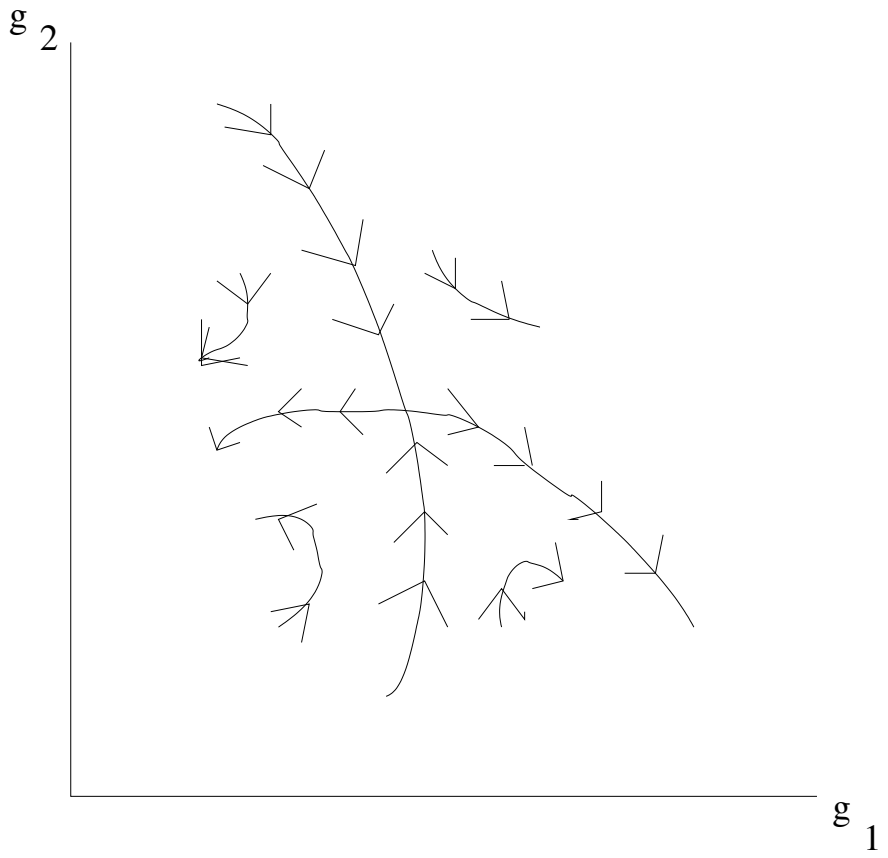

Figure 10.7: A renormalization group flow with one attractive direction and one repulsive direction.

Eq. 10.74 simplifies, and the correlation function is just

$$\begin{aligned} \bar{\Gamma}^{(n)}(sp) &= s^{d_n}\Gamma^{(n)}(p;,\bar{\lambda},\mu)\exp\left(-n\gamma_\phi(\bar{\lambda})\ln s\right) \\ &= s^{d_n-n\gamma^*_\phi}\Gamma(p) \end{aligned} \tag{10.90}$$

or

$$\Gamma(p) \sim p^{d_n-n\gamma^*_\phi}. \tag{10.91}$$

For example, the two-point function is

$$\Gamma^{(2)}(p) \sim p^{2-2\gamma^*_\phi}. \tag{10.92}$$

This behavior is called "scaling." It says that at the fixed point, the only relevant quantity with dimensions of energy is the common momentum scale of the particles participating

in the correlation function. The underlying physics has no scale. It is almost a massless theory: if there were an excitation with mass m we would expect $\Gamma^{(2)}(p) \sim p^2 + m^2$ and Eq. 10.92 is"almost" telling us that $m = 0$. However, these critical systems do not really have particles; $\Gamma^{(2)}(p)$ is not p^2. And pause to Fourier transform $\Gamma^{(2n}(p)$ back to coordinate space: all correlation functons decay with distance as a power law, not exponentially. It turns out (but this is another story for another day) that this is the physics of statistical systems tuned to be at a (second order) critical point.

10.5 Renormalization prescriptions

There is a fair amount of arbitrariness in defining m, λ and ϕ. Often a choice is made for convenience (ease of comparison with experiment). Some schemes give rise to large coefficients in higher orders of perturbation theory, and are disfavored. Sometimes it is possible to directly relate the renormalized parameters to physical quantities. For example, in QED, a physical electric charge could be defined through the Thompson cross section. Sometimes, this cannot be done so directly. QCD is an example – because of confinement there are no free quarks.

Here are several examples of prescriptions (all for ϕ^4):

First, one could define $\Gamma^{(2)}(p, m_A) = p^2 + m_A^2$ at $p^2 = m_A^2$ and $\Gamma^{(4)}(p_1, p_2, p_3, p_4) = -\lambda_A$ at $p_i = 0$. (This fixes both the charge and field renormalization.)

Next, one could use any other momentum point. For example $\Gamma^{(2)}(p, m_A) = p^2 + m_B^2$ at $p^2 = M$ and $\Gamma^{(4)}(p_1, p_2, p_3, p_4) = -\lambda_B$ at $p_i p_j = M^2(\delta_{ij} - 1/4)$. This is $s = t = u = M^2$, off mass shell, arbitrary M.)

Finally, in dimensional regularization there is the so-called modified minimal subtraction ($\overline{MS}$) prescription, which you will see in a later course.

These are all Euclidean space prescriptions. Let's now return to Minkowski space: recall $p_E = (p_E^0, \vec{p}_E) \to p_M = (ip_E^0, \vec{p}_E)$. The propagator transforms into

$$\frac{1}{p_E^2 + m^2} \to \frac{1}{-p_M^2 + m^2} \to \frac{1}{-p_M^2 + m^2 - i\epsilon} \tag{10.93}$$

using $p_M^2 = (p_0^m)^2 - (\vec{p}_M)^2$ and inserting the $i\epsilon$ to have causal propagators. Of course, there

are i's to chase in the Dyson series, too. But recall the (Euclidean) four point amplitude:

$$\bar{\Gamma}^{(4)} = -\lambda_0 + \frac{3}{32\pi^2}\lambda_0^2 \ln\frac{\Lambda^2}{\mu^2} - \frac{\lambda_0^2}{32\pi^2}[I(s) + I(t) + I(u)] \tag{10.94}$$

where

$$I(s) = \int_0^1 dx \ln[\frac{m_0^2 + sx(1-x)}{\mu^2}]. \tag{10.95}$$

Let's define our coupling by imposing the second prescription,

$$\begin{aligned} \bar{\Gamma}^{(4)}(s=t=u=M^2) &\equiv -\lambda_B \\ &= -\lambda_0 + \frac{3}{32\pi^2}\lambda_0^2 \ln\frac{\Lambda^2}{\mu^2} - \frac{3\lambda_0^2}{32\pi^2}\int_0^1 dx \ln[\frac{m_0^2 + M^2x(1-x)}{\mu^2}]. \end{aligned} \tag{10.96}$$

Away from this kinematic point the amplitude expressed in terms of λ_B is

$$\bar{\Gamma}^{(4)} = \lambda_B - \frac{\lambda_0^2}{32\pi^2}\int_0^1 dx \ln[\frac{m_0^2 + sx(1-x)}{m_0^2 + M^2x(1-x)}] + (s \to t) + (s \to u) \tag{10.97}$$

and to this order in perturbation theory, $\lambda_0 = \lambda_B$ and $m_0 = m$. The Minkowski space version of this expression is

$$\bar{\Gamma}^{(4)} = -[-\lambda_B - \frac{\lambda_B^2}{32\pi^2}(\int_0^1 dx \ln[\frac{m^2 - i\epsilon + sx(1-x)}{m^2 - i\epsilon + M^2x(1-x)}] + (s \to t) + (s \to u). \tag{10.98}$$

The new terms, beyond the λ_B piece, are the predictions of the theory for scattering away from the fiducial point. Notice, by the way, that the scattering amplitude has a branch cut beginning at $s = 4m^2$ – but that is a story for another day.

This is a renormalizable quantum field theory. All the Λ dependence has disappeared from predictions of physical processes. Let's summarize what we have found.

Our bare theory was expressed in terms of four quantities, m_0, λ_0, ϕ_0 and Λ. We had a finite number (three, in this case) of cutoff dependent quantities, and their cutoff dependence could be absorbed into a finite number of parameters which could be (and must be) determined by experiment. There is no evidence of the cutoff, when Γ's are expressed in terms of m, λ and ϕ. Then we asked, how do Green's functions change when the momentum scale changes? The physical answer (mostly) is that the couplings λ and m become scale dependent ($\bar{\lambda}(s)$, $\bar{m}(s)$). We derived a particular renormalization group equation by varying the bare Γ_0 with respect to the renormalization point μ,

$$\left[-s\frac{\partial}{\partial s} + \beta(\lambda)\frac{\partial}{\partial\lambda} + (\gamma_m - 1)m\frac{\partial}{\partial m} - n\gamma_\phi + d_n\right]\bar{\Gamma}_R^{(n)}(sp_1, sp_2 \ldots; m, \lambda, \mu) = 0. \tag{10.99}$$

(There are many alternative formulations of the renormalization group.)

If there are many couplings the equation for the beta function becomes a matrix equation involving all the couplings

$$\mu\frac{\partial\lambda_i}{\partial\mu} = \sum_j \beta_{ij}(\lambda_1, \lambda_2 \dots). \tag{10.100}$$

In this case, there is a many-dimensional space in which all the couplings flow. Under a change in scale some of them may flow into a fixed point; some may flow away.

10.6 More general couplings

Let's suppose we have more complicated couplings. Dimensional analysis tells us that

$$[\lambda_n\phi^n] = L^{-D} = \Lambda^D \tag{10.101}$$

and since the engineering dimension of ϕ is

$$[\phi] = L^{1-D/2} = \Lambda^{-(1-D/2)}, \tag{10.102}$$

we know that the dimension of the coupling is

$$[\lambda^n] = \Lambda^{n+D-nD/2} \equiv \Lambda^\epsilon. \tag{10.103}$$

Generally, couplings are dimensional. We can parameterize a coupling as $\lambda_n = \lambda_0\Lambda^\epsilon$ using the cutoff to set the energy scale. Then λ_0 is dimensionless number. How does the pure number λ_0 vary if we vary Λ (holding λ_n fixed? This is easy:

$$\Lambda\frac{\partial\lambda_n}{\partial\Lambda} = 0 = \epsilon\Lambda^\epsilon\lambda_0 + \Lambda^\epsilon\Lambda\frac{\partial\lambda_0}{\partial\Lambda} \tag{10.104}$$

or

$$\Lambda\frac{\partial\lambda_0}{\partial\Lambda} = -\epsilon\lambda_0 = \beta(\lambda_0) \tag{10.105}$$

in this case

$$\frac{\lambda_0(\Lambda')}{\lambda_0(\Lambda)} = \left(\frac{\Lambda'}{\Lambda}\right)^{-\epsilon} \tag{10.106}$$

so if $\epsilon > 0$, if we increase Λ we must decrease the bare coupling λ_0 to keep the renormalized coupling fixed.

Now suppose Λ is fixed but the momentum scale s varies.Our theory has all these couplings, but how do they run? We have just computed the beta function, so we know that

$$s\frac{\partial \lambda_n}{\partial s} = \beta(\lambda_n) = -\epsilon\lambda_n \tag{10.107}$$

or

$$\lambda_n(s) = \lambda_n(1)s^{-\epsilon} \tag{10.108}$$

The UV flow of λ_n is toward greater values if $\epsilon < 0$ or to lesser values if $\epsilon > 0$. The IR flow is opposite, λ_n flows to the origin if $\epsilon < 0$. Table 10.6 gives some examples.

interaction	n	$\epsilon(D=4)$	IR flow
$m^2\phi^2$	2	$2>0$	increases
$\lambda_4\phi^4$	4	0	stagnant (logarithmic, $\beta(\lambda) = .O(\lambda^2)$
$\lambda_6\phi^6$	6	$-2<0$	decreases

Table 10.1: Flows for various couplings

Suppose we are in $D = 4$ in a theory defined with a UV cutoff Λ, and the theory has nonrenormalizable interactions, $n > 4$ ϕ^n terms. We are interested in computing processes at energy scales far below Λ. As we go to long distance or low energy scales, the nonrenormalizable couplings run to zero! The long distance effective scalar field theory has only m^2 and λ_4 couplings. All other couplings are small like Λ^ϵ. Green's functions for processes at scale p pick up contributions of size $(p/\Lambda)^{|\epsilon|}$ (this is p^2/Λ^2 for a ϕ^6 coupling). This result is also called "universality" and it is a specific example of physics we saw earlier: an arbitrary $V_0(\phi_0)$ looks like $m^2\phi^2 + \lambda\phi^4$ at a small momentum scale $p \ll \Lambda$. This inspires a rhetorical question I heard in class from Weinberg around 1973: "Why are the theories that describe Nature renormalizable quantum field theories?" and his answer, from 5-6 years later, was "Because the scale of new physics (Λ) is at very high energy."

.

10.7 Where do Lagrangians come from, or the story of effective field theories

Imagine that we had a Theory of Everything, and also imagine that we were atomic physicists. Should we use the Theory of Everything to make predictions? That would be an expensive

undertaking, since the Theory of Everything presumably includes physics at very high energy scales.

A better alternative is to somehow remove high energy degrees of freedom from the theory and construct a "low energy" or "effective" theory, which we then use to compute atomic properties. The effective theory is only supposed to be valid up to some maximum energy scale, and it only includes the low energy degrees of freedom which are relevant for the calculations we want to do.

"Remove" has three possibilities. The first one is to integrate out degrees of freedom. For this example, work with the path integral. We start with a theory with an energy cutoff Λ whose partition function is

$$Z_\Lambda = \int [d\phi_k]_{k<\Lambda} \exp(-S(\phi)_\Lambda. \tag{10.109}$$

The functional integral includes fields defined at all energy scales up to the cutoff. We then break the integration variables $[d\phi_k]_{k<\Lambda}$ apart and perform the functional integral over the higher energy modes of ϕ, leaving behind a theory which only includes low energy degrees of freedom.

$$\begin{aligned} Z_{\Lambda_1} &= \int [d\phi_k]_{k<\Lambda_1} \left(\int [d\phi_k]_{\Lambda_1<k<\Lambda} \exp(-S(\phi)_\Lambda \right) \\ &= \int [d\phi_k]_{k<\Lambda_1} \exp(-S(\phi)_{\Lambda_1}). \end{aligned} \tag{10.110}$$

The original theory had an action $S(\phi)_\Lambda$ and a set of operators with an associated set of couplings $g_n(\Lambda)$. The lower energy action $S(\phi)_{\Lambda_1}$ is characterized by its own set of operators with couplings $g_m(\Lambda_1)$ which of course are functions of the $g_n(\Lambda)$'s. ($S(\phi)_{\Lambda_1}$ is called the low-energy or Wilsonian effective action.) An analog of renormalization group flow as described in earlier sections of these notes would be to ask what is the relation of the low scale couplings to the high scale ones, $g_m(\Lambda_1, g_n(\Lambda))$.

A second way to proceed would be to determine the couplings $g_m(\Lambda_1)$ from targeted calculations in the Theory of Everything. An example of such a calculation wold be to determine the magnetic moment of a nonrelativistic electron starting with QED and a relativistic electron.

The first possibility is very hard to do, but sometimes the construction of $S(\phi)_{\Lambda_1}$ can be carried out. The second possibility can also be very hard to implement. So there is a third

possibility: simply to ask, what is the most general low energy theory? And then to work with it. This is easy: the partiton function is built only in terms of the low energy degrees of freedom and the action is

$$S(\phi)_{\Lambda_1} = \int d^D x \sum_i g_i \mathcal{O}_i \tag{10.111}$$

where the sum runs over all local operators of the low energy degrees of freedom which are allowed by symmetry. This does not seem to be such a good idea at first glance, since the sum could in principle include an infinite number of operators (and couplings). We need some way to organize the problem, that will give us a (good) approximate action. Let's consider dimensional analysis again.

If the typical energy scale in the problem is E we can use the kinetic term to set the scale. We assume that S_0 is $O(1)$ and the relevant length scale is $1/E$, so field variables scale as

$$S_0 = \frac{1}{2}\int d^D x (\partial_\mu \phi)^2 \to [\phi] = E^{-1+D/2}. \tag{10.112}$$

Then if an operator $\mathcal{O}$ is built of M ϕ's and N derivatives, $\mathcal{O}_i$ scales as E^{δ_i} where

$$\delta_i = M(-1 + \frac{D}{2}) + N \tag{10.113}$$

The corresponding g_i has to have dimension $\Lambda^{D-\delta_i}$ to compensate. Now we replace the g_i's by dimensionless couplings $\lambda_i = \Lambda^{\delta_i - D} g_i$. We presume that the λ_i's are just a set of numbers whose size is order unity. Now, for a process occurring at scale E, we estimate that

$$\int d^D x \mathcal{O}_i \sim E^{\delta_i - D} \tag{10.114}$$

so that a term in the action is of order

$$g_i \mathcal{O}_i \sim \Lambda_i \left(\frac{E}{\Lambda}\right)^{\delta_i - D}. \tag{10.115}$$

We see that if $\delta_i > D$, this term is less and less important at small energy (where $E/\Lambda < 1$) – it is "irrelevant." Here is the vocabulary: if $\delta_i < D$ the size of the term in S grows as E falls. Such terms are called "relevant" or (old language) "super-renormalizable". If $\delta_i = D$, the size of the term remains constant as E falls. Such operators are called "marginal" or "renormalizable." Finally, operators with $\delta_i > D$ shrink as E falls; they are "irrelevant" or "nonrenormalizable." We have already seen the words "renormalizable" and "nonrenormalizable" when we discussed the dependence of observables on the cutoff Λ; here they are

again, but with a twist. The nonrenormalizable operators are less and less important as we go to ever smaller values of E.

And now for the punch line: typically, there are only a finite number of relevant and marginal operators, since the condition for not being irrelevant is

$$\delta = M(-1 + \frac{D}{2}) + N \leq D. \tag{10.116}$$

For example, in $D = 4$ scalar field theory there is one relevant operator, $m^2\phi^2$ and one marginal one, $\lambda\phi^4$.

Why do we emphasize the free action Eq. 10.112 in this analysis? It is because we are assuming that the theory is weakly coupled, so the free action determines the size of the typical fluctuations of the field variables, and then of the sizes of matrix elements. It's necessary that the coefficient of the dominant term in S is dimensionless when we rescale in order to see this. And in $\frac{1}{2}(\partial_\mu\phi)^2$, $M = N = 2$, $\delta = D$.

Another way to say this is that we are comparing the size of all operators to that of the kinetic term.

And another variation on this idea: assume that the kinetic term dominates. Scale all energies and momenta by a factor s, so lengths and time scale by $1/s$. The volume and derivatives in Eq. 10.112 scale as s^{2-D}, and fluctuations in ϕ scale as $s^{-1+D/2}$. Then the ith interaction term scales as $s^{\delta_i - D}$. If $s < 1$ and $\delta_i > D$, the operator scales to irrelevancy.

Sometimes there can be more than one "kinetic term." In membranes, for example, there are ∇^2 tension terms and ∇^4 rigidity terms, but at any given momentum, one of these will dominate and set the scaling.

	high energy theory	scale	low energy theory
(1)	string theory	$M_{string} \sim 10^{18}$ GeV	field theory of gravity and matter
(2)	grand unified theory	$M_{GUT} \sim 10^{16}$ GeV	$SU(3) \times SU(2) \times U(1)$
(3)	Weinberg-Salam model	$M_W \sim 80$ GeV	Fermi theory
(4)	QCD	0.5-1 GeV	pions (and nucleons)
(5)	lattice field theory	(lattice spacing)$^{-1}$	continuum field theory

Table 10.2: Examples of effective field theories

There are many cases where we know the high energy theory and its low energy effective field theory. Some examples of effective field theories are given in Table 10.7. In cases (2)

and (3), the high energy and low energy theories are both perturbative. In QCD, the high energy fields (quarks and gluons) are not the low energy fields (pions and nucleons). At really long distances, QCD becomes strongly interacting. The only light modes are pions. They are Goldstone bosons: Goldstone's theorem says that if a continuous global symmetry is broken, there will be massless modes corresponding to the unbroken degrees of freedom.

More about case (3) in the Table: Consider muon decay, $\mu^- \to \nu_\mu e \bar{\nu}_e$, which proceeds through a virtual W boson. The amplitude is

$$M = (-ig)^2 \bar{u}_e \Gamma^\mu v_{\nu_e} \frac{ig_{\mu\nu}}{q^2 - M_W^2} \bar{v}_{\nu_\mu} \Gamma^\nu u_\mu \tag{10.117}$$

with $\Gamma^\mu = \gamma^\mu(1-\gamma_5)$. The fundamental interaction is $\mathcal{L} = g\bar{f}_i \Gamma^\mu f_j W_\mu$ with a dimensionless coupling g. At very low energies, where $q^2 \ll m_W^2$ the expression simplifies,

$$\frac{ig_{\mu\nu}}{q^2 - M_W^2} \to -\frac{ig_{\mu\nu}}{M_W^2} \tag{10.118}$$

which is a four-fermion interaction. We could generate it from an effective Lagrangian containing the term

$$\mathcal{L} = G_F(\bar{f}_1 \Gamma^\mu f_2)(\bar{f}_3 \Gamma_\mu f_4) \tag{10.119}$$

and $G_F = g^2/M_W^2$ is a dimensionful coupling, the "Fermi coupling constant," with a value $G_f \sim 10^{-5}$ GeV^{-2}. The gauge boson has disappeared from the theory. This is called "Fermi theory," and it is what Fermi wrote down in the 1930's to describe the weak interactions. Let's check dimensions. The kinetic term gives us $[\bar{\psi}\partial\psi] = \Lambda^4$ so $[\bar{\psi}\psi] = \lambda^3$ or $[\psi] = \Lambda^{3/2}$. $[(\bar{f}\Gamma f)^2] = \Lambda^{4\times 3/2} = [\Lambda]^6$. (we call this a "dimension-6 operator."). The units of G_F are Λ^{-2} to compensate. This is an irrelevant operator, which means that it grows as the energy rises. There will be an energy scale where Fermi theory breaks down and it must be replaced by something new. The "'something new" is the W boson, of course. This sounds like we are just closing a circle, but in olden times, before the W, all we had was Fermi theory. Considerations like this, looking from the Fermi theory to higher energy, could be – and were – used to estimate the "scale of new physics" or to be more mundane, the mass of the W boson.

There are other similar stories to be told. Searches for the electron electric dipole moment are done in terms of a search for new dimension-six operators, and the limits are typically quoted as lower bounds on Λ, the scale of the new physics.

Finally, in condensed atom gases which form Bose - Einstein condensates, the high energy system is electrons and nuclei, or maybe atoms, the low energy system are the collective excitations of the BEC, a superfluid (as we saw earlier).

Presumably no field theory we encounter is complete up to arbitrarily high energies. They are all effective field theories, valid up to some cutoff scale. If that is the case, is renormalization unimportant, since the cutoff scale renders all integrals finite? No, renormalization tells us about scale dependent couplings, so we know how to move around in energy while remaining in our effective field theory description. It also tells us that low energy physics depends on the short distance theory only through its relevant and marginal couplings, and possibly through some leading irrelevant couplings, if one measures to sufficiently high accuracy.

And what are these operators? Naive scaling tells us – except that if a coupling gets large, naive scaling can fail. Recall the expression for an n-point amplitude at a fixed point

$$\begin{aligned}\bar{\Gamma}^{(n)}(sp) &= s^{d_n}\Gamma^{(n)}(p;\bar{\lambda},\mu)\exp\left(-n\gamma_\phi(\bar{\lambda})\ln s\right)\\ &= s^{d_n-n\gamma^*_\phi}\Gamma(p)\\ &= s^{d'_n}.\end{aligned} \tag{10.120}$$

If $\gamma_\phi(\lambda^*)$ is large ($O(1)$), $\delta'_n \neq \delta_n$. If all we have is perturbation theory, we may lose control of the calculation. We may not know which operators are relevant, marginal or irrelevant.

Marginal operators show the most important corrections to naive scaling. We can write

$$E\frac{\partial g}{\partial E} = \beta(g) = bg^2 + O(g^3). \tag{10.121}$$

The case where $\beta(g) = 0$ exactly is really marginal. In the case $b > 0$, g falls as E falls. This is a "marginally irrelevant" coupling. The other case of $b < 0$ where g rises as E falls, is called "marginally relevant.' Then

$$g(E) = \frac{g(\Lambda)}{1 + bg(\Lambda)\ln\frac{\Lambda}{E}} \tag{10.122}$$

and g becomes large at $E \sim \Lambda\exp(1/(bg(\Lambda))$. Something new has to happen at or below that energy. In QCD, the effects of confinement set in.

Irrelevant couplings are useful because they tell us the range of validity at high energy. When they grow to be $O(1)$, new physics has to appear.

Relevant operators are dangerous. Think about ϕ^2 in a $D = 4$ scalar field theory: $D-\delta = 2$. In $S_\Lambda(\phi)$ there is a term $(\lambda_{\phi^2}\Lambda^2)\phi^2$. Why shouldn't the pure number λ_{ϕ^2} be order unity? There has to be some kind of story to make it small, some un-natural fine tuning. But if $\lambda_{\phi^2} \sim 1$, the ϕ particle is heavy and it should not appear in the low energy effective theory

at all. This leads to an (esthetic?) rule: effective field theories must be "natural," meaning that masses should be forbidden by symmetries. Gauge interactions are allowed because gauge invariance forbids the gauge bosons from developing a mass. A symmetry like chiral symmetry, if present, forbids fermions from getting a mass. Scalars should be Goldstone bosons, or something else, like supersymmetry, should prevent them from acquiring large masses. i

In the condensed matter context, the situation is that we want to make correlation lengths large with respect to some short distance cutoff distance. To make correlation lengths diverge, the relevant operators must be tuned by hand. The analog of m^2 is presumed to vary with temperature like $T - T_c$ and only by tuning to $T = T_c$ does the system exhibit criticality.

Is the Standard Model natural? No. The high scale is either some grand unification scale ($\Lambda = 10^{16}$ GeV) or the Planck scale $\Lambda = 10^{19}$ GeV). The Higgs mass is 120 GeV, far below either of these values of Λ. But we saw from our calculation of Σ that radiative corrections could shift the squared mass of the Higgs by an amount Λ^2. Why is the Higgs (at 120 GeV) so light?

And two final questions, with answers, courtesy Joe Polchinski, hep-th/9210046:

Q1: Doesn't the infinite set of irrelevant operators mean that effective field theory has no predictive power?

A1: No, you can do calculations which are accurate to $(E/\Lambda)^{\delta-D}$, and if Λ is known, you know how accurate your calculation will be.

Q2: Does that mean that quantum field theory is an approximation which might have little to do with some underlying, more fundamental description of nature? And isn't renormalization a bad thing, because it implies that we can only probe the high energy theory through a small number of parameters?

A2 (from Joe): Nobody ever promised you a rose garden.

10.8 A few words about renormalization in QED

The Lagrangian for QED is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi - m\bar{\psi}\psi \tag{10.123}$$

(plus gauge fixing terms). We are interested first in determining the superficial degree of divergence of processes in QED. We imagine that we have a Feynman diagram in $D = 4$ dimensions with

- N vertices, which give $N - 1$ delta-functions
- $(I_e + I_\gamma - (N-1)) \int d^4k$ aggregate internal integrations, each counting as Λ^4
- E_γ external photons
- E_e external fermions
- I_γ internal photon lines – these count as Λ^{-2} from $1/q^2$ propagators
- I_e internal fermion lines. Each counts as Λ^{-1} since the propagator is $1/\not{p}$

This gives a degree of divergence of

$$\begin{aligned} d &= 4[I_e + I_\gamma - (N-1)] - 2I_\gamma - I_e \\ &= 4 + 3I_e + 2I_\gamma - 4N. \end{aligned} \tag{10.124}$$

Each vertex brings two fermion lines and one photon line together, so counting photons gives the relation

$$N = E_\gamma + 2I_\gamma \tag{10.125}$$

(each internal photon connects two vertices) and for electrons, count the number of electrons per vertex, two, times the number of vertices N to discover

$$2N = 2I_e + E_e. \tag{10.126}$$

Putting the last three equations together, we find that

$$d = 4 - \frac{3}{2}E_e - E_\gamma. \tag{10.127}$$

This is independent of the order of perturbation theory since it is independent of N. This tells us that QED is a renormalizable theory.

There are four physical quantities at our disposal: e, m, ψ and A_μ. And it turns out that there are four divergent quantities (see Fig. 10.8):

The vertex function ($E_e = 2$, $E_\gamma = 1$) has $d = 0$ and is logarithmically divergent.

The electron self energy appears superficially to have $d = 1$ but that is not quite correct. We can write

$$\Sigma(p) = A + B\not{p} + Cm + finite \tag{10.128}$$

However, chiral symmetry says that $A = 0$. If it happened that the bare m_0 were zero, the electronic part of $\mathcal{L}$ would factorize into two terms, one involving only ψ_L and $\bar{\psi}_L$ and a second term involving the right handed fields. A mass term couples left handed and right handed fields and such a term cannot be generated by radiative processes. This means that A=0 in Eq. 10.128. If there is a nonzero mass to start with, interactions can shift it, so there can be a Cm term in Σ. The factors of m and $\not{p}$ in the numerator of $\Sigma(p)$ means that B and C have $d = 0$: they are each logarithmically divergent.

Finally, we have the photon self energy,

$$\Pi_{\mu\nu} = Dg_{\mu\nu} + q_\mu q_\nu C' + Cq^2 g_{\mu\nu} + finite. \tag{10.129}$$

Again, appearances are deceiving. Current conservation (equivalently, transversality) says $q_\mu \Pi_{\mu\nu} = 0$ so $D = 0$ and

$$\Pi_{\mu\nu} = \bar{\Pi}_{\mu\nu}(q^2)[q_\mu q_\nu - q^2 g_{\mu\nu}] \tag{10.130}$$

where $\bar{\Pi}_{\mu\nu}(q^2)$ is logarithmically divergent. Thus there are four divergent quantities (L in the vertex, B and C in Σ, and $\bar{\Pi}_{\mu\nu}(q^2)$), to be absorbed into renormalization of the bare quantities e_0, m_0, ψ_0 and A_0^μ to give the (renormalized) observable ones.

We have two final processes, shown in the figure. The "photon splitting" term $\langle T(AAA)\rangle$ vanishes due to charge conjugation symmetry. The light by light scattering term has $E_\gamma = 4$ and $E_e = 0$ so it has $d = 0$ at first glance, but a transversality argument says that there are additional factors of q in the amplitude which make it finite (and very small, the amplitude begins at order e^4). (A physical argument: the process has to involve products of $F_{\mu\nu}$'s in order to be gauge invariant, and $F_{\mu\nu} \sim q_\mu A_\nu - q_\nu A_\mu$; the factors of q have to be compensated by additional momentum terms in the denominators of the Feynman graphs which render them finite.)

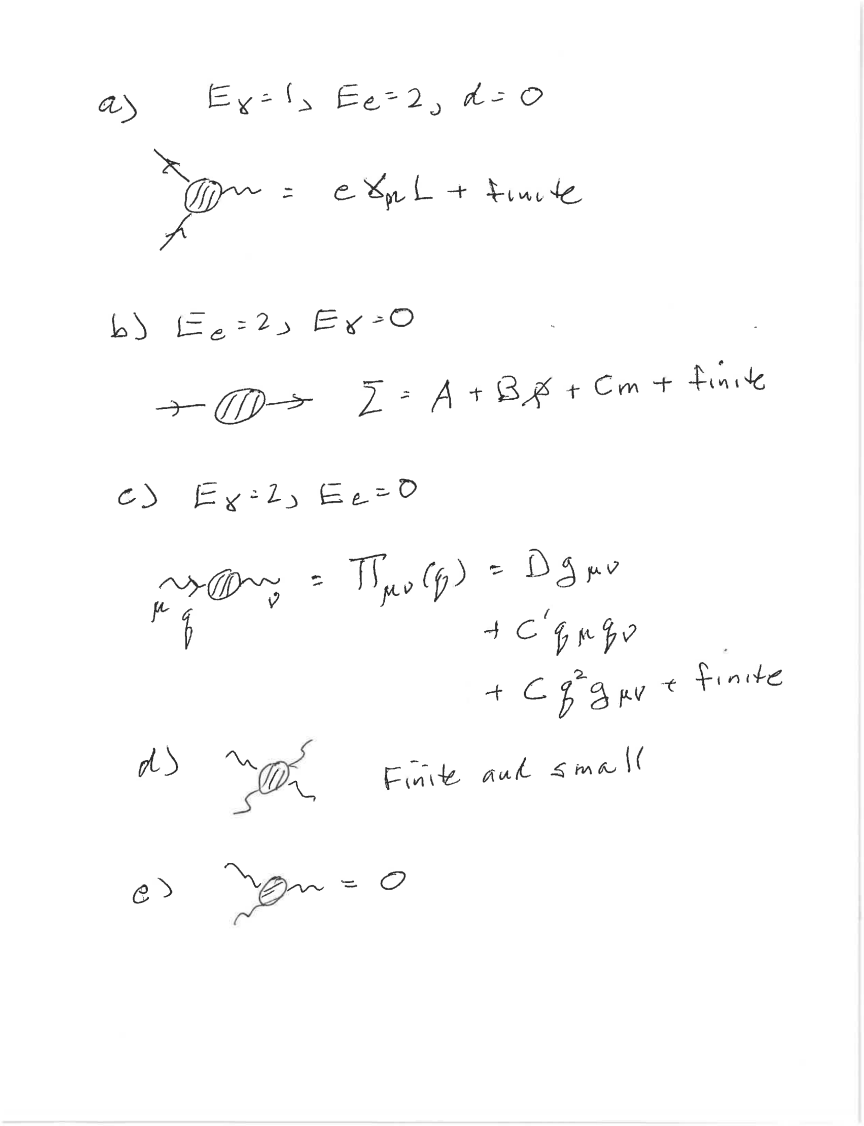

Figure 10.8: Fundamental processes in QED.

The semester is rapidly coming to an end, so let us just jump to the two "classical" (1940's) tests of QED, the anomalous magnetic moment of the electron and the Lamb shift. Both of these were discovered in atomic physics experiments.

First, $g - 2$. The electron-electron-photon vertex can be parameterized as

$$\Gamma_\mu = e[\gamma_\mu L + 2i\sigma_{\mu\nu}q_\nu F_2] \tag{10.131}$$

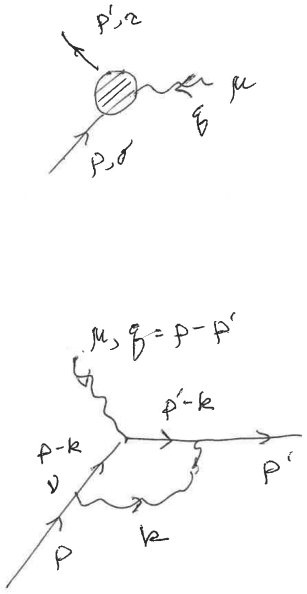

Figure 10.9: Graphs for $g-2$.

The first term gives the Dirac g-factor of two for the magnetic moment. The second term, which begins at order e^2, gives the so-called anomalous magnetic moment: the real magnetic moment is equal to the Dirac value plus a small correction,

$$\mu = \mu_{Dirac}(1 + \frac{\alpha}{2\pi} + \dots) \tag{10.132}$$

with $\mu_{Dirac} = e/(mc) = ge/(2mc)$ so that the Dirac value is $g = 2$. We can rewrite the expression for μ in terms of $g - 2$: the answer is

$$\frac{g-2}{2} = \frac{\alpha}{2\pi} + \dots . \tag{10.133}$$

The lowest order coefficient was first computed by Schwinger in 1947. The QED result is known to order $(\alpha/\pi)^4$ and at present theory and experiment (done with trapped electrons) agree to a part in 10^{12}. The $(\alpha/\pi)^4$ calculation involved 891 diagrams, individually all gauge dependent, most of them infrared and ultraviolet divergent. These days the magnetic moment of the muon is more newsworthy, since non-QED effects (the strong interactions, mostly) make a bigger contribution to the muon than they do to the electron. This makes muonic $g - 2$ a potential place to see new physics.The most recent experimental result is in arXiv:2506.03069 and a recent white paper with a theory discussion is is arXiv:2505.21476.

Let's do a skim into the calculation. With the kinematics in Fig. 10.9 we imagine writing the matrix element of the current in terms of the most general operator we can construct

$$\langle p'\tau|eJ_\mu|p\sigma\rangle = \bar{u}(p',\tau)\mathcal{O}_\mu(p',p)u(p,\sigma), \tag{10.134}$$

and we can build it with a term parameterizing the size of every vector-valued quantity in the amplitude

$$\mathcal{O}_\mu(p',p) = \gamma_\mu F_i(q^2) + 2i\sigma_{\mu\nu}q^\nu F_2(q^2) + q_\mu F_3(q^2) \tag{10.135}$$

with $q = p' - p$ being the photon's four momentum. The F_i's are form factors which in principle can be arbitrary functions of q^2. Here $\sigma_{\mu\nu} = \gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu$. Current conservation requires $\partial^\mu J_\mu = 0$ or in momentum space $q^\mu J_\mu = 0$. Expanding out Eq. 10.135, we must have

$$0 = \bar{u}(p',\tau)[\not{q}F_1 + 2i\sigma_{\mu\nu}q^\mu q^\nu F_2 + q^2 F_3]u(p,\sigma). \tag{10.136}$$

By antisymmetry, $2i\sigma_{\mu\nu}q_\mu q_\nu = 0$. We can rewrite $\not{q} = \not{p}' - \not{p}$ and use the Dirac equation to see that

$$\bar{u}(p')(\not{p}' - \not{p})u(p) = \bar{u}(p')(m - m)u(p) = 0. \tag{10.137}$$

So, to satisfy Eq. 10.136, it must be that $F_3(q^2) = 0$. Finally, at $q = p - p' = 0$ the current is $J_\mu = \bar{u}\gamma_\mu u F_1(0)$ and this says that $F_1(0)$ must be equal to the renormalized charge e_r.

The second term gives us the anomalous magnetic moment. To see this, use

$$\sigma_{ij} = \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}, \tag{10.138}$$

and pick the direction $\mu = \hat{z}$. Then

$$\begin{aligned} J_\mu A^\mu &= -\vec{J}\cdot\vec{A} \\ &= 2iF_2(q^2)\begin{pmatrix} \sigma_x q_y - \sigma_y q_x & 0 \\ 0 & \sigma_x q_y - \sigma_y q_x \end{pmatrix} A_z \\ &= 2F_2(q^2)\begin{pmatrix} \vec{\sigma}\cdot(i\vec{q}\times\vec{A}) & 0 \\ 0 & \vec{\sigma}\cdot(i\vec{q}\times\vec{A}) \end{pmatrix} \\ &= 2F_s(q^2)\vec{\sigma}\cdot\vec{B} \\ &= \mu\vec{S}\cdot\vec{B} \end{aligned} \tag{10.139}$$

with the relation between form factor and magnetic moment

$$\mu = 4F_2(0). \tag{10.140}$$

The lowest order graph is shown in Fig. 10.9. Formally, it is logarithmic divergent. However, the presence of the $\sigma_{\mu\nu}q^\nu$ in the definition of the form factor means that the integrand has to have an additional negative power of the loop momentum to make the dimensions work out. This renders the magnetic moment piece of the diagram actually finite. This can be seen if one is willing to do the calculation in a devious way!

Next, we have the Lamb shift. In the Schrödinger equation, and in the Dirac equation, the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ levels of a Coulomb bound state are degenerate. In reality, they are slightly split, with an experimental value $\Delta E(2S-2P) = 1040$ MHz. JILA people might enjoy reading the discovery paper, Lamb and Retherford, Phys. Rev. **72**, 241 (1947). The splitting is due to the virtual emission and absorption of a photon by the electron. The effect is mostly nonrelativistic and was worked out first by Bethe (while riding on a train coming back from the meeting where the experimental result was announced) – see H. Bethe, Phys. Rev. **72**, 339 (1947). Here is a sketch:

Start with a Hamiltonian for a nonrelativistic charged particle in a potential

$$H = \frac{1}{2m}(\vec{p} - \frac{e}{c}\vec{A})^2 + V(r) \tag{10.141}$$

and treat the $\frac{e}{mc}\vec{p}\cdot\vec{A}$ term as a perturbation. The energy shift of a level n is given by second order perturbation theory,

$$\Delta E_n = e^2 \int \frac{d^3k}{2k(2\pi)^3} \sum_\lambda \sum_j \frac{|\left\langle n|\frac{\vec{p}\cdot\vec{\epsilon}_\lambda}{m}|j\right\rangle|^2}{E_n - E_j - k}. \tag{10.142}$$

The calculation is done exactly like the ones for electronic transitions in a good first year graduate quantum mechanics class. Pick the two polarizations in and out of the plane defined by $\vec{p}$ and $\vec{k}$; the in-plane term gives a $\sin\theta$ and the out of plane one vanishes. Then

$$\Delta E = \frac{4\pi\alpha(2\pi)}{2(2\pi)^3}\int_0^K kdk\int_{-1}^1 d\cos\theta\sin^2\theta\sum_j \frac{|\left\langle n|\frac{\vec{p}}{m}|j\right\rangle|^2}{E_n - E_j - k} \tag{10.143}$$

cutting of the k integral at K. This gives

$$\Delta E_n = \frac{2\alpha}{3\pi}\int_0^K kdk\sum_j \frac{|\left\langle n|\frac{\vec{p}}{m}|j\right\rangle|^2}{E_n - E_j - k}. \tag{10.144}$$

If we perform the k integral, we get

$$\Delta E_n = \frac{2\alpha}{3\pi}\left(-\sum_j \frac{|\langle n|\vec{p}|j\rangle|^2}{m^2}K + \sum_j \frac{E_j - E_n}{m^2}|\langle j|\vec{p}|n\rangle|^2 \ln\frac{|E_j - E_n + K|}{|E_j - E_n|}\right). \tag{10.145}$$

The expression is strongly cutoff dependent, especially the first term.

To make sense of the result, compare this result to the same situation for a free electron. The shift in its energy is simple, since p is diagonal: it is the first term in Eq. 10.145.

$$\Delta E(free) = -\frac{2\alpha}{3\pi}\frac{p^2}{m^2}K. \tag{10.146}$$

Now pause to re-think the physics. We are asking "How is the mass of hydrogen different from $m_p + m_e - \langle V\rangle$?" where m_e is the energy of a free electron, and the answer is that it is $\Delta E = \Delta E_n - \Delta E(free)$. This quantity is

$$\Delta E = \frac{2\alpha}{3\pi}\sum_j (E_j - E_n)|\left\langle n|\frac{\vec{p}}{m}|j\right\rangle|^2 \log\frac{K}{E_j - E_n}. \tag{10.147}$$

Now for some physics. $E_j - E_n$ is a number in eV. Relativity is important at $k \sim m$, so our approximations break down there. Just set $K = m$ for the cutoff and wait for a (much harder) fully relativistic treatment of take care of the physics up there. (This took a couple of years to complete, but nowadays you can find it in books.) Next, replace the $E_j - E_n$ in the logarithm by an average $\langle E_j - E_n\rangle$ since logarithms don't depend strongly on their arguments. This gives

$$\Delta E = \frac{2\alpha}{3\pi m^2}\log\frac{m}{\langle E_j - E_n\rangle}\sum_j (E_j - E_n)|\langle j|\vec{p}|n\rangle|^2 \tag{10.148}$$

The sum is fun to do, reminiscent of the evaluation of the Thomas - Reiche - Kuhn sum rule in elementary quantum mechanics. It starts with

$$\sum_j (E_j - E_n) |\langle j|\vec{p}|n\rangle|^2 = \sum_j \langle n|\vec{p}H - H\vec{p}|j\rangle \langle j|\vec{p}|n\rangle . \tag{10.149}$$

Evaluating the commutator eliminates the j sum,

$$\begin{aligned} \sum_j \langle n|\vec{p}H - H\vec{p}|j\rangle \langle j|\vec{p}|n\rangle &= \left\langle n|\frac{1}{i}\vec{\nabla}V \cdot \frac{\vec{p}}{i}|n\right\rangle \\ &= -\int d^3x \psi_n^* \vec{\nabla}V \cdot \vec{\nabla}\psi_n \\ &= \frac{1}{2}\int d^3x (\nabla^2 V)|\psi_n|^2. \end{aligned} \tag{10.150}$$

In a Coulomb field

$$\nabla^2 V = 4\pi Z\alpha \delta^3(\vec{r}) \tag{10.151}$$

and the expression for the Lamb shift is

$$\Delta E = \frac{2\alpha}{3\pi}\frac{1}{m^2}\frac{1}{2} 4\pi Z\alpha [\log \frac{m}{\langle E_j - E_n\rangle}] |\psi(0)|^2. \tag{10.152}$$

Now we recall (or look up) that for S-wave states and in terms of the Bohr radius a_0

$$|\psi(0)|^2 = \frac{Z^3}{\pi n^3 a_0^3}. \tag{10.153}$$

The Rydberg (Ry) is $\frac{1}{2}\alpha^2 m = \frac{1}{2}\alpha/a_0$ so

$$\Delta E = \frac{8\alpha^3}{3\pi}\frac{Z^4}{n^3}\log\frac{m}{\langle E_j - E_n\rangle}\text{Ry}. \tag{10.154}$$

The quantity $\langle E_j - E_n\rangle$ is the average energy of excitation of an electron in the $2S$ state, which Bethe quotes as 17.8 eV. I don't know a clever way to get it. For the $2P$ state, $|\psi(0)|^2 = 0$, not so for the $1S$ state (so it is the state with the energy shift). Putting in numbers, we arrive at a shift of about 1040 MhZ.

I highly recommend reading Bethe's paper yourself!

Chapter 11

Closing thoughts

If you are reading this, then presumably the semester has been over for a while. Our study of quantum field theory is still incomplete, of course, but time is fleeting. There are several "foundational" topics I have left out. The most important of these are (in my opinion)

1. How to deal with fermions in path integrals. This involves strange objects called Grassman variables or anticommuting $c-$ numbers: $\psi_1\psi_2 = -\psi_2\psi_1$, $\psi_1\psi_1 = 0$.

2. Dimensional regularization for computing loop integrals. All physicists doing such calculations have a little table of "dim reg" integrals hidden away someplace to help them along.

3. The physics of spontaneous symmetry breaking and its major consequences, the presence of Goldstone bosons when a continuous global symmetry is spontaneously broken, the Higgs effect when a gauge symmetry is broken

There are a couple more topics that I would have liked to include, and which formed part of the syllabus when I occasionally taught a second semester quantum field theory class:

1. It would be nice to come to grips with the Standard Model at a moderately high level, including how renormalization works, the calculation of beta functions, and some phenomenology

2. More about critical phenomena in statistical systems, more about the renormalization group

Of course, my second set of topics reflects physics issues that I enjoy thinking about. Other people would probably come up with a different list. But that's not important. I think that if you have made it this far you are ready to identify your own set of interesting problems. While we really didn't look at anything which is currently an active area of research, you have some of the tools to begin to solve problems which involve quantum field theoretic issues in your own field of research. Enjoy!

Appendix A

Conventions for fields and scattering amplitudes

A.1 Kinematics conventions

To begin, I'll work in units where $\hbar = c = 1$. (Recall that $\hbar c = 200$ MeV-fm or 200 eV-nm for comversion to engineering units.) In these units [length] = 1/[energy] and a cross section σ which is an area is also an inverse energy-squared.

I'll work with "mostly minus" metric, so that momentum four vectors are

$$p^\mu = (E, \vec{p}); \qquad p_\mu = (E, -\vec{p}) \tag{A.1}$$

and the metric is

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{A.2}$$

so $p_\mu = g_{\mu\nu} p^\nu$. This is the standard convention in the quantum field theory or particle physics literature, as opposed to the "mostly plus" metric used in general relativity. Note that the energy-momentum relation for a free particle of mass m is

$$p^2 = p_\mu p^\mu = p^\mu p_\mu = E^2 - |\vec{p}|^2 = m^2. \tag{A.3}$$

A few other definitions:

$$\begin{aligned} x^\mu &= (t, \vec{x}) = (x_0, \vec{x}) \\ \partial_\mu &= (\frac{\partial}{\partial t}, \vec{\nabla}) \\ \partial^\mu &= (\frac{\partial}{\partial t}, -\vec{\nabla}) \\ \Box &= \partial^\mu \partial_\mu = \frac{\partial^2}{\partial t^2} - \nabla^2 \end{aligned} \tag{A.4}$$

and so on.

Typical formulas we'll encounter involve invariants such as $a \cdot b = a_\mu b^\mu$ or $\partial_\mu \phi \partial^\mu \phi$. A major annoyance in this subject is that many people (including me, sometimes) are sloppy with upper versus lower indices and put them anywhere they want. This is because, subconsciously, they know that every quantity they work with is actually going to be contracted to a scalar, $a \cdot b = a_0 b_0 - \vec{a} \cdot \vec{b}$, with the "known" minus sign for the three-vector part of the dot product.

Finally, there is something called "Euclidean space" for which $x^\mu = x_\mu = (x_0, x_1, x_2, x_3)$ and $a \cdot b = a_0 b_0 + a_1 b_a + a_2 b_2 + a_3 b_3$. Quantum field theory in Euclidean space looks very much like (or, is identcal to) statistical mechanics in four spatial directions.

A.2 Particles in a box

Begin with free particles in a box of volume $V = L^3$. A free particle amplitude is

$$\psi(x) = \langle x|k\rangle = \frac{1}{\sqrt{V}} e^{ikx} \tag{A.5}$$

The momentum in direction j is quantized as $k_j = 2\pi n_j/L$ where n_j is an integer for periodic boundary conditions. States are normalized so

$$\langle \vec{k}|\vec{k}'\rangle = \delta_{nn'} \tag{A.6}$$

and the completeness relation is

$$1 = \sum_n |k\rangle\langle k| = \int \frac{V}{(2\pi)^3} d^3k |\vec{k}\rangle\langle k|, \tag{A.7}$$

slightly abusing notation to treat k as continuous. (It is the usual $dn_x = L dk_x/(2\pi)$ story.) That is the phase space factor – the number of states between $\vec{k}$ and $\vec{k} + d\vec{k}$ is

$$dN = \frac{V}{(2\pi)^3} d^3k. \tag{A.8}$$

If $\psi(x)$ is a Schrödinger wave function, the probability density $\rho = \psi^*\psi = 1/V$. The current density is

$$\vec{J} = \frac{1}{2im} \psi^* \vec{\nabla}\psi - (\vec{\nabla}\psi)^*\psi = \frac{k}{m}\frac{1}{V} = \frac{\vec{v}}{V} \tag{A.9}$$

where $\vec{v}$ is the particle's velocity.

The Golden Rule begins by defining the transition probability per unit time $d\Gamma$ in terms of the T-matrix $\langle f|\mathcal{T}|i\rangle$ from a state $|i\rangle$ at time $t = 0$ to a state $|f\rangle$ at time T,

$$d\Gamma = \frac{1}{T} \lim_{T\to\infty} |\langle f(T)|\mathcal{T}|i(0)\rangle|^2. \tag{A.10}$$

Consider a process where $|i\rangle$ is a two particle state, $i = 1$ to 2, and $|f\rangle$ consists of n_f outgoing particles, $j = 1$ to n_f. Let's assemble all the pieces for a formula for the differential cross section.

First, in $d\Gamma$ there is a phase space factor for each outgoing particle,

$$\prod_{n=1}^{n_f} V\frac{d^3k_n}{(2\pi)^3}. \tag{A.11}$$

Next, in $\langle f(T)|\mathcal{T}|i(0)\rangle$ there will be a factor coming from the exponentials from the time dependence of the unperturbed states,

$$\int_0^T dt e^{i\Delta\omega t} = \frac{e^{i\Delta\omega T}-1}{i\Delta\omega} \tag{A.12}$$

where $\Delta\omega = \sum_{i=1}^{2} E_i - \sum_{j=1}^{n_f} E_f$. This will square (the usual squaring of the delta function story) to give the Golden Rule expression

$$d\Gamma = 2\pi\delta(\Delta\omega)|\langle f|\mathcal{T}|i\rangle|^2 \prod_{n=1}^{n_f}\frac{V}{(2\pi)^3}d^3k_n. \tag{A.13}$$

The cross section is defined as the ratio

$$d\sigma = d\Gamma \times \frac{1}{flux} \tag{A.14}$$

where the flux factor is $flux = |J| = v_{rel}/V$ and $v_{rel} = v_1 - v_2$.

Finally, it is often the case that momentum is conserved. Component by component the T-matrix will involve integrals of the form

$$\langle f|\mathcal{T}|i\rangle \propto \int_V d^3x e^{i\Delta kx} = V\delta_{\vec{n}_f,\vec{n}_i} = \prod_{i=x,y,z}\frac{e^{i\Delta k_i L}-1}{i\Delta k_i}, \tag{A.15}$$

where the first identity involves the integer-counting for the wave numbers, quantized in the box, and the second equality comes from taking k as continuous and doing the integral over the box. We must square the T-matrix and take the volume to be large. The mathematics is identical to what is encountered in squaring the frequency delta function and going to long times, Eq. A.12 from the time integral, to give Eq. A.13. Putting all the pieces together,

$$\begin{aligned} d\sigma &= \frac{V}{v_{rel}}|\langle f|\hat{T}|i\rangle|^2 \times (2\pi)^4\delta(\sum\omega_f - \sum\omega_i) \\ &\quad \times V\delta^3(\sum\vec{k}_f - \sum\vec{k}_i) \\ &\quad \times \prod_{n=1}^{n_f} V\frac{d^3k_n}{(2\pi)^3}. \end{aligned} \tag{A.16}$$

The leading V is from the flux and the V in the middle is from squaring the momentum conservation delta function. The reduced T-matrix $\langle f|\hat{T}|i\rangle$ is the full T-matrix $\langle f|\mathcal{T}|i\rangle$ with the overall momentum-conserving delta functions (pick the phrase you like best) snipped off, or divided out, or thrown away. Typically, it will involve an integral over all the relative coordinates, since overall space translation invariance is where overall momentum conservation comes from. It is proportional to $(1/\sqrt{V})^{n_f+2}$ because there are n_f+2 free particle wave functions in it, each carrying a factor of $1\sqrt{V}$ (recall Eq. A.5). And it gets squared. In the final expression, all the volume factors cancel; the final result for the cross section will be independent of the volume (as it should be).

A.3 Delta function normalization

Alternatively, we can use delta-function normalization for plane wave states in an infinite box. The normalization convention is

$$\langle\vec{k}|\vec{k}'\rangle = \delta^3(k-k') \tag{A.17}$$

so that

$$\int d^3k\langle\vec{k}|\vec{k}'\rangle = 1 \tag{A.18}$$

and the completeness relation is

$$1 = \int d^3k|k\rangle\langle k|. \tag{A.19}$$

It follows that the coordinate space wave function is

$$\langle x|k\rangle = \frac{1}{(2\pi)^{3/2}}e^{ikx}. \tag{A.20}$$

Now the phase space coounting is just

$$dN = d^3k. \tag{A.21}$$

There is a translation dictionary between the "box" T-matrix element and the "$\delta-$function" one, just because the definitions of the states, Eqs.A.5 and A.20, are different. Writing the dictionary in terms of volume-independent quantities, it is

$$[(2\pi)^3]^{(n_f+n_i)/2}\langle f|\mathcal{T}|i\rangle_{\delta-fn} = V^{(n_f+n_i)/2}\langle f|\mathcal{T}|i\rangle_{box}. \tag{A.22}$$

Either way this expression has no remaining 2π's in it. Generally, the term on each side of this formula will be proportional to $\delta^3(\sum k_i - \sum k_f)$.

To pass to the cross section, we have to square the matrix element. Unfortunately, doing this while preserving the delta functions (note we are really squaring a delta function this time) requires using wave packets. That is a little too much for me, but it is worked out in Peskin and Schroeder, Sec. 4.6. The final answer involves

$$\langle f|\mathcal{T}|i\rangle = \langle\{k_f\}|\mathcal{T}|\{k_i\}\rangle = (2\pi)^4\delta^4(\sum k_f - \sum k_i)\frac{\tilde{T}}{(2\pi)^3} \tag{A.23}$$

and

$$d\sigma = \frac{1}{v_{rel}}|\tilde{T}|^2(2\pi)^4\delta^4(\sum k_f - \sum k_i)\prod_{j=1}^{n_f} d^3k_j. \tag{A.24}$$

Note that $\tilde{T}$ has factors of 2π in it, from the normalization of the states. They can be pulled out into the phase space factor. This will give an expression identical to Eq. A.15, but without the volume factors. (They would have cancelled anyway.)

A.4 The normalization everybody uses

Finally, suppose we choose to normalize our states so that

$$\langle\vec{k}|\vec{k}'\rangle = C\delta^3(k-k') \tag{A.25}$$

where C is some convenient real constant. Then the normalization integral is

$$\int d^3k\langle\vec{k}|\vec{k}'\rangle = |C|^2 \tag{A.26}$$

and the coordinate space wave function is

$$\langle x|k\rangle = \sqrt{C}\frac{1}{(2\pi)^{3/2}}e^{ikx} \tag{A.27}$$

while the completeness relation is

$$1 = \int \frac{d^3k}{C}|k\rangle\langle k|. \tag{A.28}$$

The Schrödinger current (recall Eq. A.9) becomes

$$\vec{J} = \frac{\vec{k}}{m}C. \tag{A.29}$$

This is a nonrelativistic formula, but in general the current is just C times a current defined with some ordinary normalization, as in the earlier sections. The phase space factor is

$$dN = \frac{d^3p}{C} \tag{A.30}$$

and the cross section is

$$d\sigma_C = d\sigma_{C=1} \times \frac{1}{C^{2+n_f}}. \tag{A.31}$$

Notice that the T-matrix element will rescale–the cross section is physical and it can't change.

This convention seems stupid, but it is actually very clever. Why? It allows one to build in Lorentz invariance in an explicit way. The most commonly used choice in the literature is

$$C = 2E(k) \tag{A.32}$$

where $E(k) = \sqrt{k^2 + m^2}$ is the energy. That is, a momentum eigenstate is

$$|\vec{p}\rangle = \sqrt{2E(p)}a_p^\dagger|0\rangle. \tag{A.33}$$

(A quick check:

$$\langle\vec{q}|\vec{p}\rangle = \sqrt{2E(q)}\sqrt{2E(p)}\langle 0|a_q a_p^\dagger|0\rangle \tag{A.34}$$

and if we assume

$$[a_q, a_p^\dagger] = \delta^3(\vec{q} - \vec{p}), \tag{A.35}$$

the left hand side of Eq. A.34 is $2E_p\delta^3(\vec{q} - \vec{p})$; compare Eq. A.25.)

Why do this? It gives you several nice things:

First, the phase space factor is Lorentz invariant. It is

$$\frac{d^3p}{2E(p)(2\pi)^3} \tag{A.36}$$

per particle, evaluating the matrix element with pure exponentials and pulling out the 2π factors from the T-matrix into the phase space. Why is this Lorentz invariant? We can write it as

$$\frac{d^3p}{2E(p)} = d^3pdE\delta(E^2 - p^2 - m^2)\theta(E). \tag{A.37}$$

The $\theta(E)$ picks the positive energy solution to the delta-function ($E = \pm\sqrt{p^2 + m^2} = \pm E(p)$). The parts, $d^3pdE = d^4p_\mu$ and $\delta(p^2 + m^2)$ are each separately Lorentz invariant, so the product is, also. Now integrate E over the delta function, you get Eq. A.37.

Next, recall our definition of the field operator,

$$\phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(p)}}[a_p e^{ipx} + \dots]. \tag{A.38}$$

Contract it on a state $|k\rangle$:

$$\begin{aligned} \phi(x)|\vec{k}\rangle &= \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(p)}}[a_p e^{ipx} + \dots]\sqrt{2E(k)}a_k^\dagger|0\rangle \\ &= \frac{e^{ikx}}{(2\pi)^{3/2}}\frac{\sqrt{2\omega(k)}}{\sqrt{2\omega(k)}}|0\rangle. \end{aligned} \tag{A.39}$$

The algebra is simple, the $(2\pi)^{3/2}$ can be lumped in the phase space in Eq. A.36. This will give nice answers when you contract field operators against states.

For the Klein-Gordon equation, the continuity equation is

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0 \tag{A.40}$$

where

$$\rho = \phi^*\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t} \tag{A.41}$$

and

$$\vec{J} = \phi^*\vec{\nabla}\phi - \phi(\vec{\nabla}\phi^*). \tag{A.42}$$

We expect ρ and $\vec{J}$ form a four vector: Eq. A.40 is $\partial_\mu J^\mu = 0$ when written covariantly.

The state

$$\begin{aligned} \phi(x)\,|0\rangle &= \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(p)}}[a_p e^{ipx} + \dots]\,|0\rangle \\ &= \int \frac{d^3p}{2\omega(p)}\,|p\rangle \end{aligned} \tag{A.43}$$

(use Eqs. A.35 and A.33) is a linear superposition of single particle states. It's almost the same as the nonrelativistic convention, just with an extra $1/\omega(p) \sim 1/m$ in the nonrelativistic limit. We can define a plane wave solution

$$\begin{aligned} \Phi_k(x) &= \langle 0|\phi(x)|k\rangle \\ &= \exp(ikx). \end{aligned} \tag{A.44}$$

(Use Eq. A.39 and pull out the factors of 2π to go into the phase space integrals.) Obviously, this is "what we would expect" for the position space wave function for the single particle state $|p\rangle$.

Compute ρ and $\vec{J}$ using $\Phi_k(x)$. We find $\rho \propto E(p)$, $\vec{J} \propto \vec{p}$, that is, the probablilty current $(\rho, \vec{J}) \propto p^\mu$ obviously forms a four vector, because the components are of J^μ are proportional to p^μ.

Next, look ahead to consider Lorentz transformations. $p = (E, \vec{p})$ in one frame becomes $\Lambda p = (E', \vec{p}')$ in another frame. States must transform unitarily,

$$|\Lambda p\rangle = U(\Lambda)|p\rangle \tag{A.45}$$

and operators must transform as

$$O \to U(\Lambda) O U^{-1}(\Lambda). \tag{A.46}$$

$U(\Lambda)$ is some unitary transformation.

What is awkward is that $\delta^3(\vec{p}-\vec{q})$ is not Lorentz invariant. However, $E\delta^3(\vec{p}-\vec{q})$ is Lorentz invariant:

$$E\delta^3(\vec{p}-\vec{q}) = E'\delta^3(p'-q'). \tag{A.47}$$

This will mean that $\langle q|p\rangle = 2E(p)\delta^3(\vec{p}-\vec{q})$, the inner product of two states, will be Lorentz invariant. To check the Lorentz invariance, perform a Lorentz transformation along direction p_3. Change variables in the delta function, and use the chain rule,

$$\delta^3(p-q) = \delta^3(p'-q')\frac{dp_3'}{dp_3} \tag{A.48}$$

where $p_3' = \gamma(p_3 + \beta E)$ and $E' = \gamma(E + \beta p_3)$. Then

$$\delta^3(p-q) = \delta^3(p'-q')\gamma(1 + \beta\frac{\partial E}{\partial p_3}). \tag{A.49}$$

Since $E^2 = p_3^2 + \ldots$, this is

$$\delta^3(p-q) = \delta^3(p'-q')\frac{\gamma}{E}(E + \beta\frac{p_3}{E}E) = \delta^3(p'-q')\frac{E'}{E} \tag{A.50}$$

as required.

Creation operators are operators, and it happens that

$$U(\Lambda)a_p^\dagger U^{-1}(\Lambda) = \sqrt{\frac{E(\Lambda p)}{E(p)}}a_{\Lambda p}^\dagger. \tag{A.51}$$

This follows from

$$U(\Lambda)|\vec{p}\rangle = |\Lambda\vec{p}\rangle \tag{A.52}$$

or (insert $1 = U^{-1}(\Lambda)U(\Lambda)$)

$$U(\Lambda)a_p^\dagger\sqrt{E(p)}U^{-1}(\Lambda)U(\Lambda)|0\rangle = \sqrt{E(\Lambda p)}a_{\Lambda p}^\dagger|0\rangle. \tag{A.53}$$

This uses $U|0\rangle = |0\rangle$ – we are assuming that the vacuum is simple, the same in all frames. Anyway, now look at a field operator,

$$\phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(p)}}[a_p e^{ipx} + \dots] \tag{A.54}$$

Transform it: $x' = \Lambda x$, $p' = \Lambda p$, substitute in $a_{\Lambda p}$,

$$\phi(\Lambda x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(p)}}[\sqrt{\frac{\omega(p')}{\omega(p)}}a_{\Lambda p}e^{ik(\Lambda x)} + \dots] \tag{A.55}$$

Change variables from p to p', use $d^3p/(2\omega(p)) = d^3p'/(2\omega(p'))$, and you discover that

$$\phi(\Lambda x) = \int \frac{d^3p'}{(2\pi)^{3/2}\sqrt{2\omega(p')}}[a_{p'}e^{ip'x')} + \dots]. \tag{A.56}$$

Look at this closely, and look at Eq. A.54: it says that $\phi(x)$ really is a scalar operator. Under a Lorentz transformation, it remains unchanged. (Only its argument, x, changes.)

Box normalization is much easier to work with, but then $\phi(x)$ is not a true scalar field, in the sense of Lorentz transformations. Think about it, what is the box doing? Of course, if the original Lagrange density is Lorentz invariant, like

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - V(\phi), \tag{A.57}$$

all predictions of the theory will be Lorentz invariant. However, with a choice of normalization for states which does not know about Lorentz symmetry, the intermediate steps of a calculation will not look Lorentz invariant. The Lorentz invariance of the final result will seem to be a miracle, if you used noncovariant definitions for fields or states. It is better to have a formalism where everything you work with has simple, known transformation properties so that Lorentz invariance can be observed every step of the way. Miracles are not always good things!

Finally, in this convention, the differential cross section is

$$d\sigma = \frac{1}{4E_1E_2v_{rel}}|\tilde{T}|^2(2\pi)^4\delta^4(\sum k_f - \sum k_i)\prod_{j=1}^{n_f}\frac{d^3k_j}{(2\pi)^3 2E(k_j)} \tag{A.58}$$

and $\tilde{T}$ is the T-matrix with all extraneous factors of 2π, or $\sqrt{2E}$, and the overall momentum conserving delta function stripped off.

(Aside: Recall $v = k/E$, and with equal mass particles and in the center of mass frame, where the two incident particles have momenta $(E_{cm}/2, \pm p, 0, 0)$, $4E_1E_2v_{rel} = 4pE_{cm}$ which is $2E_{cm}^2$ or $2s$ in the extreme relativistic limit.)

Note: I seem to have introduced a Mandelstam variable. For two-to-two scattering, $p_1 + p_2 \to p_3 + p_4$, the Mandelstam invariants are

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_3 + p_4)^2 \\ t &= (p_1 - p_3)^2 = (p_4 - p_2)^2 \\ u &= (p_1 - p_4)^2 = (p_3 - p_2)^2. \end{aligned} \tag{A.59}$$

This discussion looks very robotic (and it is). But now you can use this set of conventions, and you don't need to think about where they came from. $\tilde{T}$ is often called the "invariant ampltude" M. It is a scalar function of the momenta and other four vectors (like polarizations) characterizing the scattering amplitude. Often, people try to construct a final answer for it in terms of invariants. Choosing the reference frame is often the last thing which is done.